

Efficacy of Tacrolimus in Rheumatoid Arthritis Patients Who Have Been Treated Unsuccessfully With Methotrexate

A Six-Month, Double-Blind, Randomized, Dose-Ranging Study

Daniel E. Furst,¹ Kenneth Saag,² M. Roy Fleischmann,³ Yvonne Sherrer,⁴ Joel A. Block,⁵ Thomas Schnitzer,⁶ Joel Rutstein,⁷ Andrew Baldassare,⁸ Jeffrey Kaine,⁹ Leonard Calabrese,¹⁰ Frederick Dietz,¹¹ Marshall Sack,¹² R. Gordon Senter,¹³ Craig Wiesenhutter,¹⁴ Michael Schiff,¹⁵ C. Michael Stein,¹⁶ Yoichi Satoi,¹⁷ Alan Matsumoto,¹⁸ Jacques Caldwell,¹⁹ Robert E. Harris,¹ Larry W. Moreland,² Eric Hurd,²⁰ David Yocum,²¹ and David A. Stamler⁷

Objective. To assess the efficacy, safety, and optimal dose of tacrolimus monotherapy in patients with rheumatoid arthritis (RA).

Methods. This phase II, randomized, double-blind, placebo-controlled monotherapy study was set in 12 community sites and 9 university-based sites. Two

hundred sixty-eight patients with RA who were resistant to or intolerant of methotrexate (mean dose 15.2 mg/week) and had active disease for at least 6 months (mean tender joint count 28.2, mean erythrocyte sedimentation rate 46.5 mm/hour) were randomized to receive treatment after discontinuation of methotrexate. Those who received at least 1 dose of tacrolimus were analyzed; 141 completed the study. Stable dosages of nonsteroidal antiinflammatory drugs and low-dose prednisone were allowed during treatment. All patients were given 1, 3, or 5 mg of tacrolimus or placebo once daily for 24 weeks. The American College of Rheumatology definition of 20% improvement (ACR20) and the tender and swollen joint counts at the end of treatment were the primary outcomes.

Results. ACR20 response rates demonstrated a clear dose response. The ACR20 response was observed in 15.5% of patients receiving placebo (95% confidence interval [95% CI] 7.1–23.9%), 29% of the 1 mg tacrolimus group (95% CI 18.3–39.7%) ($P < 0.058$); 34.4% of the 3 mg group (95% CI 22.7–46.0%) ($P < 0.013$), and 50% of the 5 mg group (95% CI 37.8–62.3%) ($P \leq 0.001$). The tender joint count improved statistically significantly in all tacrolimus groups. The swollen joint count, physical function, and patient-assessed pain improved statistically significantly in the 3 mg and 5 mg groups. The incidence of creatinine elevation $\geq 40\%$ above baseline levels increased in a dose-dependent manner. Drop-out rates were high (41–59%) and were more common for inefficacy in the placebo patients (71.4%), whereas they were more common for toxicity in the high-dose

Supported by Fujisawa Research Institute of America, Inc., Clinical Research Center, North America (Deerfield, Illinois), and by the Rasmuson Center for Arthritis, Orthopedics and Musculoskeletal Diseases (trial registry FK506RA-001).

¹Daniel E. Furst, MD, Robert E. Harris, MD: University of California at Los Angeles Medical School; ²Kenneth Saag, MD, Larry W. Moreland, MD: University of Alabama at Birmingham; ³M. Roy Fleischmann, MD: Rheumatology Associates, Dallas, Texas; ⁴Yvonne Sherrer, MD: Rheumatology/Immunology Arthritis Center, Fort Lauderdale, Florida; ⁵Joel A. Block, MD: Rush–Presbyterian–St. Luke's Medical Center, Chicago, Illinois; ⁶Thomas Schnitzer, MD: Northwestern University, Chicago, Illinois; ⁷Joel Rutstein, MD, David A. Stamler, MD: Arthritis Diagnosis and Treatment Center, San Antonio, Texas; ⁸Andrew Baldassare, MD: Arthritis Consultants, St. Louis, Missouri; ⁹Jeffrey Kaine, MD: Sarasota Arthritis Center, Sarasota, Florida; ¹⁰Leonard Calabrese, DO: Cleveland Clinic Foundation, Cleveland, Ohio; ¹¹Frederick Dietz, MD: Rockford Clinic, Rockford, Illinois; ¹²Marshall Sack, MD: Austin Diagnostic Clinic, Austin, Texas; ¹³R. Gordon Senter, MD: Salisbury, North Carolina; ¹⁴Craig Wiesenhutter, MD: Coeur d'Alene, ID; ¹⁵Michael Schiff, MD: Denver Arthritis Clinic, Denver, Colorado; ¹⁶C. Michael Stein, MD: Scarborough, Ontario, Canada; ¹⁷Yoichi Satoi, MD: Fujisawa Research Institute of America, Chicago, Illinois; ¹⁸Alan Matsumoto, MD: Johns Hopkins Hospital, Baltimore, Maryland; ¹⁹Jacques Caldwell, MD: Florida Arthritis and Allergy Institute, Daytona, Florida; ²⁰Eric Hurd, MD: Dallas, Texas; ²¹David Yocum, MD: University of Arizona Health Science Center, Tucson.

Address correspondence and reprint requests to Daniel E. Furst, MD, UCLA Medical School–Rheumatology Division, 1000 Veteran Avenue, Rehabilitation Center Room 32–59, Los Angeles, CA 90024. E-mail: DEFurst@mednet.ucla.edu.

Submitted for publication July 11, 2001; accepted in revised form April 16, 2002.

tacrolimus groups (31–33%). Discontinuation for creatinine elevation occurred in the 3 mg (3.1%) and 5 mg (10.9%) tacrolimus groups.

Conclusion. Tacrolimus improved disease activity in methotrexate-resistant or -intolerant patients with RA. A dose response was observed when efficacy and toxicity were assessed at different doses. The optimal dose of tacrolimus appears to be >1 mg but ≤3 mg daily.

The hallmarks of rheumatoid arthritis (RA) include synovial inflammation and joint destruction. Although various mechanisms have been recognized in the pathogenesis of RA, T cell-mediated immune activities (1,2), including limited T cell clonality, passive disease transfer by T cells, specific clonal synovial and peripheral T cell expansion to type II collagen, and an altered ratio of CD4+ memory T cells to naive T cells (3–7), appear to play an important role.

Tacrolimus (FK506) is a macrolide immunosuppressant drug that has been approved for prophylaxis of liver and kidney allograft rejection and primarily affects T cell function, making it a good candidate for treating RA. Although the actions of tacrolimus have not been fully elucidated, it binds to a cytoplasmic protein, FK binding protein, and mediates immunosuppression by inhibiting calcineurin, a calcium- and calmodulin-dependent phosphatase. The principal biologic effects of calcineurin inhibition include decreased antigen-stimulated interleukin-2 (IL-2) T cell production, decreased interferon- γ and tumor necrosis factor α (TNF α), and decreased IL-2 receptor expression on T cells (8–11). The net effect is potent inhibition of human T cell proliferation. Tacrolimus is ~100-fold more potent in inhibiting T cell proliferation than is cyclosporin A, a calcineurin inhibitor with documented efficacy in treating patients with severe RA (12,13).

Because of the known role of T cell activation in disease pathogenesis, the observed immunomodulating actions of tacrolimus, and the encouraging results of a 16-week open-label study of tacrolimus in RA patients (Fujisawa Research Institute: data on file), we assessed the efficacy, optimal dose, and safety profile of tacrolimus in methotrexate-resistant or -intolerant patients with RA. Our hypothesis was that tacrolimus (FK506) was more effective than placebo, and that there was a dose response with respect to efficacy and toxicity in the treatment of RA.

PATIENTS AND METHODS

Patients. This multicenter, randomized, double-blind, placebo-controlled, parallel-group study included male and

female patients with RA who were 18–70 years of age and had active disease for at least 6 months' duration. Patients were either intolerant of methotrexate (discontinued methotrexate due to a documented adverse event) or resistant to it (active disease continuing after at least 8 weeks of therapy with methotrexate [≥ 15 mg/week]).

An institutional review board approved the protocol at all sites. After we obtained the patients' written informed consent, those patients receiving disease-modifying antirheumatic drugs (DMARDs), including methotrexate, discontinued these agents at least 4 weeks prior to screening. At screening and baseline, patients were required to have at least 10 of 68 joints assessed as tender or painful with pressure, at least 7 of 66 joints assessed as swollen, and no more than 30% variation in the tender joint count between screening and baseline (4 weeks). Prednisone, up to 10 mg per day (or equivalent), and nonsteroidal antiinflammatory drugs (NSAIDs) were permitted, provided that the doses were stable for at least 4 weeks prior to screening. Patients with controlled hypertension could be enrolled if the systolic blood pressure was ≤ 160 mm Hg and diastolic blood pressure was ≤ 90 mm Hg at screening and baseline. Nonpregnant women of child-bearing potential and sexually active men were eligible to participate in the study upon agreeing to practice an approved method of birth control.

Study exclusion criteria were an American College of Rheumatology (ACR) (revised) functional status of IV (14), intra- or periarticular steroids administered within 4 weeks prior to screening, history of malignancy other than localized skin cancer, active infection, evidence of uncontrolled medical illness, liver dysfunction as indicated by aspartate transaminase (AST), alanine transaminase (ALT), amino alkaline phosphatase, or total bilirubin levels greater than twice the upper limit of normal, renal insufficiency (serum creatinine ≥ 124 μ moles/liter), significantly elevated cell counts indicating anemia, leukopenia, or thrombocytopenia, previous total lymphoid irradiation, or known active substance abuse.

Sample size was determined based on the expected level of improvement in the tender joint count. Fifty-one patients/group were needed to detect a mean difference (between active and placebo) of 8 in the tender joint count, with a common standard deviation of 14.2, using a 2-sided test at 80% power and $\alpha = 0.05$. To achieve this sample size, 60 patients/group, or a total of 240 patients, were targeted for enrollment.

Tacrolimus treatment. Study patients were stratified, prior to randomization, according to the reason for methotrexate failure and were randomly assigned (1:1:1:1) to receive 24 weeks of treatment with either tacrolimus or matching placebo at a once-daily dose of 1, 3, or 5 mg. The randomization scheme was generated by the Research Data Operations Department at Fujisawa Healthcare, using an internally developed SAS program (Cary, NC), and supplied to an independent packaging company which packed the hard gelatin capsules by each patient number in accordance with the scheme. The study drug was then provided to each site. The allocation sequence was placed in a sealed, opaque envelope and kept in a secured location. The randomization code was concealed from sites by attaching an emergency double-blind label to each study-drug supply kit. Compliance, defined if a subject took between 85% and 110% of the study medication, was assessed by pill count.

Clinical assessments. Efficacy evaluations, performed at baseline and at 4, 8, 12, 16, and 24 weeks after randomization, were assessed by tender joint count (maximum of 68), swollen joint count (maximum of 66, excluding the hips; swelling cannot be accurately felt in the hips and thus is not measured in these joints), patient's assessment of pain (on 100-mm visual analog scale [VAS]), patient's global assessment of disease activity (on 100-mm VAS), physician's global assessment of disease activity (on 100-mm VAS), patient's assessment of physical function (by the modified Health Assessment Questionnaire [M-HAQ], part 1 [15]), and changes in the levels of acute-phase reactants (erythrocyte sedimentation rate and C-reactive protein). These represent the components of the ACR core response criteria (16).

Safety was assessed by open-ended questioning regarding problems since the last visit, plus a complete blood cell count, measurement of serum sodium, potassium, chloride, calcium, phosphate, magnesium, bicarbonate, blood urea nitrogen, glucose, hemoglobin A_{1c}, amylase, total protein, albumin, cholesterol, triglycerides, creatinine phosphokinase, AST, ALT, alkaline phosphatase, lactic dehydrogenase, total bilirubin, and uric acid, and by urinalysis. The best measures of the glomerular filtration rate (GFR) are inulin or paraaminohippurate clearances, but these were not feasible in this trial. The GFR as measured using urinary creatinine-to-serum creatinine ratios is known to be very inaccurate on an outpatient basis (17). Therefore, creatinine clearances were calculated using the method of Cockcroft and Gault (18). Laboratory parameters were evaluated 1, 2, 4, 8, 12, 16, and 24 weeks after randomization. Physical examinations were performed regularly during the trial, and electrocardiograms were obtained at baseline and at 12 and 24 weeks.

Patients who exhibited an increase of $\geq 40\%$ from the baseline serum creatinine level were required to be retested within 1 week. Patients who had persistent creatinine elevations $\geq 40\%$ above baseline levels were discontinued from the study. Similarly, patients whose end-of-study creatinine level was $> 30\%$ above baseline levels were followed up for up to 12 weeks after study completion, or until the creatinine level either returned to baseline levels or subsided within normal limits. In addition to routine medical monitoring, an independent data safety monitoring board, consisting of 2 rheumatologists, a statistician, and an ethicist, periodically reviewed trial safety data.

Statistical analysis. The primary efficacy end points were the response to treatment according to the ACR definition of 20% improvement (ACR20) at end of treatment and change from baseline to end of treatment in the tender and swollen joint counts. ACR20 is a combined index of response that requires $\geq 20\%$ improvement in both the tender and swollen joint counts plus $\geq 20\%$ improvement in 3 of the following 5 measures: patient global or physician global assessment of disease activity, patient global measurement of pain, acute-phase reactant levels (erythrocyte sedimentation rate or C-reactive protein), or a physical function measure (M-HAQ) (16). For the ACR50 response, $\geq 50\%$ improvement is required.

Secondary efficacy end points included change from baseline to end of treatment in the patient's assessment of pain, patient's global assessment of disease activity, physician's global assessment of disease activity, patient's assessment of

physical function using the M-HAQ, and erythrocyte sedimentation rate and C-reactive protein levels. In addition, improvement according to the ACR definition of 50% improvement at end of treatment (ACR50) was assessed.

Data on the intent-to-treat subset (patients who were randomized and received at least 1 dose of study medication), which was the primary patient efficacy and safety analysis group, were analyzed. The completer subset included patients who completed the 24-week study period and had an efficacy assessment at the end of treatment or later. Primary efficacy end points were analyzed using both the intent-to-treat and completer subsets.

A test for linearity of the primary efficacy end points with increasing dose was first performed to test the hypothesis that tacrolimus was effective. Subsequent tests of each active-drug dose group compared with placebo, each at the 5% significance level, were planned in the event that a statistically significant ($P < 0.05$) linear trend was observed. The chi-square test for discrete variables and one-way analysis of variance for continuous variables were used to calculate P values.

ACR20 response rates over time were analyzed in the intent-to-treat subset using the last observation carried forward method. Primary and secondary end points used the last assessment of treatment, regardless of time. Statistical testing of the proportion of patients satisfying the ACR20 criteria was conducted using logistic regression to assess dose-response linearity and to compare the 3 tacrolimus dose groups with placebo. Factors included in the model were study treatment group and stratum at randomization (i.e., methotrexate-resistant or -intolerant). Change from baseline to end of treatment in both the tender and swollen joint counts and secondary efficacy measures were analyzed using a general linear model, including the same factors as in the logistic regression model.

The incidence of adverse events in each active-treatment dose group was compared with that in the placebo group by using Fisher's exact test (2-tailed). For key laboratory parameters and blood pressure, the frequency of within-group change in each active-treatment dose group was compared with that in the placebo group using the Cochran-Mantel-Haenszel test.

The data were gathered and monitored according to Food and Drug Administration regulations. After the database was locked, the data were analyzed by statisticians at Fujisawa Research Institute of America according to prespecified criteria.

RESULTS

There were no statistically significant differences in the demographic and baseline variables among the 4 treatment groups. The mean age of patients in each group ranged from 50.1 years to 54.1 years ($P = 0.14$), with 60.1–82.3% being women ($P = 0.15$). This study was a monotherapy trial and all patients discontinued methotrexate at least 4 weeks before starting tacrolimus. None were restarted on methotrexate during the trial.

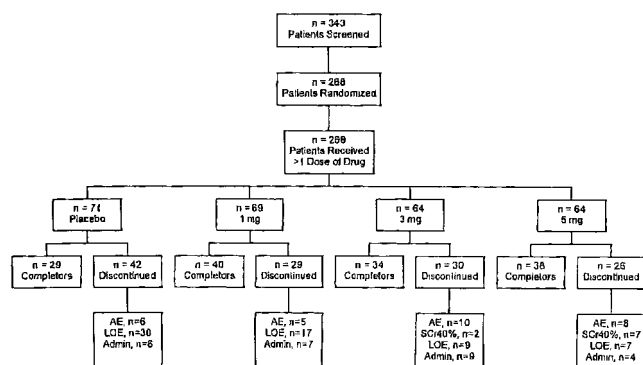


Figure 1. Flow of patients in the trial and reasons for discontinuation. AE = discontinued secondary to adverse events; LOE = discontinued secondary to lack of efficacy; Admin. = discontinued for administrative reasons; SCr40% = discontinued secondary to protocol-specified changes in the serum creatinine level $\geq 40\%$ (see Patients and Methods).

These patients either had tried methotrexate and developed a toxic reaction or had what they believed to be an inadequate response. Their mean prestudy methotrexate dose was between 14.2 mg and 16.2 mg weekly ($P = 0.19$ among groups), the mean duration of methotrexate use before washout was 21.2–22.9 months ($P = 0.99$ among groups), and 50–55% of patients discontinued methotrexate secondary to toxicity ($P = 0.90$ among groups), whereas 45–50% were resistant to methotrexate (protocol-defined active disease despite ≥ 8 weeks of methotrexate therapy at ≥ 15 mg/week [$P = 0.70$ among groups]). The mean disease duration was 9.7–11.5 years ($P = 0.13$), rheumatoid factor was positive in 65.6–74.6% of patients ($P = 0.70$), 71.9–74.6% of patients were taking up to 10 mg prednisone daily ($P = 0.34$), estimated creatine clearance rates, determined using the method of Cockcroft and Gault (18), were 99.1–99.4 ml/minute ($P = 0.99$), and body mass index values were 27.9–29.0 kg/m² ($P = 0.79$). Figure 1 describes the distribution of patients during the trial. Compliance exceeded 90% in all treatment groups.

The ACR20 response rate at the end of treatment showed a dose response. Among the placebo patients, 15.5% achieved the ACR20 response (95% confidence interval [95% CI] 7.1–23.9%), whereas among those receiving tacrolimus, 29.0% (95% CI 18.3–39.7%) in the 1 mg group ($P < 0.058$), 34.4% (95% CI 22.7–46.0%) in the 3 mg group ($P < 0.013$), and 50.0% (95% CI 37.8–62.3%) in the 5 mg group ($P \leq 0.001$) achieved this response (Figure 2 and Table 1). ACR20 response rates in the completer subset (data not shown) yielded results similar to those obtained in the intent-to-

treat analysis. The ACR50 response rate for each tacrolimus group, which ranged between 14.1% and 17.2%, was significantly higher than in those receiving placebo ($P \leq 0.05$), although no dose response was apparent for the ACR50 response rates (Figure 2 and Table 1).

The tender joint count and erythrocyte sedimentation rate improved under all 3 regimens in a dose-dependent manner ($P \leq 0.05$ –0.001). The C-reactive protein level improved slightly in all groups, and the changes were statistically significant in the 3 mg and 5 mg groups. Only the 3 mg and 5 mg tacrolimus groups were significantly different from the placebo group with respect to the swollen joint count, patient assessment of pain, and patient assessment of physical function (see Table 1 for P values). The improvements in physical function in those receiving 3 mg or 5 mg of tacrolimus, as indicated by decreases of 0.3 and 0.4 in the M-HAQ scores, respectively, were both statistically and clinically significant (Table 1) (19). The improvements in the patient and physician global assessments of disease activity were significant in patients receiving 5 mg of tacrolimus ($P \leq 0.001$ for both), and approached significance in the 3 mg group ($P = 0.051$ and $P = 0.081$, respectively). In summary, a statistically significant linear dose response across dose groups was observed for all primary and secondary efficacy end points ($P \leq 0.001$).

The overall dropout rate in this trial was 47.9%, not including adverse events associated with increased

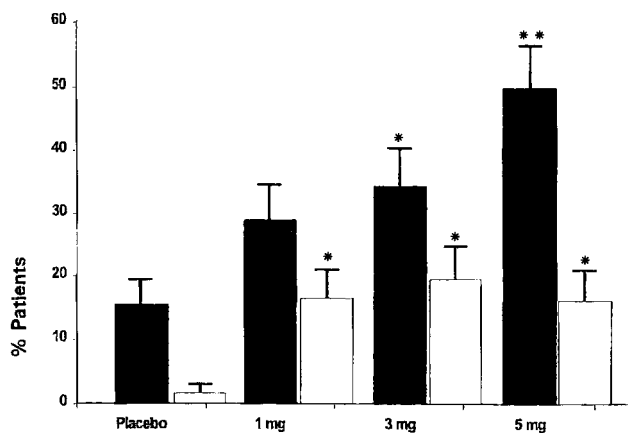


Figure 2. Percentage of patients achieving the American College of Rheumatology definition of 20% (solid bars) and 50% (open bars) improvement in prespecified criteria (16) following 24 weeks of treatment with different doses of tacrolimus (1, 3, or 5 mg) or placebo. * = $P \leq 0.05$ and ** = $P \leq 0.001$ versus placebo. Bars show the mean and SD.

Table 1. Efficacy results at end of treatment among patients with rheumatoid arthritis*

	Placebo (n = 71)		Tacrolimus					
	1 mg (n = 69)		3 mg (n = 64)		5 mg (n = 64)			
	Baseline	Change from baseline	Baseline	Change from baseline	Baseline	Change from baseline	Baseline	Change from baseline
ACR improvement response rate								
ACR20	–	15.5 (7.1–23.9)		29.0 (18.3–39.7)		34.4 (22.7–46.0)†		50.0 (37.8–62.3)‡
ACR50	–	1.4 (0–4.1)		14.5 (6.2–22.8)†		17.2 (7.9–26.4)†		14.1 (5.5–22.6)†
Response criteria								
Tender joint count (maximum 68)	28.5 ± 12.8	–1.0 ± 14.6	26.6 ± 11.1	–6.3 ± 13.0†	29.2 ± 13.0	–8.0 ± 14.2§	28.4 ± 12.4	–12.9 ± 13.4‡
Swollen joint count (maximum 66)	21 ± 9.7	–1.9 ± 10.5	19 ± 9.5	–3.8 ± 8.0	18.5 ± 8.0	–5.4 ± 10.1†	20.5 ± 9.1	–6.8 ± 8.8§
Patient pain on 100-mm VAS	68 ± 22	–5.5 ± 31	67 ± 21	–11 ± 30	66 ± 22	–16 ± 33†	67 ± 20	–24 ± 31‡
Patient global on 100-mm VAS	65 ± 24	–3 ± 30	60 ± 23	–11 ± 29	62 ± 23	–13 ± 32	59 ± 23	–21 ± 29‡
Physician global on 100-mm VAS	67 ± 18	–11 ± 23	63 ± 19	–14 ± 30	61 ± 20	–19 ± 31	67 ± 16	–28 ± 26‡
Physical function on M-HAQ (scale 1–4)	2.2 ± 0.6	0 ± 0.6	2.1 ± 0.5	–0.1 ± 0.5	2.1 ± 0.6	–0.3 ± 0.6†	2.1 ± 0.5	–0.4 ± 0.6‡
ESR, mm/hour	47 ± 31	+5 ± 25	48 ± 30	–4 ± 27†	44 ± 33	–5 ± 22†	47 ± 34	–11 ± 25‡
CRP, units/mg/liter	4.0 ± 3.4	+0.5 ± 2.4	3.5 ± 3.3	–0.3 ± 2.5	4.1 ± 4.3	–0.8 ± 2.7§	3.8 ± 3.1	–1.7 ± 2.8‡

* Values are the mean ± SD, except American College of Rheumatology 20% and 50% (ACR20 and ACR50, respectively) response rates, which are the percentage of patients meeting the improvement criteria (95% confidence intervals) considered to be the efficacy threshold. VAS = visual analog scale; M-HAQ = modified Health Assessment Questionnaire; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein.

† $P \leq 0.05$ versus placebo.

‡ $P \leq 0.001$ versus placebo.

§ $P \leq 0.01$ versus placebo.

levels of serum creatinine (these are commented upon below and showed a dose response on the adverse event curve). Among the placebo group, dropouts accounted for 59% of entered patients (42 of 71), of which 71% (30 of 42) were for inefficacy. In the 1 mg tacrolimus group, dropouts occurred at a frequency of 42% (29 of 69), among which 59% (17 of 29) were for lack of efficacy and 17% (5 of 29) were because of adverse events. For the 3-mg tacrolimus-treated patients, the dropout rate was 47% (30 of 64), but now adverse events accounted for 33% (10 of 30), whereas 30% (9 of 30) were for inefficacy. Finally, the 5 mg tacrolimus group dropout rate was 41% (26 of 64), with adverse events accounting for 31% (8 of 26) of the discontinued patients and only 27% of the patients stopping secondary to inefficacy. Thus, inefficacy accounted for most of the dropouts in the placebo group, whereas adverse events became the more prominent reason for discontinuation in the tacrolimus groups as the dose increased. The intent-to-treat analysis gave the same results as the completer analysis, so the high dropout rate did not bias the results.

Adverse events not requiring a dose change or specific therapy were very common in all groups, includ-

ing the placebo group, and in fact, occurred in 75% of patients receiving placebo (Table 2). These rates of adverse events are consistent with those observed in other studies, even in patients not receiving any drugs (20) (Table 2). In the tacrolimus groups, the incidence of patients experiencing adverse events was between 89% and 93% (14–23% higher than in the placebo group), and no dose response was apparent. Although no change in test medications was required, some adverse events were clearly more frequent in some dose groups, namely diarrhea (28.1% versus 11.3–15.6% in the 5 mg tacrolimus versus other groups), nausea (18.8% versus 5.6–15.9% in the 3 mg tacrolimus versus other groups), tremor (21.9% versus 0–4.3% in the 5 mg tacrolimus versus other groups), anxiety (10.9% versus 1.4–1.6% in the 5 mg tacrolimus versus other groups), and urinary tract infections (12.5% versus 0–9.4% in the 3 mg tacrolimus versus other groups). Discontinuations due to toxicity occurred in 8.5%, 7.2%, 15.6%, and 12.5% of patients in the placebo, 1 mg, 3 mg, and 5 mg groups, respectively.

Discontinuations due to toxicity most frequently occurred because of gastrointestinal system adverse

Table 2. Incidence of adverse events among the 4 treatment groups*

Body system	Placebo (n = 71)	Tacrolimus		
		1 mg (n = 69)	3 mg (n = 64)	5 mg (n = 64)
All systems	74.6	92.8†	90.6‡	89.1‡
Body as a whole	36.6	47.8	50.0	42.2
Cardiovascular	15.5	15.9	18.8	17.2
Digestive	35.2	49.3	57.8†	57.8†
Diarrhea	11.3	11.6	15.6	28.1‡
Dyspepsia	7.0	17.4	20.3‡	9.4
Nausea	5.6	15.9	18.8‡	14.1
Hematologic	1.4	7.2	0	6.3
Metabolic	4.2	5.8	10.9	14.1
Musculoskeletal	11.3	13.0	15.6	7.8
Nervous system	22.5	26.1	23.4	48.4†
Anxiety	1.4	1.4	1.6	10.9‡
Headache	11.3	10.1	20.3	15.6
Tremor	0	4.3	3.1	21.9§
Respiratory	15.5	26.1	20.3	12.5
Skin	16.9	17.4	7.8	10.9
Special senses	8.5	4.3	4.7	9.4
Urogenital¶	4.2	11.6	21.9†	20.3†
Urinary tract infection	1.4	0	12.5‡	9.4
GU symptoms	0	0	6.3	6.3
GU signs	1.4	5.7	4.8	4.8

* Values are the percentage of patients. More than 1 adverse event can be reported by a single patient, so the sum of terms may exceed 100%. Percentages are rounded and reported to the nearest whole number. *P* values are based on Fisher's exact test (2-tailed) comparing each active group with placebo. Genitourinary (GU) symptoms include incontinence as well as frequency and urgency of urination. GU signs include stone, metrorrhagia, pyuria, and hematuria.

† *P* ≤ 0.01 versus placebo.

‡ *P* ≤ 0.05 versus placebo.

§ *P* ≤ 0.001 versus placebo.

¶ Excluding creatinine or proteinuria.

events, although 1 or 2 instances of numerous causes of discontinuation occurred (Tables 2 and 3). Detailed perusal raised the probability that nervous system adverse events occurred more frequently in the 5 mg tacrolimus dose group, among which 4 patients discontinued therapy. Discontinuations mandated for elevated levels of serum creatinine showed a dose response and are detailed below.

Because of previous clinical experience with tacrolimus in the setting of transplantation, serum glucose levels, creatinine levels, and blood pressure were carefully monitored. The mean change in serum glucose concentration from baseline to end of treatment ranged from -2.1 to +8.4 mg/dl and hemoglobin A_{1c} changes ranged from 0.0 to 0.2, with no significant differences between any active treatment versus placebo.

The frequencies of change in systolic and diastolic blood pressure were analyzed separately. There were no significant changes in systolic blood pressure for

any treatment group compared with placebo. For diastolic blood pressure, the frequency of upward shifts was higher in the 3 mg and 5 mg groups versus placebo (*P* = 0.001 and *P* = 0.007, respectively), although the absolute changes were relatively small. For example, after 20 weeks of treatment, the mean increase in diastolic blood pressure was +2.0 mm Hg (SD 9.4) and +5.2 mm Hg (SD 7.4) for the 3 mg and 5 mg groups, respectively. Despite the noted diastolic blood pressure changes, treatment-emergent hypertension was not more common in these dose groups.

Figure 3 details the percentage of patients experiencing increases in serum creatinine levels relative to baseline during the study. The protocol stipulated that a 40% creatinine increase required repeat testing within 1 week; if the repeat value exceeded the 40% threshold, discontinuation from the study was required. We used the 40% criterion because based on the potential effects of concomitant medications and normal laboratory variation when serum creatinine is low, we considered it to be more appropriate than a 30% threshold. Creatinine elevation ≥40% above baseline levels occurred at some time during treatment in 7% (5 of 71), 8.7% (6 of 69), 18.8% (12 of 64), and 28.1% (18 of 64) of the placebo, 1 mg, 3 mg, and 5 mg groups, respectively. These elevations frequently improved when retested, thus not requiring discontinuation of study medication. Creatinine elevation led to discontinuation, as required by predefined criteria, in only the 3 mg group (3.1% [n = 2]) and 5 mg group (10.9% [n = 7]). In 8 of the 9 patients for whom postdiscontinuation followup was possible, the creatinine levels returned to within 40% of baseline values and within normal limits by 4 weeks of followup. Creatinine levels returned to within 0.3 mg of baseline values in 4 of the 8 patients who were followed up for 8 weeks postdiscontinuation. In the other 4 patients, creatinine values were improving at the final protocol-required measurement (8 weeks after discontinuation) and were within normal limits.

There were no statistically significant differences between any active treatment and placebo with respect to changes in hemoglobin levels, white blood cell count, liver enzymes, or cholesterol levels (data not shown). The frequency of increase in serum potassium was more common in the 5 mg group than in those receiving placebo (*P* ≤ 0.034).

A similar number of patients in each dose group had protocol deviations. Common types of deviations included failure to obtain laboratory data, electrocardiograms, or vital signs, visits outside specified time windows, and missed doses of study medication. These

Table 3. Incidence of adverse events leading to discontinuation*

	Placebo (n = 71)	Tacrolimus		
		1 mg (n = 69)	3 mg (n = 64)	5 mg (n = 64)
Total	6 (8.5)	7 (10.1)	12 (18.8)	8 (12.5)
Body as a whole				
Asthenia		1 (1.4)		1 (1.6)
Infection	1 (1.4)	1 (1.4)	2 (3.1)	
Digestive system				
Anorexia/nausea/vomiting	1 (1.4)	3 (4.3)	3 (4.7)	
Dyspepsia				1 (1.6)
Diarrhea		1 (1.4)		
Liver function tests		1 (1.4)		
Gastric ulcer			1 (1.6)	
Miscellaneous	1 (1.4)	1 (1.4)		1 (1.6)
Any musculoskeletal RA/arthritis/joints	1 (1.4)	4 (5.8)	3 (4.7)	1 (1.6)
Nervous system		1 (1.4)		
Parasthesias/abnormal vision				2 (3.1)
Tremor/twitching		1 (1.4)		2 (3.1)
Dizziness/headache			2 (3.1)	
Respiratory system		1 (1.4)	1 (1.6)	
Skin				
Rash	2 (2.8)		1 (1.6)	
Skin ulcer		1 (1.4)		
Urogenital				
Kidney calculus	1 (1.4)		1 (1.6)	
Pyelonephritis	1 (1.4)			

* Values are the no. (%) of patients. RA = rheumatoid arthritis.

protocol deviations had no significant impact on the study findings.

DISCUSSION

Tacrolimus suppresses T cell immunity, as indicated by decreased production of IL-2 and TNF α ; although it does so by binding to a cytoplasmic protein different from that of cyclosporine (21). It is important

to note that this study does not directly compare tacrolimus with cyclosporin A.

The purpose of this phase II trial was to explore the efficacy and tolerability of tacrolimus by identifying clear responses not confounded by concomitant DMARD therapy. The population targeted for this study comprised patients who do not tolerate or sufficiently respond to methotrexate, since this is the population who will likely use this therapy. Many patients had been taking >15 mg/week methotrexate for >20 months and had taken corticosteroids, yet still had active disease.

Given the fact that methotrexate had been either toxic or insufficiently effective at the doses thought appropriate at the time that this study was designed (at the present time, higher doses are recommended), it was felt to be ethical to withdraw the methotrexate for this study and have patients remain off the methotrexate during the trial, although low-dose corticosteroids and NSAIDs were allowed. It is possible that were the trial to be designed today, background DMARD therapy would be continued.

The ACR20 response in this 6-month study revealed a dose-response curve, with greatest efficacy observed in the 3 mg and 5 mg groups. Independent of

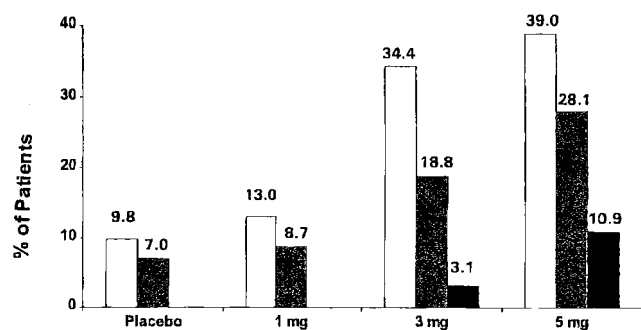


Figure 3. Percentage of patients with increases in the serum creatinine levels relative to baseline. Open bars indicate a $\geq 30\%$ rise, shaded bars indicate a $\geq 40\%$ rise, and solid bars indicate discontinuations due to a $\geq 40\%$ rise.

adverse events, these dose levels are likely to be the most effective. Although it is tempting to statistically compare the active-drug dose levels with regard to efficacy end points, the study was not designed or powered to detect such differences. Nonetheless, the linear dose response observed for all primary and secondary efficacy end points indicates differences between the active-drug dose levels that may be detected in larger trials.

The ACR50 measure eliminated much of the placebo effect, since only 1.4% of placebo patients achieved this degree of response, whereas 14.1–17.2% of tacrolimus patients achieved an ACR50 response. Since some patients receiving 1 mg tacrolimus achieved an ACR50 response, even this dose level may be effective (as defined by the ACR50 response rate) in a subset of patients.

Adverse events not leading to any change in test medications were common and occurred in 75% of patients given placebo and in up to 93% among those treated with 1–3 mg tacrolimus. The frequent occurrence of adverse events may be a function of how carefully one asks about the type of “adverse events.” For example, in one study where “adverse events” were ascertained among subjects who were not receiving any drugs, ~15% of normal volunteers experienced “adverse events” despite taking no medications during 1 week (20). During a 6-month study, therefore, the frequency of side effects in the order of magnitude found in this study would come as no surprise, if such effects are carefully sought.

Closer perusal of these events can be helpful particularly in determining a variety of gastrointestinal and genitourinary side effects, as well as, possibly, some central nervous system side effects, which were more common in the higher-dose tacrolimus groups in our study. In the 5 mg group, anxiety and tremor were more common, while in the 3 mg group, nausea and dyspepsia were more common (Tables 2 and 3). The lack of a dose response mitigates their importance. Nevertheless, one should look more carefully for these side effects in future studies.

Urinary tract infections were reported in 15 subjects. Of these, laboratory evidence of an actual urinary tract infection was found in only 3 patients (1 taking placebo, 2 taking 5 mg tacrolimus). Thus, an increased incidence of infection did not seem to occur. Of 3 discontinuations attributed to urogenital side effects, 2 occurred in the placebo group (Table 3). However, some sort of genitourinary discomfort, including incontinence, dysuria, or urination frequency, did occur in 4–6 patients per tacrolimus group versus 1 patient in

the placebo group. As for the gastrointestinal and central nervous system symptoms, one should examine the occurrence of these symptoms closely in future studies to see if they represent real adverse events from this drug. Finally, the nonsignificant upward trend in fasting blood glucose and hemoglobin A_{1c} levels should prompt careful monitoring of these parameters in subsequent investigations.

The effect of tacrolimus on renal function was carefully monitored. Criteria were established a priori for premature discontinuation based on creatinine increases above baseline levels. Forty percent increases in serum creatinine above baseline levels occurred in 7%, 9%, 19%, and 28% of patients in the placebo, 1 mg, 3 mg, and 5 mg regimens, respectively, at some time during treatment. Of these patients, 100% in the placebo and 1 mg groups and 83% and 61% in the 3 mg and 5 mg groups, respectively, showed normalization in their serum creatinine levels despite continued tacrolimus therapy. In most cases, creatinine increases were transient during treatment, indicating that either these changes were unrelated to the test medication or adaptive mechanisms allowed renal function to normalize. For all patients who discontinued treatment prematurely for creatinine elevation and who had adequate followup, the creatinine levels normalized within 4 weeks of discontinuing the study medication.

The aggregate data indicate that 5 mg of tacrolimus is more nephrotoxic than are the other regimens. The 3 mg dose may also be slightly more nephrotoxic than 1 mg, although the number of patients who discontinued due to nephrotoxicity was very small ($n = 2$ for 3 mg versus none at 1 mg). Because paraaminohippurate or iothalamate clearances were not a reliable measure (creatinine clearance using 24-hour urine samples has been shown to be a relatively inaccurate estimate of renal function, particularly in an outpatient setting) (17), the level of creatinine in the serum was used as the surrogate measure of renal function.

One might question whether background NSAID use potentiated the FK506-related toxicity, particularly the nephrotoxicity. NSAID use at baseline was similar across dose groups (68–81%; $P = 0.18$), thus precluding a bias toward more NSAID use in one group over any other as a possible cause. Furthermore, the frequency of NSAID use was not higher among patients who developed creatinine elevations, nor was it higher among those who discontinued the study due to creatinine elevation when compared with patients who did not develop such increases. (Approximately half of the patients discontinuing treatment for creatinine elevation

were receiving concomitant NSAIDs.) Thus, NSAID use did not appear to influence the creatinine results.

Dropout rates in this study were relatively high, 41–59% (Figure 1), with many dropouts being for lack of efficacy, especially in the placebo and 1 mg groups. This may reflect the patient's and/or the physician's expectation of a rapid response. Analysis of the completer subset yielded results similar to those in the intent-to-treat subset, indicating that the dropout rate did not bias the results. In fact, the predominant reasons for discontinuation, inefficacy in the placebo and 1 mg dose groups and adverse events in the 3 mg and 5 mg dose groups, support the conclusion of the study.

Although 5 mg of tacrolimus was more effective than the other regimens and might be the most effective dose for some patients, its toxicity precludes it from being the optimal dose in a general patient population. Although the 3 mg dose appeared more effective than the 1 mg dose with regard to the magnitude of improvement in the components of the ACR20 response criteria, these doses were not statistically separable by either the ACR20 or the ACR50 response criteria. Furthermore, 3 mg tacrolimus was slightly more toxic than 1 mg. Therefore, one may speculate that a tacrolimus dose >1 mg but <3 mg daily may be optimal for treating this population of RA patients, although this hypothesis needs to be tested in a well-controlled clinical trial that would include doses lower than 3 mg daily. Therefore, additional studies will be needed to place tacrolimus in the rheumatologic armamentarium.

ACKNOWLEDGMENTS

The authors would like to thank the following individuals for their valuable assistance on this project: James Hazel for project management, and Debbie Granner, Jane Lopez, and Sheila K. Van Slochteren for manuscript preparation.

REFERENCES

1. Weyand CM, Goronzy JJ. Pathogenesis of rheumatoid arthritis. *Med Clin North Am* 1997;81:29–55.
2. Fox DA. The role of T cells in the immunopathogenesis of rheumatoid arthritis. *Arthritis Rheum* 1997;40:598–609.
3. Stamenkovic I, Stegagno M, Wright KA, Krane SM, Amento EP, Colvin RB, et al. Clonal dominance among T-lymphocyte infiltrates in arthritis. *Proc Natl Acad Sci U S A* 1998;85:1179–83.
4. De Keyser F, Elewaut D, Vermeersch J, DeWever N, Cuvelier C, Veys EM. The role of T cells in rheumatoid arthritis. *Clin Rheumatol* 1995;14(Suppl 2):5–9.
5. Sekine T, Kato T, Masuko-Hongo K, Nakamura H, Yoshino S, Nishioka K, et al. Type II collagen is a target antigen of clonally expanded T cells in the synovium of patients with rheumatoid arthritis. *Ann Rheum Dis* 1999;58:446–50.
6. Wagner UG, Koetz K, Weyand CM, Goronzy JJ. Perturbation of the T cell repertoire in rheumatoid arthritis. *Proc Natl Acad Sci U S A* 1998;95:14447–52.
7. Goronzy JJ, Zettl A, Weyand CM. T cell receptor repertoire in rheumatoid arthritis. *Int Rev Immunol* 1998;17:339–63.
8. Kino T, Hatanaka H, Miyata S, Inamura N, Nishiyama M, Yajima T, et al. FK506, a novel immunosuppressant isolated from a *Streptomyces*. II. Immunosuppressive effect of FK506 in vitro. *J Antibiot (Tokyo)* 1987;40:1256–65.
9. Kino T, Inamura N, Sakai F, Nakahara K, Goto T, Okuhara M, et al. Effect of FK506 on human mixed lymphocyte reaction in vitro. *Transplant Proc* 1987;19(Suppl 6):36–9.
10. Sawada S, Suzuki G, Kawase Y, Takaku F. Novel immunosuppressive agent, FK506: in vitro effects on the cloned T cell activation. *J Immunol* 1987;139:1797–803.
11. Tocci MJ, Matkovich DA, Collier KA, Kwok P, Dumont F, Lin S, et al. The immunosuppressant FK506 selectively inhibits expression of early T cell activation genes. *J Immunol* 1989;143:718–26.
12. Woo J, Sewell HF, Thomson AW. The influence of FK506 and low-concentration cyclosporin on human lymphocyte activation, antigen expression and blastogenesis: a flow cytometric analysis. *Scand J Immunol* 1990;31:297–304.
13. Tugwell P, Bombardier C, Gent M, Bennett KJ, Bensen WG, Carrette S, et al. Low-dose cyclosporine versus placebo in patients with rheumatoid arthritis. *Lancet* 1990;335:1051–5.
14. Hochberg MC, Chang RW, Dwosh I, Lindsey S, Pincus T, Wolfe F, et al. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. *Arthritis Rheum* 1992;35:498–502.
15. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26:1346–53.
16. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727–35.
17. Blumenthal HP, Nelson DM, Spiegel HE, Givens SS, Weissman L. Compliance assessment in healthy volunteers collecting 24 hour urine samples [abstract]. *Clinical Research* 1984;32:239A.
18. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976;16:31–41.
19. Wells G, Boers M, Shea B, Tugwell P, Westhovens R, Saurez-Almazor M, et al, for the OMERACT/ILAR Task Force on Generic Quality of Life, Life Outcome Measures in Rheumatology. Sensitivity to change of generic quality of life instruments in patients with rheumatoid arthritis: preliminary findings in the generic health OMERACT study. *J Rheumatol* 1999;26:217–21.
20. Reidenberg MM, Lowenthal DT. Adverse non-drug reactions. *N Engl J Med* 1968;279:678–9.
21. Morris RE. Rapamycin: FK506's fraternal twin or distant cousin. *Immunol Today* 1991;12:137–40.