A Retrospective Analysis of Therapy for Acute Graft-Versus-Host Disease: Initial Treatment

By Paul J. Martin, Gary Schoch, Lloyd Fisher, Vera Byers, Claudio Anasetti, Frederick R. Appelbaum, Patrick G. Beatty, Kristine Doney, George B. McDonald, Jean E. Sanders, Keith M. Sullivan, Rainer Storb, E. Donnall Thomas, Robert P. Witherspoon, Pavel Lomen, John Hannigan, and John A. Hansen

We have reviewed results of therapy in 740 patients with grades II-IV acute graft-versus-host disease (GVHD) after allogeneic marrow transplantation. At the beginning of therapy, 597 patients (81%) had rash, 369 (50%) had liver dysfunction and 396 (54%) had gut dysfunction. Initial treatment was with glucocorticoids (n = 531), cyclosporine (n = 170), antithymocyte globulin (ATG) (n = 156) or monoclonal antibody (n = 3) either singly (n = 633) or in combination (n = 107). Parameters of GVHD severity in each organ were recorded weekly, and evaluation of response was made using values at the initiation of secondary treatment or, for patients without such treatment, using values on day 29 of primary treatment or the last recorded value before death, whichever occurred first. Minimal criteria for improvement or progression were defined for each organ, but no attempt was made to define liver or gut outcome if another complication such as venocclusive disease or infectious enteritis was present. Improvement rates were 43% for skin disease, 35% for evaluable liver disease and 50% for evaluable gut disease. Overall complete or partial responses were seen in 44% of patients. Multivariate analyses were carried out to identify patient, disease or treatment factors associated with likelihood of overall improvement and likelihood of response in at least one organ. A similar analysis was also carried out to identify covariates associated with time to treatment failure (defined as initiation of secondary therapy or death not due to relapse of malignancy). In all three models, GVHD prophylaxis using cyclosporine combined with methotrexate was associated with favorable GVHD treatment outcome compared to prophylaxis with either agent alone, and treatment with glucocorticoids or cyclosporine was more successful than treatment with ATG. Other factors associated with unfavorable outcome in the model of time to treatment failure and also entered in one of the response models were recipient HLA disparity with the donor, presence of a liver complication other than GVHD, and early onset of GVHD. Results of this analysis indicate that glucocorticoids represent the best initial therapy available for treatment of acute GVHD, although much room for improvement remains.

ACUTE GRAFT-versus-host disease (GVHD) contributes substantially to the morbidity and mortality of allogeneic marrow transplantation.[1-3] Virtually all patients receive prophylactic immunosuppressive medications after transplantation[4] or have T cells depleted from donor marrow in order to minimize the risk of acute GVHD.[5] Although numerous studies have demonstrated the efficacy of a variety of approaches for prophylaxis, GVHD nonetheless remains a frequently encountered complication of marrow transplantation, particularly when the donor is not an HLA-identical sibling[6,7] or when prophylaxis cannot be administered because of toxicity.[8] In the past, acute GVHD has been most often treated with glucocorticoids[4,9-17] or antithymocyte globulin (ATG).[12,18] Cyclosporine has also been employed in patients who have not received this agent for GVHD prophylaxis.[13,19]

There are relatively few studies of GVHD treatment, compared to the large number of studies assessing GVHD prophylaxis. This disparity reflects in part the difficulty and complexity of evaluating the results of treatment. The first reported human trial made extensive use of case summaries in order to describe results.[18] Only three randomized trials have ever been reported,[12,13,20] and the largest contained only 77 patients. One large retrospective study has been reported recently, but durable complete response was the only endpoint analyzed.[21] Most studies have not clearly defined minimal criteria for improvement or progression of GVHD in each organ, and few published studies have given explicit consideration to the effects of other posttransplant complications involving the gut or liver. In some studies, the time of the evaluation was not defined, an issue complicated by the fact that GVHD can show a waxing and waning course. Furthermore, an individual organ such as the skin or gut can demonstrate regional differences in response to treatment, and improvement of GVHD in one organ can be accompanied by worsening in another. At any given time point, some patients may already have received additional therapy due to lack of satisfactory response to initial treatment, while others may have died. Thus a valid classification system for assessing the success of GVHD treatment must be capable of dealing with the complexity seen in such patients.

We have reviewed treatment results for patients who developed grades II-IV acute GVHD after allogeneic marrow transplantation. The purpose and goals of this study were to define the natural history of acute GVHD, to determine patient and disease characteristics that influence therapeutic response and survival, to assess treatment factors

From the Division of Clinical Research, the Fred Hutchinson Cancer Research Center, the Department of Medicine, University of Washington, Seattle, WA, and Xoma Corporation, Berkeley, CA.

Submitted March 22, 1990; accepted June 15, 1990.

Supported by Public Health Service Grants CA 18029, CA 18221, CA 15704, HL 36444 and CA 09515 awarded by the National Institutes of Health, DHHS and by Xoma Corporation. E.D.T. is the recipient of Research Career Award AI 02425 from the National Institute of Allergy and Infectious Diseases.

Presented in part at the 31st Annual meeting of the American Society of Hematology, Atlanta, GA, December 2-5, 1989.

Address reprint requests to Paul J. Martin, MD, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" *in accordance with 18 U.S.C. section 1734 solely to indicate this fact.*

0006-4971/90/7608-0020$3.00/0

that influence outcome, and to assess methods for analysis of future treatment trials. Data were collected retrospectively and then analyzed according to predefined criteria using two endpoints, treatment response and time to treatment failure. Results with these endpoints were concordant with each other and successfully identified patient, disease and treatment factors associated with favorable or unfavorable outcome after therapy of acute GVHD.

MATERIALS AND METHODS

Patients. Records were abstracted and case report forms were prepared for all patients who were given systemic immunosuppressive treatment for acute GVHD of at least grade II severity after a first allogeneic marrow transplant at the Fred Hutchinson Cancer Research Center (FHCRC) between September 1975 and October 1987. Details of the pretransplant regimens of cyclophosphamide with or without total body irradiation have been published,[22] and the GVHD prophylaxis regimens have been described.[22-26] All patients gave informed written consent for treatment according to protocols approved by the FHCRC Institutional Review Board.

Staging of organ dysfunction. Although subsequent improvement or deterioration was evaluated only for organs without complications other than GVHD, severity of dysfunction in the skin, liver and gut at the onset of treatment was categorized in all patients regardless of whether GVHD or other complications accounted for any abnormality. Thus, this categorization differs fundamentally from the organ staging system proposed by Glucksberg[27] that describes the severity of GVHD per se, assuming that no other causes of abnormality are present. For all organs, stage 0 indicates normal function. Skin severity was staged according to the extent of skin surface area involvement: grade I, ≤25%; grade II, 26% to 50%; grade III, 51% to 75%, grade IV, >75%. Severity of hepatic dysfunction was categorized according to the serum bilirubin concentation: grade I, 2.0 to 2.9 mg/dL; grade II, 3.0 to 5.9 mg/dL; grade III 6.0 to 14.9 mg/dL; grade IV, >15 mg/dL. Severity of gut dysfunction was categorized according to the volume of diarrhea and the presence of cramps or visible blood in the stool: grade I, diarrhea volume 500 to 999 mL/day; grade II, 1,000 to 1,499 mL/day; grade III, diarrhea volume ≥1,500 mL/day or cramps or visible blood; grade IV, simultaneous presence of any two or all three of the criteria for grade III severity. Diarrhea volume was measured as the average for the day of evaluation and the two preceding days in order to minimize errors caused by large day-to-day variation. Stool volumes greater than 500 mL were not considered if urinary mixing was noted. The scales for categorizing skin and liver severity are similar to those described by Glucksberg for staging GVHD, but the scale for gut severity is different. The Glucksberg criteria require a diarrhea volume >1,500 mL/day for stage III gut GVHD and >2000 mL/day for stage IV gut GVHD.[27] In the present study, abdominal cramping and visible stool blood were used as criteria for stage III gut dysfunction because diarrhea volume showed marked daily variation and was influenced by other factors such as oral intake and antidiarrheal medications. For statistical analysis, a numerical score of 0 to 4 equivalent to the severity stage was assigned for each organ.

Treatment. Acute GVHD was generally treated with methylprednisolone (2 mg/kg/day), ATG (15 mg/kg every other day for 6 doses) or cyclosporine (3 mg/kg/day intravenously or 12.5 mg/kg/day orally).[11-13,18,20] Glucocorticoid treatment was continued at full doses for at least 14 days. Thereafter, doses were tapered over variable periods of time as allowed by the clinical response. Cyclosporine was given at prescribed or at maximally tolerated doses for 50 days. Thereafter, doses were decreased by 5% per week until discontinuation after 180 days. Cyclosporine was generally used only for patients who had not received this agent for GVHD prophylaxis. Methotrexate and cyclosporine administered for GVHD prophylaxis were continued during GVHD treatment unless there was toxicity. Secondary therapy for GVHD was initiated whenever the attending physician made a change in treatment because the response to primary therapy was less than satisfactory. In general, secondary treatment was started if GVHD was progressive anytime after at least three days of initial therapy, if unimproving grades III-IV GVHD persisted after at least seven days of initial therapy or if unimproving grade II GVHD persisted after at least 14 days of initial therapy. Glucocorticoid doses increased because of GVHD recurring during the pretaper or taper phase of primary treatment with glucocorticoids and extension of ATG therapy beyond the originally prescribed course were considered secondary therapy. Treatment of chronic GVHD was not considered secondary therapy.

Assessment of complications other than GVHD. The presence of liver or gut complications other than GVHD was determined by reviewing results of extensive clinical evaluation of each patient, including cultures and biopsies, together with autopsy information, if available. Skin biopsies were evaluated at the time of treatment in 96% of patients in this study. Liver biopsies were evaluated before the start of treatment in 3% of patients and afterwards in 9% of patients. Gut biopsies were evaluated before the start of treatment in 9% of patients and afterwards in 25% of patients. Full or partial autopsies were carried out in 282 (53%) of the 531 patients who died. Clinical and pathologic criteria for judging the presence of liver or gut complications other than GVHD have been described elsewhere.[28-30]

Measurements of response. Measurements of involved skin surface area, serum bilirubin, stool volume (three day average), and creatinine were recorded weekly through day 29 of treatment or until death or a change in therapy made because of unsatisfactory response to treatment. Additional notations were made recording the presence of visible blood in the stool, or abdominal cramping and the use of dialysis. Response was evaluated by using a nonfixed time point that allowed a maximum period of observation for assessment of outcome in each patient. Thus, responses were evaluated according to values at the initiation of secondary treatment or for patients without such treatment, according to values on day 29 of primary treatment or the last recorded values before death, whichever occurred first.

No attempt was made to define response in the liver or gut if a complication other than GVHD was known to be present in the respective organ at any time during initial treatment even if there was biopsy evidence of GVHD. Skin disease was considered improved if there was resolution of rash or decrease of involved surface area by ≥25%. Progressive skin disease was defined as an increase in involved surface area by ≥25%. Liver disease was considered improved if there was a decrease in serum bilirubin to less than 2 mg/dL for patients with baseline values of 2 to 4 mg/dL, a decrease of ≥2 mg/dL for patients with baseline values of 4 to 8 mg/dL, or ≥25% decrease in serum bilirubin for patients with baseline values ≥8 mg/dL. Progressive liver disease was defined as an increase of serum bilirubin by ≥2 mg/dL for patients with baseline values <8 mg/dL or ≥25% increase in serum bilirubin for patients with baseline values ≥8 mg/dL. Because an increase or decrease in serum bilirubin can reflect altered renal function,[31] improvement and progression of liver disease were not scored if the serum bilirubin and serum creatinine both increased or decreased such that the bilirubin: creatinine ratio was changed by <25%. Likewise, liver disease was considered nonevaluable if a "stable" bilirubin was accompanied by an increase or decrease in serum creatinine such that the bilirubin: creatinine ratio was changed by >25%. In this situation, an improvement or deterioration in hepatic function could have been masked by an opposite change in renal function. The 25% threshold for the ratio

was arbitrarily judged to represent the smallest change that could be reliably detected. Gut disease was considered improved if there was resolution of diarrhea or decrease in the three day average stool volume by ≥500 mL with clearing of cramps and bleeding if present. Clearing of any cramps and bleeding was considered as evidence of improvement in patients who had diarrhea volumes <500 mL but not in patients who had unchanged diarrhea volumes ≥500 mL. Progressive gut disease was defined as an increase in the three day average stool volume by ≥500 mL or the development of new cramps or bleeding. Stool volumes ≥500 mL were not considered if urinary mixing was noted. Factors such as the platelet count and the amount of oral intake or antidiarrheal medications were not taken into account as possibly affecting stool blood and diarrhea volume. For all organs, assessment of treatment response was made entirely according to clinical criteria regardless of biopsy or autopsy findings.

Overall response categories were determined for each patient. Organs not evaluable because of complications other than GVHD or because of inability to determine response were not considered. Complete response (CR) was defined as resolution of GVHD in all evaluable involved organs with no subsequent additional treatment given for acute GVHD. Partial response (PR) was defined in two ways: as improvement in at least one evaluable organ without deterioration in others, or as resolution of GVHD in all evaluable organs with a requirement for additional treatment because of abnormality persisting in a nonevaluable organ. Mixed response (MR) was defined as improvement in at least one evaluable organ with deterioration in at least one other. Overall progression (P) was defined as deterioration in at least one evaluable organ without improvement in others, while no change (NC) was defined as the absence of any difference sufficient to meet minimal criteria for improvement or deterioration in any evaluable organ after treatment.

Statistical analysis. Multivariate logistic regression analysis[32] was used to identify factors associated with response categories. Multivariate proportional hazards regression analysis[33] was used to identify factors associated with time to treatment failure defined as either the initiation of secondary treatment (any change in therapy made because of lack of satisfactory response) or death not due to relapse of malignancy, whichever occurred first. In some protocols, patients were treated with a succession of agents (eg ATG followed by glucocorticoids).[20] This was not considered a treatment failure unless the originally prescribed regimen had been changed due to lack of satisfactory response. Covariates considered for the multivariate models included patient characteristics (age, diagnosis, recipient HLA disparity, and patient and donor CMV serology), chronology of GVHD and treatment (interval times from transplant to onset of GVHD and treatment, interval time between onset and treatment), severity of organ dysfunction at the beginning of treatment (percent skin involvement, serum bilirubin, gut severity score, sum total severity score for all organs, presence of complications other than GVHD in the liver or gut, creatinine ≥ 2 mg/dL), treatment characteristics (agents used for GVHD prophylaxis, continuity of prophylaxis, laminar air flow isolation and agents administered for GVHD therapy) and calendar year of transplant. Covariates were entered in the model in a step-up progression until the P value for the improvement X^2 was >.05 for all remaining factors not included in the model. All P values are two-sided without adjustment for multiple comparisons. Relative risks are adjusted for the influence of other factors entered into the final model. Thus these values estimate the independent incremental risk added by each factor after other factors have been taken into account.

RESULTS

Patient characteristics. Between September 1975 and October 1987, 1,986 patients received an allogeneic marrow transplant at the FHCRC. During this 12 year period, 740 patients (37%) received primary treatment for acute GVHD, and 427 of the 740 (19%) received secondary treatment. Demographic characteristics of the patients treated for GVHD are summarized in Table 1. The median age of these 740 patients, their sex distribution and the proportions in various diagnostic categories were similar to the corresponding demographic characteristics of the overall 1,986 patient population (data not shown). As might be expected, the proportion of patients with HLA-partially matched donors was greater among the population treated for GVHD (31%) than in the overall population (9%) (data not shown). Until May 1980, virtually all patients received a standard regimen of methotrexate as prophylaxis for GVHD.[22] Studies of prophylaxis with cyclosporine were initiated in June 1980,[23,24] and studies of prophylaxis with a combined methotrexate/cyclosporine regimen were initiated in June 1983.[25,26] Small numbers of patients were given either unmodified[34] or T-cell-depleted marrow[35] with no GVHD prophylaxis, and in certain studies, ATG or glucocorticoids were included in the prophylaxis regimen. In many patients it was necessary to reduce or omit one or more doses of methotrexate or cyclosporine, generally because of renal impairment or severe mucositis. Some patients were given glucocorticoids, generally methylprednisolone (1 mg/kg/day), as substitute

Table 1. Patient Characteristics (n = 740)

Median patient age, years (range)	23 (0.8-62)
Patient sex, n (%)	
Male	463 (63)
Female	277 (37)
Pretransplant diagnosis, n (%)	
Acute nonlymphoblastic leukemia	230 (31)
Acute lymphoblastic leukemia	191 (26)
Chronic myelogenous leukemia	172 (23)
Aplastic anemia	55 (7)
Other	92 (12)
Number of recipient HLA disparities, n (%)	
0*	514 (69)
1	111 (15)
2	89 (12)
3	26 (4)
Prescribed prophylaxis, n (%)	
Methotrexate	400 (54)
Cyclosporine	41 (6)
Methotrexate/Cyclosporine	281 (38)
No prophylaxis, T-cell depletion	4 (1)
No prophylaxis, unmodified marrow	14 (2)
Supplemental ATG	27 (4)
Supplemental glucocorticoids	24 (3)
Interruption of original prophylaxis, n (%)	412 (56)
Substitution of glucocorticoids as prophylaxis, N (%)	73 (10)

*Forty-nine donors were HLA-phenotypically identical haploidentical relatives and 14 were HLA-phenotypically identical unrelated individuals. All other donors in this category were HLA-genotypically identical siblings.

prophylaxis when the originally prescribed regimen could not be administered because of toxicity or pretransplant renal impairment.

Severity of organ dysfunction and response to treatment. The onset of acute GVHD occurred at a median of 17 days after transplantation, and treatment was initiated at a median of 24 days. At the beginning of treatment, 81% of patients had rash, 50% had liver dysfunction, and 54% had gut symptoms (Table 2). A majority of patients had skin disease of at least stage III severity, while fewer patients had moderate or severe (stages III-IV) liver or gut dysfunction. Renal impairment (creatinine ≥2.0 mg/dL) was present in 87 patients (11%), and 39 (5%) were on dialysis when GVHD treatment was started. Initial treatment was with glucocorticoids (n = 531), cyclosporine (n = 170), ATG (n = 156), or monoclonal antibody (n = 3) either singly (n = 633) or in combination (n = 107). For a variety of reasons (usually the presence of a complication other than GVHD), outcome could not be evaluated in 58% of the 488 patients with liver dysfunction and in 29% of the 507 patients with gut symptoms (Table 3). Improvement or resolution of GVHD in the respective organ was seen in 43% of patients with skin disease, 35% of patients with evaluable liver disease and in 50% of patients with evaluable gut disease. Progression of involvement was more likely for liver and gut disease than for skin disease. Among the 236 patients with other posttransplant complications involving the liver, 75 (32%) had moderate or severe (stages III-IV) hepatic dysfunction before GVHD treatment, and 26% showed improvement in hepatic function after treatment for GVHD ($P < .05$ compared to patients without other liver complications) (data not shown). Among the 125 patients with other posttransplant complications involving the gastrointestinal tract, 52 (42%) had moderate or severe (stages III-IV) gut dysfunction before GVHD treatment, and only 16% showed improvement in gut function after treatment for GVHD ($P < .000001$ compared to patients without other gastrointestinal complications) (data not shown). These findings illustrate the difficulty in evaluating organs with multiple complications.

Correlation between response categories and nonrelapse mortality. For each patient, an overall response category was defined according to outcome in evaluable organs. In some patients, a response category could not be assigned because none of the involved organs could be evaluated. It should also be recognized that according to the classification criteria, a patient with resolution of rash and diarrhea but an increase in bilirubin would be considered to have a complete response if the liver could not be evaluated because of venocclusive disease, even though hepatic GVHD may also have been present. To assess the validity of the classification system, we analyzed nonrelapse mortality for patients in each response category (Fig 1). Patients with complete response had low mortality, equivalent to that of patients who had no GVHD (not shown). Patients with partial or mixed response had intermediate mortality, while patients with stable disease or progression had the highest mortality. Mortality rates for patients within the two subcategories of partial response (see Materials and Methods) were equivalent to each other. Nonevaluable patients had mortality equivalent to that of patients with stable disease or progression (not shown). The excellent correlation between response category and subsequent nonrelapse mortality supports the validity of the classification system. Since patients with mixed response had better survival than those with progressive disease, it appears that improvement in at least one evaluable organ represents an important indicator of survival in patients treated for acute GVHD.

Table 2. Severity of Organ Dysfunction at Onset of First Line Treatment

	Organ		
Severity*	Skin (%)	Liver (%)	Gut (%)
0	140 (19)	371 (50)	335 (46)
I	92 (12)	100 (14)	176 (24)
II	114 (15)	148 (20)	59 (8)
III	384 (52)	90 (12)	124 (17)
IV	7 (1)	31 (4)	37 (5)
Missing	3	0	9

*Severity is categorized according to criteria described in Materials and Methods and does not necessarily correspond to GVHD stage: 120 patients had abdominal pain and 51 patients had visible stool blood as reasons for stage III-IV gut severity. Data indicate the numbers of patients in each category. Numbers in parentheses indicate the percent in each category.

Multivariate analyses of factors associated with treatment outcome. Overall, 18% of patients had a complete response, 44% showed overall improvement (CR + PR) and 57% showed improvement in at least one organ (CR + PR + MR) (Table 4). From data not shown, it was observed that improvement of skin and liver GVHD was less frequent when only ATG was used for treatment compared to other treatments ($P < .01$ for skin, $P < .02$ for liver). The less favorable response to treatment with ATG was confirmed in the analysis of overall response categories (Table 4). Results were similar when patients who had received glucocorticoids as substitute prophylaxis were excluded from the analysis. By the methods used for analysis in this study, there was no advantage seen for treatment with multiple agents compared to treatment with single agents.

During certain time periods, the assignment of treatment was made as part of a randomized prospective trial (n = 162),[12,13,20] while at other times, certain agents (eg, cyclosporine or ATG) were not available. Otherwise, the assignment of different treatments was made according to the preference of the attending physician. Thus, the unfavorable response to treatment with ATG could reflect an unrecognized selection bias. For example, the mean total severity score for all organs in patients treated with ATG was 5.7, compared with 5.3 for patients treated with glucocorticoids ($P < .01$) and 4.7 for patients treated with cyclosporine ($P < .001$). As a means of addressing this issue, multivariate logistic regression analysis was carried out in order to identify factors associated with overall improvement (complete or partial response) (Table 5), and a separate analysis was carried out to identify factors associated with response in at least one organ (complete, partial or mixed response)

Table 3. Response to First Line Treatment of Acute GVHD Categorized by Individual Organs

Involvement and Outcome	Organ (n = 740)		
	Skin	Liver	Gut
Definite Involvement	632 [85]	250 [34]	376 [51]
Evaluable	624	207	359
Resolution	202 (32)	62 (30)	158 (44)
Improvement	66 (11)	11 (5)	22 (6)
No change	266 (43)	32 (15)	44 (12)
Progression	90 (14)	102 (49)	135 (38)
Non-evaluable	8	43	17
Early death†	8	2	4
Interference¶	0	41	13
Uncertain Involvement	9 [1]	238 [32]	131 [18]
Data missing	9	2	6
Other complications present‡	0	236	125
Uninvolved*	99 [13]	252 [34]	233 [31]

Data indicate the numbers of patients in each category. Numbers in brackets indicate percent of total patients in each category. Numbers in parentheses indicate percent of evaluable patients in each category.

*Some organs with normal function at the beginning of treatment subsequently developed abnormality. Hence the numbers of uninvolved organs in this table are smaller than the corresponding numbers in Table 2.

†Early death was defined as occurring within 6 days after beginning treatment for GVHD.

‡Of the 236 patients not evaluable because of hepatic complications other than GVHD, 157 had venocclusive disease, 55 had viral hepatitis, 10 had bacterial infections or sepsis, 13 had fungal or other infections and 8 had other pathologic processes. Of the 125 patients not evaluable because of gut complications other than GVHD, 55 had viral enteritis; 37 had bacterial enteritis; 32 had fungal enteritis; and 11 had other pathologic processes.

¶Interference in the evaluation of liver disease occurred because of a change in renal function, and interference in the evaluation of gut disease occurred because of urinary mixing with stool in situations where there was no other basis for judging outcome (see Materials and Methods).

(Table 6). As an alternative approach, a multivariate proportional hazards regression analysis was used to identify factors associated with time to treatment failure (Table 7) defined as either the initiation of secondary treatment or death not due to recurrence of malignancy, whichever occurred first.

As might have been expected, there was good concordance among covariates that entered the two response models (Tables 5 and 6). In both models, combined cyclosporine/methotrexate prophylaxis was associated with a higher probability of improvement after treatment for GVHD, whereas use of ATG for treatment was associated with a lower probability of improvement. It also appears that GVHD treatment was less successful among patients who had received glucocorticoids as substitute prophylaxis. The apparent beneficial effect of more severe skin disease was unexpected. This finding may simply indicate that measureable improvement in skin disease occurs more readily in patients with extensive involvement than in patients with limited involvement. Paradoxically, severe gut dysfunction was associated with increased likelihood of response in at least one organ (Table 6). This may reflect the higher response rate seen for evaluable gut disease than for liver or skin disease (Table 3). Also, in deriving overall response categories, gut outcome was censored when gastrointestinal complications other than GVHD were present, thereby excluding the unfavorable results seen in such patients.

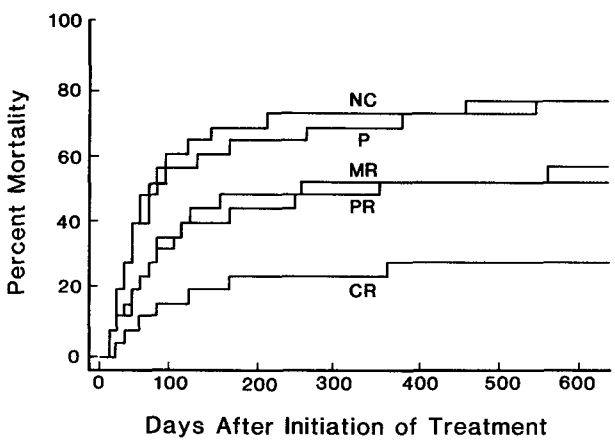

Fig 1. Nonrelapse mortality of patients categorized according to treatment outcome. CR indicates complete response; PR, partial response; MR, mixed response; NC, no change; P, progression ($P < .0001$ for equality among groups). Numbers of patients in each category are given in Table 4.

Most covariates that entered the model of time-to-treatment failure (Table 7) also entered one or both response models (Tables 5 and 6). Thus, combined cyclosporine/methotrexate prophylaxis was associated with a decreased treatment failure rate compared with prophylaxis using either agent alone, while use of ATG for treatment was associated with an increased failure rate compared to treatment with other agents. These two variables entered as covariates in all three models. The presence of a complication other than GVHD in the liver and recipient HLA disparity were both associated with an increased treatment failure rate (Table 7) and with a decreased likelihood of response in at least one organ (Table 6). Increased interval time from transplant to onset of treatment for GVHD was associated with a decreased treatment failure rate (Table 7) and with an increased likelihood of overall improvement (Table 5). A higher total organ severity score at the initiation of treatment and the use of glucocorticoids in the prophylaxis regimen were also associated with an increased treatment failure rate

Table 4. Overall Response to Treatment

Agent* (n)	Evaluable	Outcome				
		CR	CR + PR	CR + PR + MR	NC	P
ATG (61)	56	6 (11)	13 (23)	19 (34)	14 (25)	23 (41)
Cyclosporine (78)	75	16 (21)	32 (43)	42 (56)	18 (24)	15 (20)
Glucocorticoids (491)	456	92 (20)	216 (47)	271 (59)	82 (18)	103 (23)
Single Agents (633)	590	114 (19)	262 (44)	334 (57)	114 (19)	142 (24)
Combined Agents (107)	94	17 (18)	42 (45)	58 (62)	10 (11)	26 (28)
Overall (740)	684	123 (18)	304 (44)	392 (57)	124 (18)	168 (25)

Abbreviations: CR, complete response; PR, partial response; MR, mixed response; NC, no change; P, progression; NE, not evaluable. ATG, antithymocyte globulin. Data indicate numbers of patients in each category. Numbers in parentheses indicate the percent of evaluable patients in each category. The difference in outcome distribution between ATG and other treatments is statistically significant ($P < .01$).

*Results are shown for single agent treatment with ATG, cyclosporine or glucocorticoids; 95 patients received ATG as part of a combined regimen, 92 patients received cyclosporine as part of a combined regimen, and 40 patients received glucocorticoids as part of a combined regimen.

(Table 7), but neither of these factors entered either of the response models. Figure 2 illustrates Kaplan-Meier estimates of the probability of treatment failure according to presence of a liver complication other than GVHD, type of GVHD treatment, number of recipient HLA disparities, and interval time from transplant to initiation of treatment for GVHD. The favorable influence of methotrexate/cyclosporine prophylaxis on time-to-treatment failure was not detectable by univariate testing of the entire patient cohort because, for historical reasons, this regimen was preferentially used for patients with HLA disparity. Thus, the beneficial effect of methotrexate/cyclosporine prophylaxis was offset by the detrimental effects of HLA disparity and other factors. In the multivariate analysis, methotrexate/cyclosporine prophylaxis reached statistical significance and entered the model after the influence of other factors such as HLA disparity had been taken into account (Table 7).

DISCUSSION

The current study presents a retrospective analysis of results for initial treatment of acute GVHD. Our approach to this analysis has incorporated several unique features designed to improve the yield of valid information and the accuracy of interpretation. First, responses were evaluated at a standardized but nonfixed single time point. This avoided both the selection of an arbitrarily fixed time point and the complexity of analyzing multiple time points. The evaluation time point selected for our study allows a maximum period of time for measurement of response for each patient, limited by initiation of secondary treatment, or in patients without such treatment, by death or a 29-day endpoint. Second, responses were assessed by using quantitative measures of improvement or deterioration rather than changes in staging category. The criteria were clearly defined, objective, and uniformly applied. Third, no attempt was made to evaluate changes in the liver or gut if another complication was present, even though this meant excluding organs from a large number of patients. This avoided any inconsistent or subjective judgments as to whether improvement or deterioration reflected GVHD or other processes.

Concerns can be raised about the validity and limitations of methods used to evaluate response to treatment for GVHD. In any analysis of response, the time of assessment and criteria for improvement or progression may not reflect clinical practice or bedside judgment. Furthermore, the rates of improvement and progression are not easily described. The attempt to improve accuracy by excluding organs with complications other than GVHD assumes that other organs remain evaluable in such patients. Moreover, this approach also assumes either that the sensitivity for detecting such complications is high or that undetected complications have a negligible influence on the response to treatment for GVHD. Concerns about the validity of these assumptions are allayed by the close correlation between response categories and nonrelapse mortality.

Unlike the analysis of response, the analysis of time-to-treatment failure as indicated by the institution of secondary therapy reflects actual bedside judgment. We also included death from causes other than relapse as an indication of treatment failure. This assumes, perhaps correctly, that most

Table 5. Factors Associated With Overall Improvement (Complete or Partial Response)

Risk Factor*	Relative Risk†	95% CI	P‡
Methotrexate/Cyclosporine prophylaxis	1.48	1.24-1.76	.00001
ATG treatment only	0.52	0.37-0.73	.0002
Glucocorticoid substitute prophylaxis	0.64	0.47-0.86	.004
Skin disease (per 25% involvement)	1.18	1.06-1.32	.004
Time to treatment (per week after transplant)	1.10	1.02-1.19	.01

*Risk factors are listed in the order in which they entered the multivariate model.

†Values > 1.0 indicate higher likelihood of overall improvement; values < 1.0 indicate lower likelihood. Relative risks for dichotomous covariates are calculated with reference to patients not having the indicated risk factor. Values for nondichotomous covariates are calculated as the incremental risk per unit indicated in parentheses.

‡Values of *P* resulted from testing the hypothesis that the relative risk was equivalent to 1.0.

Table 6. Factors Associated With Improvement in at Least One Organ (Overall Complete, Partial or Mixed Response)

Risk Factor*	Relative Risk†	95% CI	P‡
ATG treatment only	0.55	0.41-0.75	.0002
Hepatic complication other than GVHD	0.82	0.68-0.98	.034
Methotrexate/Cyclosporine prophylaxis	1.61	1.32-1.95	.00001
Gut complication other than GVHD	0.68	0.55-0.86	.001
Gut severity score (per grade)	1.24	0.91-1.42	.003
Glucocorticoid substitute prophylaxis	0.68	0.50-0.93	.015
Recipient HLA disparity (per locus)	0.74	0.60-0.91	.006
Creatinine ≥ 2.0 mg/dL	0.53	0.30-0.92	.025
Skin disease (per 25% involvement)	1.13	1.01-1.27	.032

*Risk factors are listed in the order in which they entered the multivariate model.

†Values > 1.0 indicate higher likelihood of improvement in at least one organ, while values < 1.0 indicate lower likelihood. Relative risks for dichotomous covariates are calculated with reference to patients not having the indicated risk factor. Values for nondichotomous covariates are calculated as the incremental risk per unit indicated in parentheses.

‡Values of *P* resulted from testing the hypothesis that the relative risk was equivalent to 1.0.

deaths following an episode of GVHD are related directly or indirectly to either GVHD or its treatment. A second assumption is that deaths due to unrelated causes are relatively infrequent and randomly distributed among patient groups. The analysis of time-to-treatment failure has the advantage that all patients remain evaluable irrespective of complications other than GVHD. Failure rates can be defined precisely and risk factors can be analyzed by multivariate proportional hazards regression. The remarkable similarity of results between the multivariate analyses of treatment response and time to treatment failure suggests the validity of either approach. Covariates appearing in both types of models likely represent the factors most reliably associated with outcome.

Results of the present study indicate that glucocorticoids represent the best initial therapy currently available for treatment of acute GVHD, particularly in patients receiving cyclosporine for prophylaxis. The finding that treatment with glucocorticoids was more successful than treatment with ATG was unexpected, since a previous randomized study comparing corticosteroids with ATG in 37 patients showed no detectable difference between the two agents for treatment of GVHD.[12] Outcome measures were not the same for the two studies, and this may account for the difference in results. In the randomized study, results were evaluated at 7 days post-therapy and at discharge or death, and no attempt was made to account for visceral complications other than GVHD. Although time-to-treatment failure was not evaluated, it was noted that 12 of 17 patients treated with ATG required further therapy for GVHD, compared with 10 of 20 patients treated with glucocorticoids. Finally, the randomized study may have lacked power for detecting the difference in efficacy between the two agents. This difference may result partly from the extended duration of steroid treatment compared with ATG treatment, which is usually limited to a period of twelve days.

The association between methotrexate/cyclosporine prophylaxis and favorable GVHD treatment outcome was consistent with previous randomized studies demonstrating that this regimen can reduce the overall maximum severity of acute GVHD compared with results using either methotrexate or cyclosporine alone.[24,25] Maximum severity of GVHD reflects both the probability of developing GVHD and the likelihood of controlling the disease after it has developed. For similar reasons, the associations between HLA-disparity and poor GVHD treatment outcome might have been predicted from previous studies showing an increased probability of acute GVHD and greater overall maximum severity among patients transplanted from HLA-nonidentical donors compared with those transplanted from HLA-identical donors.[6] Plausible biologic explanations can be offered for certain other associations observed in the multivariate models. For example, early onset of moderate or severe GVHD may indicate the activation of a large number of alloresponsive T-cell clones with some being relatively resistant to inhibition by immunosuppressive agents. Some observations were unexpected and remain less well explained. For example, patients with liver complications other than GVHD

Table 7. Factors Associated With Time to Treatment Failure

Risk Factor*	Relative Risk†	95% CI	P‡
Hepatic complication other than GVHD	1.52	1.26-1.82	.00001
ATG treatment only	1.97	1.48-2.63	.00001
Recipient HLA disparity (per locus)	1.24	1.11-1.37	.0001
Initial total organ severity score (per unit)	1.08	1.03-1.13	.0021
Methotrexate/Cyclosporine prophylaxis	0.71	0.59-0.87	.0007
Time to treatment (per week after transplant)	0.94	0.90-0.98	.004
Glucocorticoid prophylaxis	1.71	1.04-2.81	.033

*Risk factors are listed in the order in which they entered the multivariate model.

†Values greater than 1.0 indicate higher likelihood of treatment failure (nonrelapse mortality or initiation of secondary treatment; values less than 1.0 indicate lower likelihood. Relative risks for dichotomous covariates are calculated with reference to patients not having the indicated risk factor. Values for nondichotomous covariates are calculated as the incremental risk per unit indicated in parentheses.

‡Values of *P* resulted from testing the hypothesis that the relative risk was equivalent to 1.0.

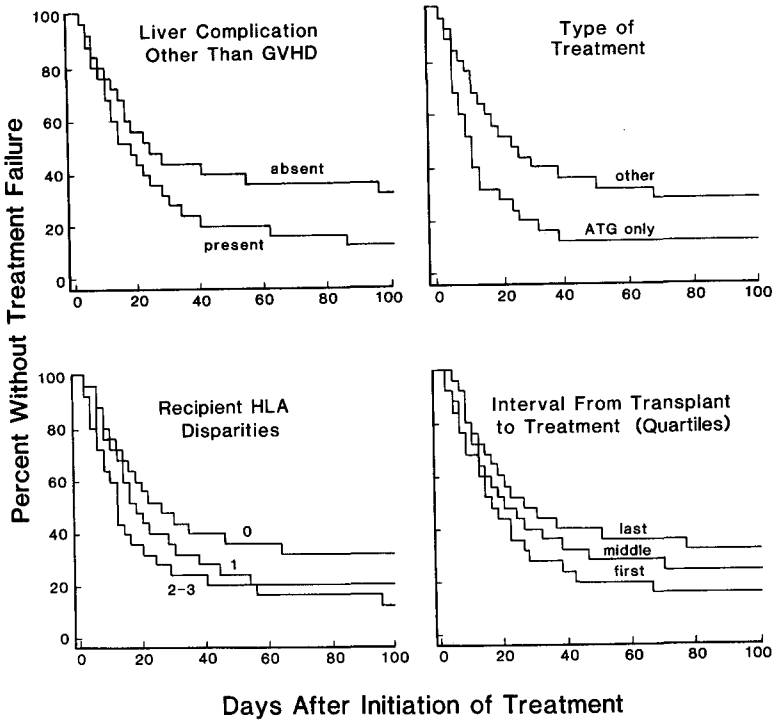

Fig 2. Association between risk factors and treatment failure. Survival plots indicate the proportion of patients alive and not requiring secondary treatment for acute GVHD. Deaths due to relapse are censored. For purposes of illustration, time interval from transplant to onset of treatment for GVHD is divided into quartiles. The first quartile encompasses the 25% of patients who had the shortest time intervals (ie, earliest treatment onset); the last quartile encompasses the 25% of patients who had the longest time intervals (ie, latest treatment onset).

responded poorly to GVHD treatment as reflected both in a lower probability of response in at least one organ and in a higher treatment failure rate.

Our results demonstrate that treatment of acute GVHD remains unsatisfactory. Among patients treated with glucocorticoids, less than half showed durable overall improvement, and results were less favorable for patients treated with ATG. A large proportion of patients required secondary treatment, and many died from GVHD itself or from infections associated with GVHD and immunosuppressive treatment. However, patients who had a complete and durable remission of GVHD showed no sequelae and had the same low nonrelapse mortality as patients who had no GVHD. We found no evidence that multiagent treatment with combinations of glucocorticoids, cyclosporine and ATG offers any measurable advantage over single agent treatment, but it remains possible that other agents or combinations can provide improved results. The factors we found to be associated with treatment outcome should be considered in the design of future clinical trials of GVHD treatment. These include recipient HLA disparity, the type of prophylaxis, time to onset of GVHD and presence of hepatic complications other than GVHD. Analysis of time-to-treatment failure could represent a useful approach for evaluating results of future studies.

ACKNOWLEDGMENT

We thank the physicians, nurses and staff of the Fred Hutchinson Cancer Research Center and Swedish Hospital Medical Center who cared for the patients. We especially thank Barbara Newton and Dara Parvin for excellent data retrieval, Larry Pietrelli for data auditing, and Alison Sell for assistance in preparation of the manuscript.

REFERENCES

1. Thomas ED, Storb R, Clift RA, Fefer A, Johnson FL, Neiman PE, Lerner KG, Glucksberg H, Buckner CD: Bone-marrow transplantation. N Engl J Med 292:832 and 895, 1975

2. Storb R, Prentice PR, Buckner CD, Clift RA, Appelbaum F, Deeg J, Doney K, Hansen JA, Mason M, Sanders JE, Singer J, Sullivan KM, Witherspoon RP, Thomas ED: Graft-versus-host disease and survival in patients with aplastic anemia treated by marrow grafts from HLA-identical siblings. N Engl J Med 308:302, 1983

3. Clift RA, Buckner CD, Thomas ED, Kopecky KJ, Appelbaum FR, Tallman M, Storb R, Sanders J, Sullivan K, Banaji M, Beatty PG, Bensinger W, Cheever M, Deeg J, Doney K, Fefer A, Greenberg P, Hansen JA, Hackman R, Hill R, Martin P, Meyers J, McGuffin R, Neiman P, Sale G, Shulman H, Singer J, Stewart P, Weiden P, Witherspoon R: The treatment of acute non-lymphoblastic leukemia by allogeneic marrow transplantation. Bone Marrow Transp 2:243, 1987

4. Vogelsang GB, Hess AD, Santos GW: Acute graft-versus-host disease: Clinical characteristics in the cyclosporine era. Medicine 67:163, 1988

5. Martin PJ, Kernan NA: T cell depletion for the prevention of graft-vs.-host disease, in Burakoff S, Deeg MJ, Ferrara J, Atkinson K (eds): Graft-vs.-Host Disease: Immunology, Pathophysiology, and Treatment. New York, NY, Marcel Dekker, 1990, p 371

6. Beatty PG, Clift RA, Mickelson EM, Nisperos B, Flournoy N, Martin PJ, Sanders JE, Storb R, Thomas ED, Hansen JA: Marrow transplantation from related donors other than HLA-identical siblings. N Engl J Med 313:765, 1985

7. Hansen JA, Clift RA, Thomas ED, Buckner CD, Storb R, Giblett ER: Transplantation of marrow from an unrelated donor to a patient with acute leukemia. N Engl J Med 303:565, 1980

8. Gale RP, Bortin MM, van Bekkum DW, Biggs JC, Dicke KA, Gluckman E, Good RA, Hoffman RG, Kay HEM, Kersey JH, Marmont A, Masaoka T, Rimm AA, van Rood JJ, Zwaan FE: Risk factors for acute graft-versus-host disease. Br J Haematol 67:397, 1987

9. Marmont AM, Damasio EE, Bacigalupo A, Giordano D, Rossi E, Reali G, Gay A, Dagna-Bricarelli F, Brema F, Carella AM, Santini G: A to O bone marrow transplantation in severe aplastic anaemia: Dynamics of blood group conversion and demonstration of early dyserythropoiesis in the engrafted marrow. Br J Haematol 36:511, 1977

10. O'Reilly RJ, Dupont B, Pahwa S, Grimes E, Smithwick EM, Pahwa R, Schwartz S, Hansen JA, Siegal FP, Sorell M, Svejgaard A, Jersild C, Thomsen M, Platz P, L'Esperance P, Good RA: Reconstitution in severe combined immunodeficiency by transplantation of marrow from an unrelated donor. N Engl J Med 297:1311, 1977

11. Weiden PL, Doney K, Storb R, Thomas ED: Anti-human thymocyte globulin (ATG) for prophylaxis and treatment of graft-versus-host disease in recipients of allogeneic marrow grafts. Transplant Proc 10:213, 1978

12. Doney KC, Weiden PL, Storb R, Thomas ED: Treatment of graft-versus-host disease in human allogeneic marrow graft recipients: A randomized trial comparing antithymocyte globulin and corticosteroids. Am J Hematol 11:1, 1981

13. Kennedy MS, Deeg HJ, Storb R, Doney K, Sullivan KM, Witherspoon RP, Appelbaum FR, Stewart P, Sanders J, Buckner CD, Martin P, Weiden P, Thomas ED: Treatment of acute graft-versus-host disease after allogeneic marrow transplantation: A randomized study comparing corticosteroids and cyclosporine. Am J Med 78:978, 1985

14. Prentice HG, Bateman SM, Bradstock KF, Hoffbrand AV: High dose methylprednisolone therapy in established acute graft-versus-host disease. Blut 41:175, 1980

15. Kendra J, Barrett AJ, Lucas C, Joshi R, Joss V, Desai M, Halil O, Rogers TR, Hobbs JR, Hugh-Jones K: Response of graft-versus-host disease to high doses of methylprednisolone. Clin Lab Haematol 3:19, 1981

16. Bacigalupo A, van Lint MT, Frassoni F, Veneziano MPG, Avanzi G, Vitale V, Marmont AM: High dose bolus methylprednisolone for the treatment of acute graft-versus-host disease. Blut 46:125, 1983

17. Kanojia MD, Anagnostou AA, Zander AR, Vellekoop L, Spitzer G, Verma DS, Jagannath S, Dicke KA: High-dose methylprednisolone treatment for acute graft-versus-host disease after bone marrow transplantation in adults. Transplantation 37:246, 1984

18. Storb R, Gluckman E, Thomas ED, Buckner CD, Clift RA, Fefer A, Glucksberg H, Graham TC, Johnson FL, Lerner KG, Neiman PE, Ochs H: Treatment of established human graft-versus-host disease by antithymocyte globulin. Blood 44:57, 1974

19. Powles RL, Barrett AJ, Clink H, Kay HEM, Sloane J, McElwain TJ: Cyclosporin A for the treatment of graft-versus-host disease in man. Lancet 2:1327, 1978

20. Deeg HJ, Loughran TP, Storb R, Kennedy MS, Sullivan KM, Doney K, Appelbaum FR, Thomas ED: Treatment of human acute graft-versus-host disease with antithymocyte globulin and cyclosporine with or without methylprednisolone. Transplantation 40:162, 1985

21. Weisdorf D, Haake R, Blazar B, Miller W, McGlave P, Ramsay N, Kersey J, Filipovich A: Treatment of moderate/severe acute graft-versus-host disease after allogeneic bone marrow transplantation: An analysis of clinical risk features and outcome. Blood 75:1024, 1990

22. Thomas ED, Buckner CD, Banaji M, Clift RA, Fefer A, Flournoy N, Goodell BW, Hickman RO, Lerner KG, Neiman PE, Sale GE, Sanders JE, Singer J, Stevens M, Storb R, Weiden PL: One hundred patients with acute leukemia treated by chemotherapy, total body irradiation, and allogeneic marrow transplantation. Blood 49:511, 1977

23. Deeg HJ, Storb R, Thomas ED, Flournoy N, Kennedy MS, Banaj M, Appelbaum FR, Bensinger WI, Buckner CD, Clift RA, Doney K, Fefer A, McGuffin R, Sanders JE, Singer J, Stewart P, Sullivan KM, Witherspoon RP: Cyclosporine as prophylaxis for graft-versus-host disease: A randomized study in patients undergoing marrow transplantation for acute nonlymphocytic leukemia. Blood 65:1325, 1985

24. Storb R, Deeg HJ, Thomas ED, Appelbaum FR, Buckner CD, Cheever MA, Clift RA, Doney KC, Flournoy N, Kennedy MS, Loughran TP, McGuffin RW, Sale GE, Sanders JR, Singer JW, Stewart PS, Sullivan KM, Witherspoon RP. Marrow transplantation for chronic myelocytic leukemia: A controlled trial of cyclosporine-versus-methotrexate for prophylaxis of graft-versus-host disease. Blood 66:698, 1985

25. Storb R, Deeg HJ, Whitehead J, Appelbaum F, Beatty P, Bensinger W, Buckner CD, Clift R, Doney K, Farewell V, Hansen J, Hill R, Lum L, Martin P, McGuffin R, Sanders J, Stewart P, Sullivan K, Witherspoon R, Yee G, Thomas ED: Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft-versus-host disease after marrow transplantation for leukemia. N Engl J Med 314:729, 1986

26. Storb R, Deeg HJ, Farewell V, Doney K, Appelbaum F, Beatty P, Bensinger W, Buckner CD, Clift R, Hansen J, Hill R, Longton G, Lum L, Martin P, McGuffin R, Sanders J, Singer J, Stewart P, Sullivan K, Witherspoon R, Thomas ED: Marrow transplantation for severe aplastic anemia: Methotrexate alone compared with a combination of methotrexate and cyclosporin for prevention of acute-graft-versus-host disease. Blood 68:119, 1986

27. Glucksberg H, Storb R, Fefer A, Buckner CD, Neiman PE, Clift RA, Lerner KG, Thomas ED: Clinical manifestations of graft-versus-host disease in human recipients of marrow from HLA matched sibling donors. Transplantation 18:295, 1974

28. McDonald GB, Sharma P, Matthews DE, Shulman HM, Thomas ED: Venocclusive disease of the liver after bone marrow transplantation; diagnosis, incidence, and predisposing factors. Hepatology 4:116, 1984

29. McDonald GB, Shulman HM, Sullivan KM, Spencer GD: Intestinal and hepatic complications of human bone marrow transplantation. Gastroenterology 90:460, 770, 1986

30. Wolford JL, McDonald GB: A problem-oriented approach to intestinal and liver disease after marrow transplantation. J Clin Gastroenterol 10:419, 1988

31. Yasmineh, WG, Killeen AA, Filipovich AH: Conjugated hyperbilirubinemia and renal function in bone marrow transplant recipients. Arch Pathol Lab Med 112:245, 1988

32. Hosmer DW, Lemeshow S: Applied Logistic Regression. New York, NY, Wiley, 1989

33. Kalbfleisch JD, Prentice RL: The statistical analysis of failure time data. New York, NY, Wiley, 1980

34. Sullivan KM, Deeg HJ, Sanders J, Klosterman A, Amos D, Shulman H, Sale G, Martin P, Witherspoon R, Appelbaum F, Doney K, Stewart P, Meyers J, McDonald GB, Weiden P, Fefer A, Buckner CD, Storb R, Thomas ED: Hyperacute graft-versus-host disease in patients not given immunosuppression after allogeneic marrow transplant. Blood 67:1172, 1986

35. Martin PJ, Hansen JA, Torok-Storb B, Durnam D, Przepiorka D, O'Quigley J, Sanders J, Sullivan KM, Witherspoon RP, Deeg HJ, Appelbaum FR, Stewart P, Weiden P, Doney K, Buckner CD, Clift R, Storb R, Thomas ED: Graft failure in patients receiving T cell-depleted HLA-identical allogeneic marrow transplants. Bone Marrow Transplant 3:445, 1988