



Published in final edited form as:

*N Engl J Med.* 2013 July 4; 369(1): 20–31. doi:10.1056/NEJMoa1215555.

## Urinary-Cell mRNA Profile and Acute Cellular Rejection in Kidney Allografts

Manikkam Suthanthiran, M.D., Joseph E. Schwartz, Ph.D., Ruchuang Ding, M.D., Michael Abecassis, M.D., Darshana Dadhania, M.D., Benjamin Samstein, M.D., Stuart J. Knechtle, M.D., John Friedewald, M.D., Yolanda T. Becker, M.D., Vijay K. Sharma, Ph.D., Nikki M. Williams, B.S., Christina S. Chang, B.S., Christine Hoang, B.S., Thangamani Muthukumar, M.D., Phyllis August, M.D., M.P.H., Karen S. Keslar, M.S., Robert L. Fairchild, Ph.D., Donald E. Hricik, M.D., Peter S. Heeger, M.D., Leiya Han, M.D., M.P.H., Jun Liu, Ph.D., Michael Riggs, Ph.D., M.P.H., David N. Ikle, Ph.D., Nancy D. Bridges, M.D., and Abraham Shaked, M.D., Ph.D. for the Clinical Trials in Organ Transplantation 04 (CTOT-04) Study

### Investigators

Weill Cornell Medical College (M.S., R.D., D.D., V.K.S., C.S.C., C.H., T.M., P.A.), Columbia University College of Physicians and Surgeons (B.S.), and Mount Sinai School of Medicine (P.S.H.), New York; Stony Brook University, Stony Brook, NY (J.E.S.); Northwestern University Feinberg School of Medicine (M.A., J.F.) and the University of Chicago (Y.T.B.), Chicago; Emory University, Atlanta (S.J.K.); National Institute of Allergy and Infectious Diseases, Bethesda, MD (N.M.W., N.D.B.); Cleveland Clinic (K.S.K., R.L.F.) and University Hospitals Case Medical Center (D.E.H.), Cleveland; Pharmaceutical Product Development, Wilmington (L.H., J.L.), and Rho Federal Systems, Chapel Hill (M.R., D.N.I.) — both in North Carolina; and the University of Pennsylvania School of Medicine, Philadelphia (A.S.)

### Abstract

**Background**—The standard test for the diagnosis of acute rejection in kidney transplants is the renal biopsy. Noninvasive tests would be preferable.

**Methods**—We prospectively collected 4300 urine specimens from 485 kidney-graft recipients from day 3 through month 12 after transplantation. Messenger RNA (mRNA) levels were measured in urinary cells and correlated with allograft-rejection status with the use of logistic regression.

**Results**—A three-gene signature of 18S ribosomal (rRNA)—normalized measures of CD3 mRNA and interferon-inducible protein 10 (IP-10) mRNA, and 18S rRNA discriminated between biopsy specimens showing acute cellular rejection and those not showing rejection (area under the curve [AUC], 0.85; 95% confidence interval [CI], 0.78 to 0.91;  $P < 0.001$  by receiver-operating-characteristic curve analysis). The cross-validation estimate of the AUC was 0.83 by bootstrap resampling, and the Hosmer–Lemeshow test indicated good fit ( $P = 0.77$ ). In an external-validation data set, the AUC was 0.74 (95% CI, 0.61 to 0.86;  $P < 0.001$ ) and did not differ significantly from the AUC in our primary data set ( $P = 0.13$ ). The signature distinguished acute cellular rejection from acute antibody-mediated rejection and borderline rejection (AUC, 0.78; 95% CI, 0.68 to 0.89;  $P < 0.001$ ). It also distinguished patients who received anti-interleukin-2 receptor antibodies from those who received T-cell-depleting antibodies ( $P < 0.001$ ) and was

Copyright © 2013 Massachusetts Medical Society.

Address reprint requests to Dr. Suthanthiran at Hamad bin Khalifa Medical Research Bldg., Division of Nephrology and Hypertension, Department of Medicine and Transplantation Medicine, New York Presbyterian Hospital–Weill Cornell Medical Center, 525 E. 68th St., Box 3, New York, NY 10065, or at msuthan@med.cornell.edu.

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

diagnostic of acute cellular rejection in both groups. Urinary tract infection did not affect the signature ( $P = 0.69$ ). The average trajectory of the signature in repeated urine samples remained below the diagnostic threshold for acute cellular rejection in the group of patients with no rejection, but in the group with rejection, there was a sharp rise during the weeks before the biopsy showing rejection ( $P < 0.001$ ).

**Conclusions**—A molecular signature of CD3 mRNA, IP-10 mRNA, and 18S rRNA levels in urinary cells appears to be diagnostic and prognostic of acute cellular rejection in kidney allografts. (Funded by the National Institutes of Health and others.)

Kidney Transplantation is considered the best available treatment for patients with end-stage renal disease (ESRD), but acute rejection, a leading cause of new cases of ESRD, undermines its full benefits.<sup>1-3</sup> Acute rejection is diagnosed by means of needle biopsy. Over time, this invasive procedure has become safer, and biopsy interpretation more standardized.<sup>4</sup> Nevertheless, bleeding and subsequent graft loss still occur, and sampling errors and inter-observer variability in biopsy reading remain problematic.<sup>5</sup> Repeated biopsies to assess the recipient's status pose challenges, including feasibility and cost.

Immunosuppressive drugs effectively treat acute rejection; a noninvasive means of diagnosing this reversible cause of graft failure would be advantageous. Furthermore, noninvasive screening that foretells acute rejection before loss of kidney function is clinically detectable might reduce rejection-associated graft damage, and ongoing characterization of the immune status could help minimize the effects of insufficient or excess immunosuppression.

The multicenter Clinical Trials in Organ Transplantation 04 (CTOT-04) study was designed to investigate whether urinary-cell levels of messenger RNA (mRNA) encoding for the CD3 chain, perforin, granzyme B, proteinase inhibitor 9, CD103, interferon-inducible protein 10 (IP-10), and the chemokine receptor CXCR3, ascertained at the time of biopsy, are diagnostic of acute rejection and to determine whether mRNA profiles of sequential urine specimens obtained at clinically stable time points predict the development of acute rejection. The rationale for the present study was provided by the immunobiology of the proteins encoded by these mRNAs and by data from single-center studies suggesting that measurement of these mRNA levels in urine offers a noninvasive means of diagnosing acute rejection of kidney allografts.<sup>6</sup>

## Methods

### Trial Design

In the prospective observational CTOT-04 study, sponsored by the National Institutes of Health (NIH), we enrolled 497 patients selected for kidney transplantation at five clinical sites. A total of 492 patients received a kidney transplant, and 4300 urine specimens were collected from 485 patients for urinary-cell mRNA profiling (Fig. 1) (see the Supplementary Appendix, available with the full text of this article at NEJM.org, for patient-enrollment information and recipient and donor characteristics).

The institutional review board at each site approved the study, and all the patients provided written informed consent. An NIH-sponsored statistical analysis and clinical coordinating center was responsible for data management and data analysis.

### Urine Samples and mRNA measurement

After transplantation, urine was collected on days 3, 7, 15, and 30 and in months 2, 3, 4, 5, 6, 9, and 12; as well as at the time of each kidney-allograft biopsy and 2 weeks thereafter.

Urine-cell pellets were prepared at the clinical sites, stored at  $-80^{\circ}\text{C}$ , and shipped to the Gene Expression Monitoring (GEM) Core at Weill Cornell Medical College, New York.

The staff at GEM Core isolated RNA from the pellets and assessed RNA quantity and purity (Table S3 in the Supplementary Appendix). Absolute levels of the mRNAs prespecified in the study protocol (CD3, perforin, granzyme B, proteinase inhibitor 9, CD103, IP-10, CXCR3, and transforming growth factor  $\beta$  1 [TGF- $\beta$  1]) and 18S ribosomal RNA (rRNA) were quantified in preamplification-enhanced real-time quantitative polymerase-chain-reaction (PCR) assays with the use of oligonucleotide primers and TaqMan probes (Table S4 in the Supplementary Appendix) designed by the GEM Core, and the results (mRNA copies per microgram of total RNA and 18S rRNA copies [ $\times 10^{-6}$ ] per microgram of total RNA) were reported to the statistical analysis and clinical coordinating center. The staff members at GEM Core were unaware of the clinical information, including the results of kidney-allograft biopsies, before transfer of the mRNA data set to the statistical analysis and clinical coordinating center.

Urine specimens were classified as passing quality control if the 18S rRNA copy number was greater than or equal to  $5 \times 10^7$  per microgram of total RNA isolated from the urine pellet and if the TGF- $\beta$  1 mRNA copy number was greater than or equal to 100 copies per microgram of total RNA isolated from the urine pellet. If either threshold was not met, the specimen was classified as failing quality control.

### **Allograft Biopsy Specimens and Matched Urine Specimens**

A total of 410 of 423 biopsies performed in 220 patients yielded specimens that were adequate for evaluation. Among these biopsies, 321 were performed because of clinical signs of rejection and 89 were surveillance biopsies. The on-site pathologist used the Banff schema<sup>7</sup> for classification of the biopsy specimens, and results were recorded with the use of a form supplied by the statistical analysis and clinical coordinating center (Table S5 in the Supplementary Appendix). Figure 1 shows the findings for the 298 biopsy specimens that had matched urine samples (urine collected from 3 days before to 1 day after the biopsy). The Supplementary Appendix lists the findings for all 410 biopsy specimens that were adequate for biopsy-based diagnosis.

Patients who did not undergo biopsy were classified as having stable graft function if the average serum creatinine level was less than or equal to 2.0 mg per deciliter (180  $\mu\text{mol}$  per liter) for available assessments 6, 9, and 12 months after transplantation, with no graft loss or death during the first 12 months after transplantation, no treatment for acute rejection, and no evidence of cytomegalovirus (CMV) or polyomavirus type BK (BKV) infection.

### **External-Validation Data Set**

Matched urine samples for all 24 biopsy specimens showing acute cellular rejection according to the Banff schema and a random selection of 47 biopsy specimens showing no evidence of rejection, obtained from 64 kidney-graft recipients who had been enrolled in the NIH-sponsored Clinical Trials in Organ Transplantation 01 (CTOT-01) study, were used by the statistical analysis and clinical coordinating center to construct an external-validation data set. Characteristics of the 64 patients (Table S2B in the Supplementary Appendix) and details of the external-validation study are provided in the Supplementary Appendix.

### **Statistical Analysis**

Each mRNA measure was analyzed before and after normalization with 18S rRNA copies ( $\times 10^{-6}$ ) per microgram of total RNA and then  $\log_{10}$ -transformed to reduce positive

skewness. Kruskal–Wallis and Mann–Whitney tests were used to compare levels across diagnoses.

Logistic regression was used to identify parsimonious subsets of the eight 18S-normalized mRNA measures — CD3, perforin, granzyme B, proteinase inhibitor 9, CD103, IP-10, CXCR3, and TGF- $\beta$ 1 — and 18S rRNA from matched urine samples that discriminated between biopsy specimens showing acute cellular rejection and those showing no rejection. Acute cellular rejection was defined as rejection of Banff grade IA or higher, and biopsy specimens showing no rejection were those classified by the on-site pathologist as showing no histologic features of rejection (Table S5 in the Supplementary Appendix). From those models in which each predictor was significant at a P value of less than 0.05, we provisionally selected the one with the greatest log-likelihood ratio and the greatest area under the receiver-operating-characteristic (ROC) curve as the best-fitting model.<sup>8</sup> Regression estimates from this model defined a diagnostic signature, and we used the area under the curve (AUC), sensitivity, and specificity to evaluate the ability of this signature to discriminate between biopsy specimens showing acute cellular rejection and those showing no rejection.

Generalizability of the fitted model to other data sets was evaluated with the use of bootstrap-resampling methods.<sup>9</sup> Logistic regression with backward elimination was used to identify the best subset of the 18S-normalized mRNA measures and the 18S rRNA measure in each of 500 data sets obtained by sampling with replacement from the original data set.<sup>10</sup> The best subset model was then fit to 500 additional bootstrap samples from which cross-validated<sup>9</sup> measures of discrimination (i.e., the AUC) and model fit (i.e., calibration-curve intercept and slope)<sup>8</sup> and a locally estimated scatterplot-smoothed (loess) calibration plot<sup>9</sup> were obtained.

We compared the prospective trajectories of the diagnostic signature in three groups. For patients who underwent a biopsy showing acute cellular rejection and who had had no prior biopsy showing borderline changes or antibody-mediated rejection, we included all quality-control–passed urine samples collected from the time of transplantation until 4 days before the first biopsy showing acute cellular rejection or until 400 days after transplantation, whichever came first. For patients who underwent biopsy but never had a biopsy finding that showed rejection, we included all quality-control–passed urine samples collected during the first 400 days after transplantation. The third group included patients who never underwent biopsy and who satisfied the criteria for stable graft function. For each group, we plotted the results of a loess model predicting the diagnostic signature from the days after transplantation, with the patient as a covariate.

We used the same methodologic approach to compare the retrospective trajectories of the diagnostic signature, looking backward from the time of biopsy in two groups — the group with biopsy findings showing acute cellular rejection and the group with biopsy findings showing no rejection. All analyses were performed with the use of SAS software, version 9.3 (SAS Institute), or RMS software, version 2.12.2 (<http://cran.r-project.org/web/packages/rms>).

## Results

### Patients and Samples

A total of 4300 urine samples were collected from 485 patients for urinary-cell mRNA profiling; 3559 samples passed quality control and 741 did not. A total of 220 patients underwent 410 kidney-allograft biopsies, including 321 biopsies performed at the five participating sites because of clinical signs of rejection (for-cause biopsies). Of 89

surveillance biopsies, 88 were performed at Northwestern University Feinberg School of Medicine, Chicago, where surveillance biopsies were part of the standard of care. Many patients had more than one biopsy performed, often with different diagnoses (Fig. 1).

Among the 265 patients who did not undergo biopsy, 202 met the criteria for stable graft function, of whom 201 had urine specimens that passed quality control. A total of 63 patients did not meet the criteria for stable graft function: 47 patients had no data on serum creatinine levels beyond 5 months after transplantation; 4 had serum creatinine levels at 6, 9, and 12 months that averaged more than 2.0 mg per deciliter; 9 were treated for CMV infection, BKV infection, or both; 1 had graft loss within the first 12 months after transplantation; and 2 died within the first 12 months after transplantation.

### mRNA Levels in Urinary Cells

We investigated whether urinary-cell levels of mRNA were diagnostic of acute cellular rejection according to the Banff schema. We compared mRNA levels in 43 urine samples that were matched to 43 biopsy specimens showing acute cellular rejection (38 for-cause and 5 surveillance biopsy specimens from 34 patients) with 163 urine samples matched to 163 biopsy specimens not showing rejection (107 for-cause and 56 surveillance biopsy specimens from 126 patients) and with 1540 urine samples from 201 patients with stable graft function who did not undergo biopsy. To determine whether a urine sample obtained at the time of biopsy was diagnostic of acute cellular rejection, only matched urine samples were included for patients with biopsy specimens, whereas all urine samples that passed quality control were included for patients with stable graft function and no biopsy.

Figure 2 shows that 18S-normalized levels of mRNA for CD3, perforin, granzyme B, and IP-10 in urinary cells differed significantly among the three study groups ( $P < 0.001$  for each mRNA). Pairwise group comparisons showed that the levels of mRNA for CD3, perforin, granzyme B, and IP-10 in patients with biopsy specimens showing acute cellular rejection were higher than in those with biopsy specimens showing no rejection ( $P < 0.001$  for each mRNA) and in those with stable graft function who did not undergo biopsy ( $P < 0.001$  for each mRNA).

Levels of 18S rRNA were also higher in patients with biopsy specimens showing acute cellular rejection than in those who had biopsy specimens showing no rejection ( $P < 0.001$ ) and those who had stable graft function and did not undergo biopsy ( $P < 0.001$ ) (Table S6A in the Supplementary Appendix). Although nonnormalized levels of the mRNAs for CD103, CXCR3, TGF- $\beta$  1, and proteinase inhibitor 9 were significantly associated with a diagnosis of acute cellular rejection (Table S6B in the Supplementary Appendix), these associations became nonsignificant after normalization by 18S rRNA.

### Development of a Three-Gene Diagnostic Signature

**Fitting the Model**—A three-gene model of 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA (all log-transformed values) was the best-fitting parsimonious model, yielding the following diagnostic signature:

$$-6.1487 + 0.8534 \log_{10}(\text{CD3}\epsilon / 18\text{S}) + 0.6376 \log_{10}(\text{IP-10} / 18\text{S}) + 1.6464 \log_{10}(18\text{S})$$

where the units of measurement in the PCR assays for CD3 mRNA and IP-10 mRNA were copy number per microgram of total RNA, and the units for 18S rRNA were number of copies ( $\times 10^{-6}$ ) per microgram of total RNA. In the equation,  $-6.1487$  was the intercept, and  $0.8534$ ,  $0.6376$ , and  $1.6464$  were the slopes (coefficients), respectively, for the  $\log_{10}(\text{CD3} / 18\text{S})$ ,  $\log_{10}(\text{IP10} / 18\text{S})$ , and  $\log_{10}(18\text{S})$  values in the best-fitting logistic-regression model.

The intercept and slopes have no intrinsic units of measurement. A diagnostic score of  $-1.213$ , with the use of this equation, was the cutoff point that maximized the combined sensitivity and specificity (Youden's index<sup>11</sup>) of the signature to discriminate between biopsy specimens showing acute cellular rejection and those showing no rejection.

ROC curve analysis showed that this three-gene signature yielded an AUC of 0.85 (95% confidence interval [CI], 0.78 to 0.91;  $P < 0.001$ ). With the use of the cutoff point of  $-1.213$ , this diagnostic signature has 79% sensitivity (95% CI, 67 to 91) and 78% specificity (95% CI, 71 to 84) to discriminate between biopsy specimens showing acute cellular rejection and those showing no rejection (Fig. 3A). The Hosmer–Lemeshow test<sup>12</sup> indicated an excellent fit of this model to the data (chi-square = 4.84 with 8 df,  $P = 0.77$ ). The three-gene signature also discriminated between the group of patients with biopsy specimens showing acute cellular rejection and the group of patients with stable graft function who did not undergo biopsy (Fig. 3B).

**Bootstrapped Model Selection and Internal Validation**—Bootstrap validation of this three-gene model yielded a cross-validated estimate of the AUC of 0.83, which is an estimate of the expected value of the AUC in independent samples (i.e., samples not used to derive the diagnostic signature). The calibration-curve intercept and slope of  $-0.06$  and  $0.92$ , respectively, revealed that the predicted probabilities of a biopsy showing acute cellular rejection, across the range of the diagnostic signature, tended to be only very slightly higher than the actual probabilities (Fig. 3C) and that the likelihood that the model was overfitted was small. The loess-smoothed estimates of the unadjusted and cross-validated calibration curves were overlaid on a diagonal reference line representing perfect model calibration (Fig. 3C). The close correspondence of the two curves to the reference line shows excellent fit and reflects the above interpretation of the intercept and slope estimates of the calibration curve.

**External Validation**—Among the 71 complementary DNA (cDNA) samples assayed for the levels of transcripts included in the diagnostic signature, 24 samples (from 21 patients) showed acute cellular rejection and 47 samples (from 43 patients) showed no rejection. Of the 24 specimens showing acute cellular rejection (from 17 for-cause and 7 surveillance biopsies), 11 were classified as Banff grade IA, 4 as grade IB, 7 as grade IIA, 1 as grade IIB, and 1 as grade III. Of the 47 biopsy specimens not showing acute cellular rejection, 19 were from for-cause biopsies and 28 from surveillance biopsies. Among the 71 specimens constituting the external-validation data set, the 36 for-cause biopsy specimens were from seven of the eight clinical sites that participated in the CTOT-01 study, and the 35 surveillance biopsy specimens were from five of these eight sites.

The ROC curve of the three-gene signature discriminating between biopsy specimens showing acute cellular rejection and those showing no rejection in the external-validation data set had an AUC of 0.74 (95% CI, 0.61 to 0.86;  $P < 0.001$ ) (Fig. 3D). This AUC was lower than the AUC of 0.85 (95% CI, 0.78 to 0.91;  $P < 0.001$ ) in the CTOT-04 data set (Fig. 3A), but the difference between the two AUCs was not significant ( $P = 0.13$ ). With the use of the cutoff value of  $-1.213$  for the diagnostic signature in the CTOT-04 study, acute cellular rejection in the external-validation data set was predicted with a specificity of 72% (95% CI, 62 to 83) and a sensitivity of 71% (95% CI, 53 to 89), values that were also lower than those in the CTOT-04 data set but not significantly ( $P > 0.35$  for both comparisons).

### Prospective Trajectory of Diagnostic Signature

Figure S2 in the Supplementary Appendix displays the loess-smoothed, average within-person prospective trajectories of the diagnostic signature (i.e., trajectories of the signature

as a function of the time since transplantation), with 95% confidence intervals, in the three groups of patients. The trajectories for the group of patients with specimens showing no rejection and for the group of patients with stable graft function and no biopsy were flat and remained below the  $-1.213$  threshold that was diagnostic of acute cellular rejection throughout the first 400 days after transplantation. However, a progressive increase in the diagnostic score was seen in the urine samples from patients in whom acute cellular rejection developed.

Thus, even after the exclusion of all urine samples obtained after the development of acute cellular rejection and those that were matched to biopsy specimens showing acute cellular rejection and had been used to develop the diagnostic signature, there was a clear signal by approximately 80 days after transplantation that values were elevated in patients in whom acute cellular rejection subsequently developed. After approximately 160 days, the average value for the patients in whom acute cellular rejection subsequently developed was greater than or equal to the threshold level that was diagnostic for acute cellular rejection (Fig. S2D in the Supplementary Appendix).

### Retrospective Trajectory of Diagnostic Signature

Figure 4 shows the loess-smoothed, average within-person retrospective trajectories of the diagnostic signature (i.e., trajectories of the signature as a function of the time before biopsy), with 95% confidence intervals, for the group of patients with biopsy specimens showing acute cellular rejection (Fig. 4A) and the group with specimens showing no rejection (Fig. 4B). There was a significant difference between the trajectories for the two groups, with the signature remaining flat and well below the diagnostic threshold during the 270 days preceding a biopsy in the group with specimens showing no rejection, whereas a marked increase was observed in the diagnostic signature during the 20-day period leading up to the first specimen showing acute cellular rejection (Fig. 4C and 4D) ( $P < 0.001$ ).

### Additional Features of the Diagnostic Signature

When urine samples matched to biopsy specimens showing acute cellular rejection were compared with urine samples matched to biopsy specimens showing borderline changes, acute antibody-mediated rejection, or chronic allograft nephropathy, the signature was diagnostic of acute cellular rejection with 71% specificity (95% CI, 55 to 87) and 79% sensitivity (95% CI, 67 to 91) (AUC, 0.78; 95% CI, 0.68 to 0.89;  $P < 0.001$ ). The signature was diagnostic in patients who underwent for-cause biopsies as well as in those who underwent surveillance biopsies and was similarly diagnostic across transplantation sites (the site-by-signature interaction was not significant,  $P = 0.30$ ). The score of the diagnostic signature decreased after antirejection therapy for acute cellular rejection ( $P = 0.05$ ), but it was not associated with the Banff grade for acute cellular rejection ( $P = 0.83$ ) (Table S7 in the Supplementary Appendix). The urine samples from patients who received induction therapy with interleukin-2 receptor antibodies, as compared with those from patients who received T-cell-depleting antibodies, had a higher diagnostic score ( $P < 0.001$ ) (Fig. S3 in the Supplementary Appendix), especially during the first month after transplantation, but the signature was diagnostic of acute cellular rejection with either type of induction therapy.

The diagnostic signature was not associated with urinary tract infection ( $P = 0.69$ ), blood infection ( $P = 0.94$ ), or CMV infection ( $P = 0.56$ ) but was associated with BKV infection ( $P = 0.03$ ). The mean diagnostic score at 4 to 6 months was associated with a decline of 30% or more in renal-allograft function from 6 to 12 months (odds ratio, 2.66; 95% CI, 1.45 to 4.87;  $P = 0.002$ ). Details of these and additional features of the signature are provided in the Supplementary Appendix.

## Discussion

Our data indicate that acute cellular rejection, a treatable cause of kidney-allograft failure, can be diagnosed noninvasively and accurately with the use of a three-gene signature. Furthermore, this parsimonious diagnostic signature measured in urine specimens obtained longitudinally from patients with normal allograft biopsies and from patients with clinically stable graft function was relatively flat and distinct from the progressive increase observed in specimens from patients in whom biopsy-confirmed acute cellular rejection developed later.

The refinements in the PCR assay — preamplification of cDNA and absolute quantification of mRNA copies — enabled precise measurement of multiple mRNAs. The results of this study, which show that increased levels of mRNA for CD3, perforin, granzyme B, proteinase inhibitor 9, CD103, IP-10, and CXCR3 are associated with acute cellular rejection, confirm, extend, and independently validate data from previous single-center studies showing the diagnostic usefulness of measuring these mRNAs.<sup>6</sup> Also, the elevated levels of these mRNAs are consistent with the role of cytotoxic T cells<sup>13</sup> and chemokines<sup>14</sup> in allograft rejection.

Our strategy of absolute quantification, rather than relative quantification with the use of the comparative  $C_T$  method,<sup>15</sup> showed that the level of 18S rRNA, an integral component of translational machinery for protein synthesis,<sup>16</sup> was higher in patients with biopsy specimens showing acute cellular rejection than in those with specimens showing no rejection and in those who had stable graft function and did not undergo biopsy. Heightened levels of 18S rRNA in patients with acute cellular rejection may reflect the activated state of the T cells mediating rejection and the proportion and state of differentiation of the cells (e.g., highly differentiated renal tubular epithelial cells vs. activated lymphocytes) contributing to the urine-cell pellet.

Robust yardsticks for measuring the immune status of the transplant recipient have not been established. The relatively flat trajectory of the diagnostic signature in patients in whom acute cellular rejection did not develop, in contrast to the increasing trajectory in those in whom acute cellular rejection developed, is a potential tool for monitoring immune status and, ultimately, for adjusting immunosuppressive therapy according to immune status. The finding that the three-gene signature may reflect the potency of immunosuppressive therapy offers opportunities for an immune-surveillance tool for monitoring the patient after transplantation, with the levels reflecting the potency of immunosuppressive regimens and a marked rise in the mRNA levels observed in the weeks before the biopsy showing acute cellular rejection serving as potential triggers for preemptive antirejection therapy.

Although acute cellular rejection is frequently treatable, it is a precursor of chronic rejection and graft loss.<sup>17</sup> Preventive strategies include immunosuppression, initiated at the time of transplantation with adjustments in medications made on the basis of drug levels, drug toxicity, and clinical events (e.g., infection). The marked increase in the trajectory of the diagnostic signature in the weeks preceding acute cellular rejection, in addition to foreshadowing the development of rejection, may offer opportunities to test this approach for preemptive therapy, before irreversible tissue damage occurs.

There are limitations to our study. Biopsy-matched urine samples were not collected for 112 of the 410 biopsy specimens, and 54 of the 298 biopsy-matched urine samples did not pass quality-control thresholds. The number of patients with antibody-mediated rejection was small, which prevented in-depth evaluation of the usefulness of urinary-cell mRNA profiling for diagnosing antibody-mediated rejection. However, the signature distinguished acute cellular rejection from antibody-mediated rejection, borderline changes, and other changes.



In conclusion, the diagnostic signature calculated from the mRNA levels of genes relevant to acute cellular rejection described in this study may provide a direct measure of risk (the predicted probability that a biopsy would reveal acute cellular rejection) and a means of assessing immune status with repeated assessments.<sup>12,18</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

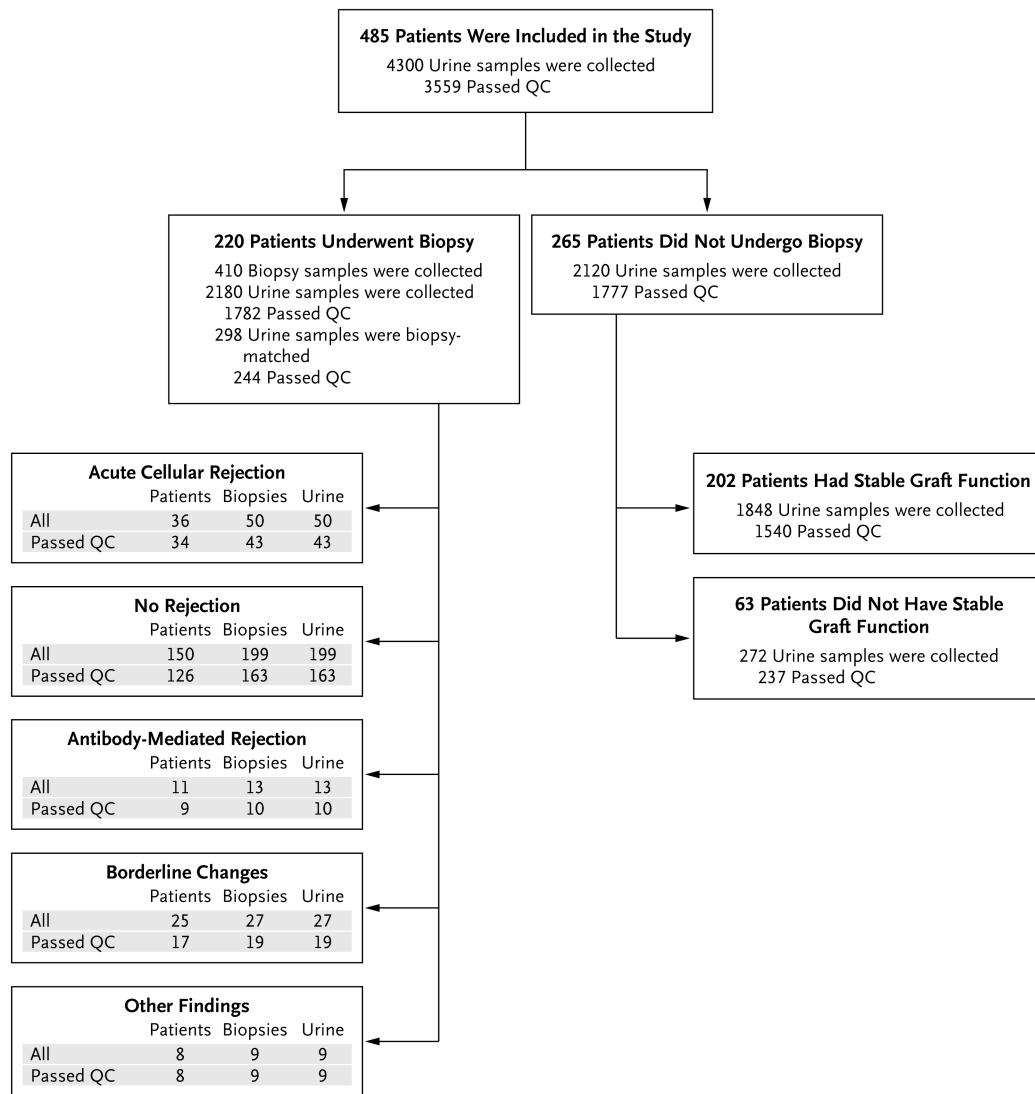
Supported in part by grants from the National Institutes of Health (U01AI63589 and R37AI051652) and the Qatar National Research Foundation (NPRP 08-503-3-111) and by a Clinical and Translational Science Center Award (UL1TR000457, to Weill Cornell Medical College).

We thank Ms. Jane I. Charette (Northwestern University), Ms. Debra McCorristan (University of Pennsylvania), Dr. Amy Sundberg (University of Wisconsin Hospitals and Clinics), Mr. Jonathan Kim (New York Presbyterian Hospital–Columbia University Medical Center), and Ms. Catherine Snopkowski and Ms. Tessa Scott (New York Presbyterian Hospital–Weill Cornell Medical Center) for contributions to the execution of the study; and CTOT-01 investigators Drs. Peter Nickerson, David Rush, and Ian Gibson (University of Manitoba), Dr. Richard N. Formica (Yale University School of Medicine), Dr. Emilio Poggio (Cleveland Clinic), Dr. Kenneth A. Newell (Emory University School of Medicine), and Dr. Jens Goebel (Cincinnati Children's Hospital) for contributions that enabled the performance of the external validation of the diagnostic signature developed in this study.

## References

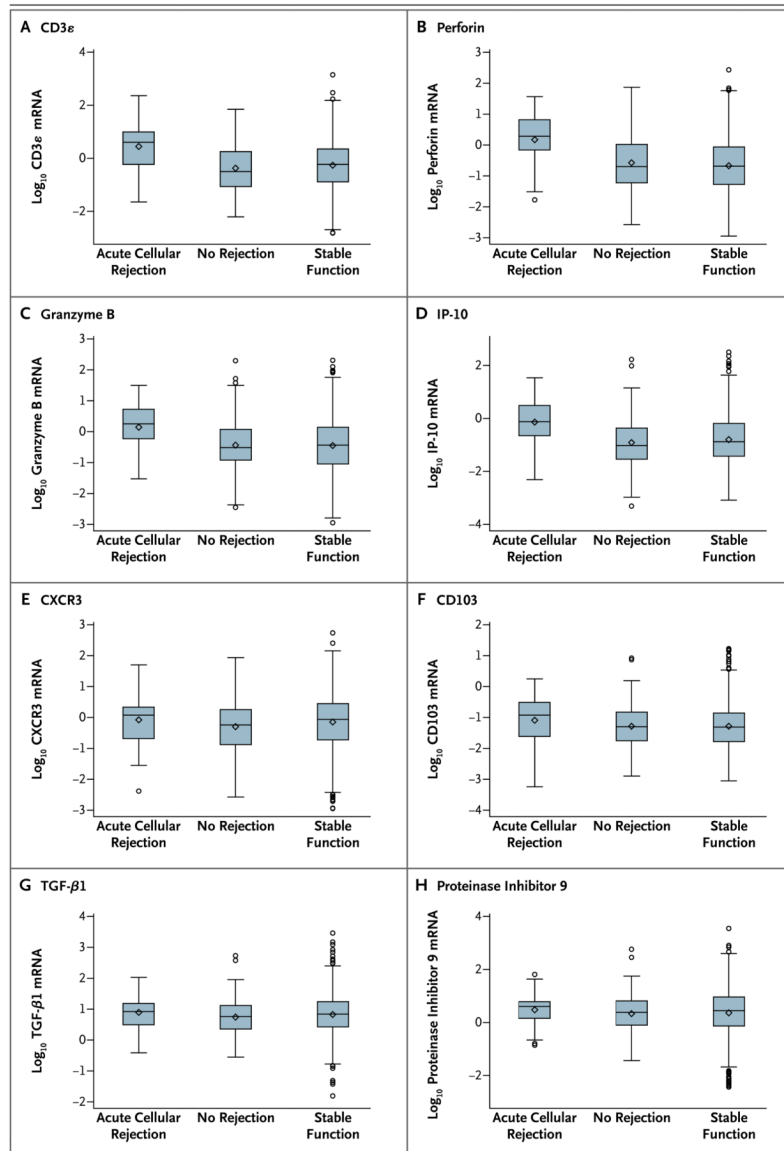
- Hariharan S, Johnson CP, Bresnahan BA, Taranto SE, McIntosh MJ, Stablein D. Improved graft survival after renal transplantation in the United States, 1988 to 1996. *N Engl J Med*. 2000; 342:605–12. [PubMed: 10699159]
- Tonelli M, Wiebe N, Knoll G, et al. Systematic review: kidney transplantation compared with dialysis in clinically relevant outcomes. *Am J Transplant*. 2011; 11:2093–109. [PubMed: 21883901]
- Langone AJ, Chuang P. The management of the failed renal allograft: an enigma with potential consequences. *Semin Dial*. 2005; 18(3):185–7. [PubMed: 15934959]
- Williams WW, Taheri D, Tolkoff-Rubin N, Colvin RB. Clinical role of the renal transplant biopsy. *Nat Rev Nephrol*. 2012; 8:110–21. [PubMed: 22231130]
- Furness PN, Taub N. International variation in the interpretation of renal transplant biopsies: report of the CERTAPAP Project. *Kidney Int*. 2001; 60:1998–2012. Erratum, *Kidney Int* 2001; 60:2429. [PubMed: 11703620]
- Anglicheau D, Suthanthiran M. Noninvasive prediction of organ graft rejection and outcome using gene expression patterns. *Transplantation*. 2008; 86:192–9. [PubMed: 18645476]
- Racusen LC, Colvin RB, Solez K, et al. Antibody-mediated rejection criteria — an addition to the Banff 97 classification of renal allograft rejection. *Am J Transplant*. 2003; 3:708–14. [PubMed: 12780562]
- Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010; 21:128–38. [PubMed: 20010215]
- Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15:361–87. [PubMed: 8668867]
- Austin P, Tu J. Bootstrap methods for developing predictive models. *Am Stat*. 2004; 58:131–7.
- Le CT. A solution for the most basic optimization problem associated with an ROC curve. *Stat Methods Med Res*. 2006; 15:571–84. [PubMed: 17260924]
- Tripepi G, Jager KJ, Dekker FW, Zoccali C. Statistical methods for the assessment of prognostic biomarkers (part II): calibration and re-classification. *Nephrol Dial Transplant*. 2010; 25:1402–5. [PubMed: 20167948]

13. Strom TB, Tilney NL, Carpenter CB, Busch GJ. Identity and cytotoxic capacity of cells infiltrating renal allografts. *N Engl J Med.* 1975; 292:1257–63. [PubMed: 1093024]
14. Schenk AD, Rosenblum JM, Fairchild RL. Chemokine-directed strategies to attenuate allograft rejection. *Clin Lab Med.* 2008; 28:441–54. [PubMed: 19028262]
15. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.* 2008; 3:1101–8. [PubMed: 18546601]
16. Lodish HF. Translational control of protein synthesis. *Annu Rev Biochem.* 1976; 45:39–72. [PubMed: 786155]
17. Almond PS, Matas A, Gillingham K, et al. Risk factors for chronic rejection in renal allograft recipients. *Transplantation.* 1993; 55:752–6. [PubMed: 8475548]
18. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007; 115:928–35. [PubMed: 17309939]



**Figure 1. Patients, Biopsy Results, and Urine Samples**

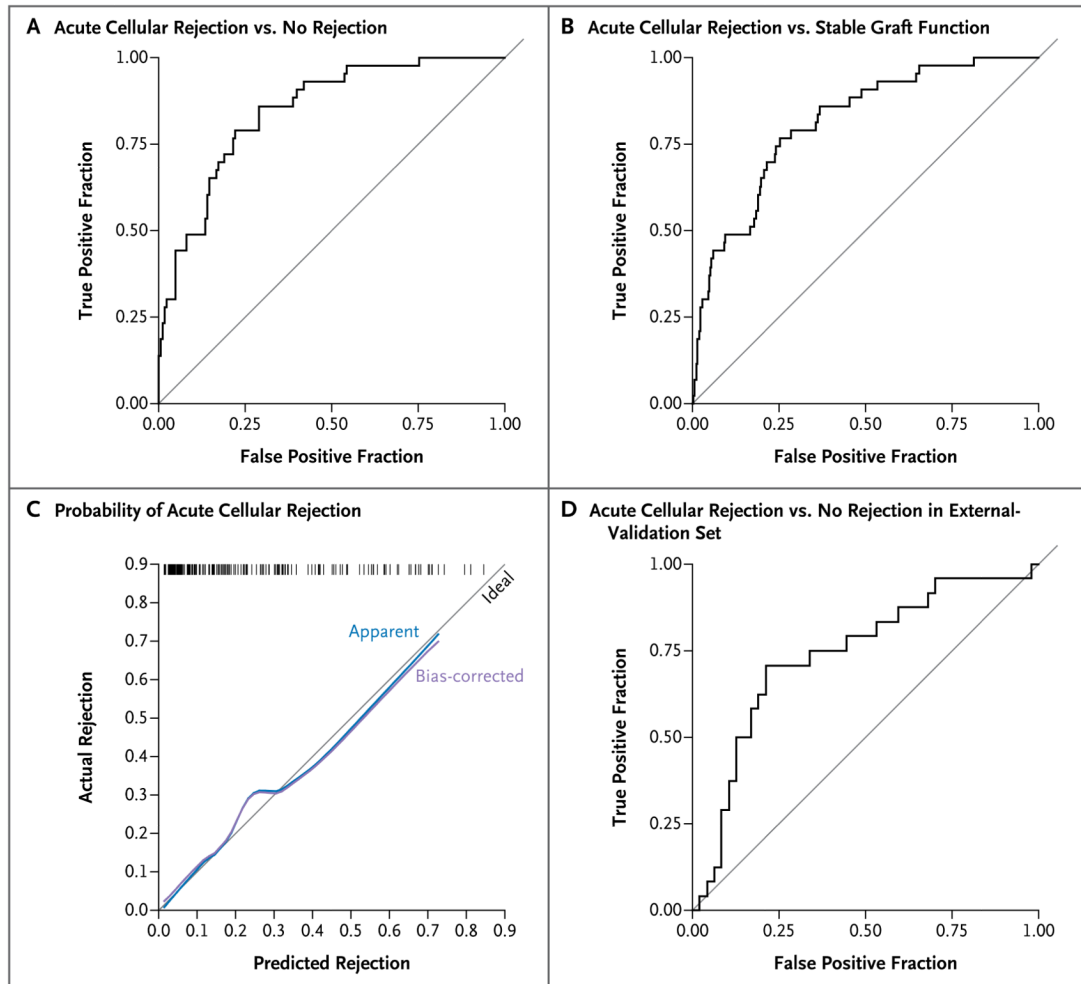
A total of 4300 urine samples were collected from 485 patients for urinary-cell messenger RNA (mRNA) profiling after transplantation on days 3, 7, 15, and 30; in months 2, 3, 4, 5, 6, 9, and 12; and at the time of kidney-allograft biopsy and 2 weeks thereafter. Of the 4300 urine specimens, 3559 were classified as passing quality control (QC) and 741 were classified as not passing. A total of 220 patients underwent 410 kidney-allograft biopsies, and 265 did not undergo biopsy. The numbers of patients with biopsy-matched urine samples (urine samples that were collected from 3 days before to 1 day after biopsy and that passed QC) are shown for patients with acute cellular rejection (defined as Banff grade IA or higher), for those without any rejection features in the biopsy sample, for those with acute antibody-mediated rejection, for those with borderline changes, and for those with other biopsy findings. The number of patients listed under different diagnostic categories exceeds the 220 patients who underwent biopsy because several patients had multiple specimens with different diagnoses. Among the 265 patients who did not undergo biopsy, 202 met the criteria for stable graft function, of whom 201 had urine samples that passed QC. Patient-enrollment information and recipient and donor characteristics are provided in the Supplementary Appendix.



**Figure 2. (facing page). Levels of mRNA in Urinary Cells**

Box-and-whisker plots show the  $\log_{10}$ -transformed ratios of mRNA copies per microgram of total RNA to 18S ribosomal RNA (rRNA) copies ( $\times 10^{-6}$ ) per microgram of total RNA for CD3, perforin, granzyme B, interferon-inducible protein 10 (IP-10), CXCR3, CD103, transforming growth factor  $\beta$  1 (TGF- $\beta$  1), and proteinase inhibitor 9 in 43 urine samples matched to 43 biopsy specimens (from 34 patients) showing acute cellular rejection, 163 urine samples matched to 163 biopsy specimens (from 126 patients) showing no rejection, and 1540 longitudinally collected urine samples from 201 patients with stable graft function who did not undergo biopsy. The horizontal line within each box represents the median, the bottom and top of each box represent the 25th and 75th percentile values, and the I bars represent the 10th and 90th percentile values; the diamond indicates the mean, and circles indicate outliers. The mRNA levels of CD3, perforin, granzyme B, and IP-10 differed significantly among the three groups ( $P < 0.001$  for all comparisons), but not the levels of CXCR3 ( $P = 0.06$ ), CD103 ( $P = 0.13$ ), TGF- $\beta$  1 ( $P = 0.11$ ), and proteinase inhibitor 9 ( $P = 0.38$ ). P values are based on the Kruskal–Wallis test, with the  $\log_{10}$ -transformed, 18S-normalized mRNA levels treated as the dependent variable. Pairwise group comparisons by

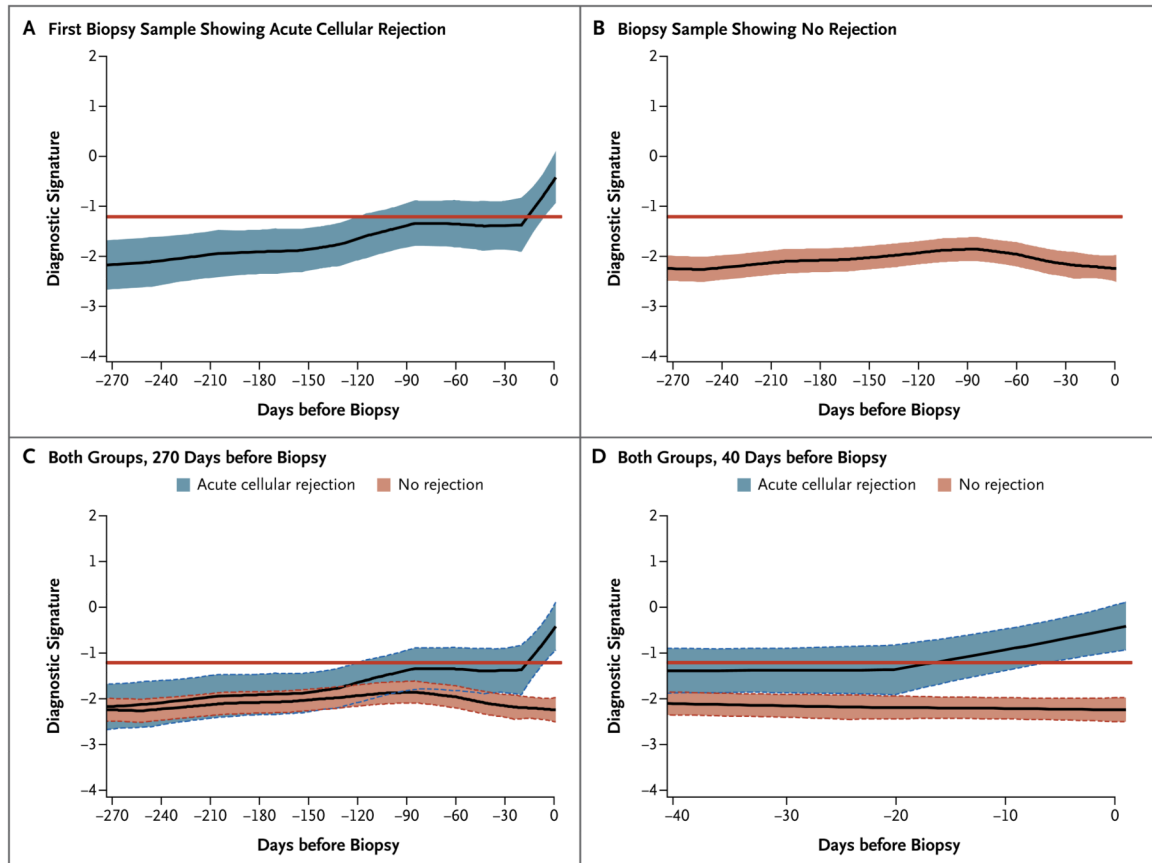
means of the Mann–Whitney test showed that the mRNA levels for CD3 , perforin, granzyme B, and IP-10 in patients with acute cellular rejection were significantly higher than the levels in those with specimens showing no rejection ( $P<0.001$  for each mRNA) and in those with stable graft function ( $P<0.001$  for each mRNA).



**Figure 3. Receiver-Operating-Characteristic Curves and Calibration Curve for the Diagnostic Signature**

The fraction of true positive results (sensitivity) and the fraction of false positive results (1 - specificity) for the diagnostic signature (calculated from  $\log_{10}$ -transformed values for 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA) as a biomarker of acute cellular rejection are shown in Panels A and B. Panel C shows the calibration plot based on bootstrap validation; vertical lines at the top of the plot indicate individual observations in the data set. In a comparison of the group of patients who had biopsy specimens showing acute cellular rejection with the group of patients who had biopsy specimens showing no rejection, the area under the curve (AUC) was 0.85 (95% CI, 0.78 to 0.91) (Panel A). In a comparison of the group of patients who had biopsy specimens showing acute cellular rejection with the group of patients who had stable graft function, the AUC was 0.81 (95% CI, 0.75 to 0.87) (Panel B). An analysis that included only specimens matched to biopsies performed because of clinical signs of rejection (38 specimens showing acute cellular rejection vs. 107 showing no rejection) showed that the three-gene signature was diagnostic of acute cellular rejection with 80% specificity (95% CI, 73 to 88) and 79% sensitivity (95% CI, 66 to 92) (AUC, 0.85; 95% CI, 0.79 to 0.92;  $P < 0.001$ ). The combination of perforin and IP-10 mRNAs or perforin and CD3 mRNAs predicted acute cellular rejection almost as well as the combination of 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA. The addition of perforin to the combination of 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA did not

improve the prediction. For the combination of 18S-normalized perforin mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA, the AUC was 0.84 (95% CI, 0.78 to 0.90;  $P < 0.001$ ), and for the combination of 18S-normalized CD3 mRNA, 18S-normalized perforin mRNA, and 18S rRNA, the AUC was 0.84 (95% CI, 0.76 to 0.91;  $P < 0.001$ ). Bootstrap validation confirmed that the best model consisted of 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA as predictors. Cross-validated estimates of the AUC and calibration-curve intercept and slope were 0.83,  $-0.06$ , and  $0.92$ , respectively. The loess-smoothed estimates of the cross-validated and unadjusted calibration curves are overlaid on a diagonal reference line representing perfect model calibration (Panel C). The three-gene model of 18S-normalized CD3 mRNA, 18S-normalized IP-10 mRNA, and 18S rRNA, with all values  $\log_{10}$ -transformed, was superior to any of the single-gene models considered. The ROC curve of the three-gene signature discriminating between specimens showing acute cellular rejection and those showing no rejection in the external-validation data set had an AUC of  $0.74$  (95% CI,  $0.61$  to  $0.86$ ;  $P < 0.001$ ) (Panel D), which was not significantly lower than the AUC of  $0.85$  in the primary data set ( $P = 0.13$ ).



#### Figure 4. Retrospective Trajectory of Diagnostic Signature

The average within-person retrospective trajectory of the diagnostic signature (i.e., the trajectory as a function of the time before biopsy) in urine samples obtained at or before biopsy that passed quality control are shown for the group of 38 patients with first biopsy specimens showing acute cellular rejection (201 urine samples) (Panel A) and the group of 113 patients with specimens showing no rejection (833 urine samples) (Panel B). Only specimens obtained during the first 400 days after transplantation were included. The diagnostic signature remained relatively flat and well below the  $-1.213$  threshold that was diagnostic of acute cellular rejection during the 270 days before biopsy in the group of patients with findings showing no rejection (Panel C). There was a significant difference in the trajectories between the two groups, with a marked increase in the diagnostic signature during the 20-day period before the first specimen showing acute cellular rejection ( $P < 0.001$ ) (Panel D). The y-axis values are diagnostic-signature scores without intrinsic units of measurement; they were calculated from the logistic-regression equation  $(-6.1487 + 0.8534 \log_{10}[\text{CD3} / 18\text{S}] + 0.6376 \log_{10}[\text{IP-10}/18\text{S}] + 1.6464 \log_{10}[18\text{S}])$  as follows. Absolute levels of CD3 mRNA, IP-10 mRNA, and 18S rRNA in the cells from each urine sample were measured by polymerase-chain-reaction assay, with the units of measurement being copies per microgram of total RNA for each mRNA measure and copies ( $\times 10^{-6}$ ) per microgram of total RNA for 18S rRNA. The mRNA copy numbers were 18S-normalized by dividing the mRNA copy number by the 18S rRNA copy number in the same sample, and the ratio was  $\log_{10}$ -transformed. In all the panels, the black lines indicate the trajectory, the colored bands the 95% confidence interval, and the red lines the diagnostic threshold.