Published in final edited form as:

Clin Trials. 2008 ; 5(6): 607–616. doi:10.1177/1740774508098326.

USE OF BIOLOGICAL ASSIGNMENT IN HEMATOPOIETIC STEM CELL TRANSPLANTATION CLINICAL TRIALS

Brent R. Logan, Ph.D.

Division of Biostatistics Medical College of Wisconsin Milwaukee, WI USA

Eric S. Leifer, Ph.D.

Office of Biostatistics Research National Heart, Lung, and Blood Institute National Institute of Health Bethesda, MD USA

Chris Bredeson, M.D.

Medical College of Wisconsin Milwaukee, WI USA

Mary M. Horowitz, M.D., M.S.

Center for International Blood and Marrow Transplant Research Medical College of Wisconsin Milwaukee, WI USA

Marian Ewell, Sc.D.

Emmes Corporation Rockville, MD USA

Shelly Carter, Sc.D.

Emmes Corporation Rockville, MD USA

Nancy L. Geller, Ph.D. [on behalf of Blood and Marrow Transplant Clinical Trials Network]

Office of Biostatistics Research National Heart, Lung, and Blood Institute National Institute of Health Bethesda, MD USA

Abstract

Background—When comparing treatments for a specific illness, it is sometimes impractical or impossible to conduct a randomized clinical trial (RCT). Biological assignment trials are one alternative design. In hematopoietic stem cell transplantation (HCT) trials, a human leukocyte antigen (HLA)-matched sibling donor is considered optimal, but such donors are available for only 20-30% of otherwise eligible patients. Rather than randomizing only those with a matched sibling donor, in a recent multiple myeloma trial, the type of HCT each patient received was biologically based, i.e., chosen according to whether or not the patient had a matched sibling donor.

Purpose—This article describes the design and implementation of biological assignment trials as well as their advantages and disadvantages.

Methods—We focus on several aspects of such trials, including efficiency of trial duration, ethical issues, and potential sources of bias. Statistical issues are considered including sample size calculations, monitoring for biased enrollment, and adjustments for imbalances in patient characteristics. A multiple myeloma trial is used as an illustration.

Results—Although they often require a larger sample size, biological assignment trials can provide substantial efficiency in terms of study duration over randomized trials when accrual to a randomized trial would be slow. Determination of sample size requires consideration of the anticipated proportion of patients with a biologically favored (HLA-matched sibling) donor. An add-on randomization of

patients without a matched sibling donor may alleviate ethical concerns about applicability of study results to all patients regardless of whether the biological assignment groups differ with respect to outcome.

Limitations—Prognostic factor imbalance and enrollment bias can occur in a biological assignment trial. Statistical adjustment for potential imbalance in prognostic factors is important, as is monitoring center accrual for enrollment bias and performing an appropriate intention-to-treat analysis.

Conclusions—A biological assignment trial can be a reasonable way to compare treatments which are biologically based, such as HLA-matched sibling transplants, when the gold-standard randomized trial design is impractical or impossible. Implementing such a trial requires careful consideration of the ethical issues and potential biases.

INTRODUCTION

Hematopoietic stem cell transplantation (HCT) is a complex and intensive set of therapies which include the administration of blood stem cells collected either from healthy donors or the patients themselves. HCT is used to treat a variety of life-threatening diseases, including leukemia, lymphoma, multiple myeloma and aplastic anemia [1]. Risks of morbidity and mortality after HCT are considerable and vary in type, timing, frequency and severity according to the type of stem cells and donor used as well as other patient-, disease- and therapy-related factors. Autologous transplantation, in which the patient receives his or her own cells after high-dose chemotherapy, has a relatively low rate of transplant-related mortality. Although autologous transplantation is effective in some diseases incurable with conventional chemotherapy, recurrent malignancy is a major cause of treatment failure. Recurrences generally occur one or more years after transplantation. Allogeneic transplantation, in which the patient receives cells from a suitably human leukocyte antigen (HLA)-matched healthy donor (usually a sibling) offers a greater chance of eradicating malignancy but confers a higher risk of transplant-related mortality due to slower hematopoietic and immunologic recovery and graft-versus-host disease. Transplant-related deaths generally occur in the first year after transplantation. Which type of transplantation (autologous versus allogeneic) is best for patients with a given disease is a pressing question. Because recurrence rates and other outcomes depend heavily on the type of disease, trials to address this question must be disease specific.

In a biological assignment trial, treatment assignment is biologically based rather than randomly assigned. In the case of HCT, this means that patients who have a suitable HLA matched sibling donor will be assigned to the allogeneic transplant arm, while patients who do not have such a donor will be assigned to another treatment, such as autologous transplantation. The existence of an HLA-matched sibling donor depends on the availability of a healthy sibling who received the same HLA genes as the patient from their common parents. Since the selection of the sibling donor and recipient genes from their parents is a random process at the time of conception, biological assignment trials are sometimes called *genetic* or *Mendelian randomization trials* [2-4]. From the trial design standpoint, it is hoped that, conditional on observable prognostic factors such as patient age and baseline disease status, the existence of a matched sibling donor is independent of patient prognosis.

Recently, the term *Mendelian randomization* has also been used to describe a study design in genetic epidemiology to reduce confounding when estimating the effect of a risk factor on disease [5-7]. Although both biological assignment and Mendelian randomization utilize the Mendelian principle of random selection of genes from parent to offspring, the genetic epidemiology studies use the actual genotype of the patients to estimate the effect of a risk factor on disease, while biological assignment studies use matching of the sibling donor and recipient HLA (and not the specific HLA genes) to estimate the effect of the type of transplant

on outcome. HLA matching status itself is not anticipated to be biologically related to outcome. Biological assignment trials have been reported in a wide variety of diseases for which HCT is considered an acceptable treatment including myeloid and lymphoid leukemias, myelodysplastic syndromes and multiple myeloma [8-23]. However, biological assignment should not be equated with randomly assigning a treatment upon enrollment to a study.

In this paper, we discuss issues which may arise in biologic assignment trials including feasibility of patient accrual, sample size considerations, adjustment for prognostic factors, monitoring for enrollment biases, intention-to-treat analyses and ethics. These issues are illustrated using experiences from the Blood and Marrow Transplant Clinical Trials Network (BMT CTN) protocol #0102 entitled, "A Trial of Tandem Autologous Stem Cell Transplants +/- Post Second Autologous Transplant Maintenance Therapy Versus Single Autologous Stem Cell Transplant Followed by Matched Sibling Nonmyeloablative Allogeneic Stem Cell Transplant for Patients with Multiple Myeloma" [24]. In this trial, patients with an HLA-matched sibling donor were assigned to receive an autologous transplant (using one's own stem cells) followed by an allogeneic transplant (using stem cells from the donor) (auto-allo). Patients without an HLA-matched sibling donor were assigned to receive two sequential (tandem) autologous transplants (auto-auto). This study design is illustrated in Figure 1.

NEED FOR ALTERNATIVE DESIGNS IN CLINICAL TRIALS IN HEMATOPOIETIC STEM CELL TRANSPLANTATION

A substantial impediment to conducting definitive trials of HCT strategies is the fact that relatively few patients are eligible for this treatment. Of the approximately 20,000 new cases of multiple myeloma (the most common indication for autologous HCT) diagnosed yearly in the United States and 13,000 new cases of acute myelogenous leukemia (AML, the most common indication for allogeneic HCT) [25] only about 3,500 myeloma patients and 2,000 AML patients receive a transplant. The total number of patients undergoing HCT in North America is about 17,000 annually (7,000 allogeneic and 10,000 autologous) [1]. These procedures are done in more than 300 transplant centers.

Age is an important factor in successful HCT outcome. HCT, especially from allogeneic donors, is generally considered only for younger patients despite the fact that most patients (>70%) with hematologic malignancies are older than 50 years. Only 26 percent of autologous transplantations are in patients older than 60 years and only 19 percent of allogeneic transplants are in patients older than 50 years [1].

Also, allogeneic HCT is limited to those patients with suitable donors, ideally HLA-matched sibling donors. Reports in the literature indicate that 25-44% of otherwise eligible transplant candidates have an HLA-matched sibling donor [8-23], although the true rate may be lower because of some selection bias in these studies. There is increasing use of alternative donors such as HLA-mismatched family members or matched unrelated donors, although transplant-related mortality rates are higher with alternative donors. Banked unrelated donor umbilical cord blood cells are the newest source of alternative donor cells for HCT and may extend application of HCT to many patients previously unable to find a donor. Nevertheless, HLA-matched sibling transplants account for about 60% of all allogeneic transplantations [1]. Transplants from unrelated donors or mismatched family members comprise the remaining 40%.

ADVANTAGES OF BIOLOGIC ASSIGNMENT TRIALS

Feasibility of patient accrual in a limited patient population

When the patient population is limited, making randomized assignment impractical, a biologic assignment trial may be considered so that a sufficient number of trial subjects may be enrolled in a timely manner. As an example, consider the most common indication for allogeneic HCT, acute myeloid leukemia (AML) in first remission. There are only about 500 HLA-matched sibling transplants for AML in first remission yearly in the United States. Overall survival is about 60% at three years [26]. To test a new approach with 80% power to detect a 15% absolute improvement over that figure would require randomizing about 300 patients. To complete accrual in three years would require that about 20% of *all* AML patients receiving HCT in first remission be enrolled at a large number of transplant centers, which is far higher than the 3-5% of U.S. adult cancer patients that enroll on clinical trials in general [27-29]. For other less accepted or less common indications, accrual of patients is even more difficult. Consequently there is interest in alternatives to the randomized trial for studying HCT strategies. For the AML example, allowing all patients otherwise eligible but without a donor to be enrolled would increase the potential study population to about 1500 per year instead of 500 per year.

Sample size considerations and trial duration

Feasibility of accrual in a timely fashion was the primary reason biological assignment was used for protocol 0102. Potential accrual of multiple myeloma patients was first estimated from data collected by the Center for International Blood and Marrow Transplant Research (CIBMTR) which maintains an observational database of HCT outcomes.

The proportion of multiple myeloma patients with an HLA-matched sibling donor was estimated to be 20-30%, based on reports in the literature [8-23]. The total number of patients expected to enroll was 254 annually based on CIBMTR data and discounted estimates of accrual at participating transplant centers. The annual number of patients with a matched sibling donor expected to enroll was estimated to range from 51 (assuming a 20% matched sibling donor rate) to 76 (assuming a 30% rate) over the study period. Table 1 contrasts the sample sizes and years of accrual necessary for a two-arm biological assignment trial with a two-arm randomized trial as a function of the proportion of patients with an HLA-matched sibling donor. The randomized trial would only include patients with an HLA-matched sibling donor (randomized in a 1:1 ratio to auto-allo versus auto-auto transplantation).

The sample sizes were selected to have 90% power to detect a difference between three year progression-free survival (PFS) of 60% (auto-allo arm) vs. 45% (auto-auto arm) at the two-sided type I error rate of 5%, using a comparison of binomial proportions. Using binomial proportions for the sample size calculations avoids making assumptions about the shapes of the hazard functions of the two treatment groups. If it were believed that the hazard functions were approximately proportional, a sample size calculation based on the log-rank statistic would be more efficient. Such is not the case for allogeneic and autologous transplantation. The risk of early transplant-related mortality is higher after allogeneic HCT while the risk of later cancer-related mortality is higher after autologous HCT. Given these known differences in post-transplant hazard rates, we picked as an endpoint PFS at a time (3 years) when we expected most events would have occurred in both arms.

As shown in Table 1, when 25% of the patients have an HLA-matched sibling donor available, a biological assignment trial would require 615 patients (461 on the auto-auto arm, 154 on the auto-allo arm) and 2.4 years of accrual, while a randomized trial would require 462 patients (231 on each arm) and 7.3 years of accrual. Allowing three years of follow-up on each patient leads to a 5.4 year biologic assignment trial as compared to 10.3 years for a randomized trial.

The potential for advances in therapy and in understanding of prognostic factors over a 10 year period make it unlikely that the results of a trial of that duration would be relevant, hence the focus on minimizing the accrual period rather than the total sample size.

To ensure that an adequate proportion of patients were enrolled to the auto-allo arm (i.e., patients having matched-sibling donors), the protocol investigators required that 150 auto-allo patients be included in the trial. Table 2 gives the total number of patients required as a function of the proportion of patients enrolled in the auto-allo arm. As in Table 1, power is calculated to detect a difference between three year progression-free survival (PFS) of 60% (auto-allo arm) and 45% (auto-auto arm) at the two-sided type I error rate of 5%, using a comparison of binomial proportions.

Note that when the number of patients with an HLA-matched sibling donor enrolled on the study is fixed, the total sample size decreases as the proportion of patients with a HLA-matched sibling donor increases. A conservative approach to planning such a trial would be to assume that the proportion of patients with an HLA-matched sibling donor is at the lower end of the expected range which, for protocol 0102, was 20%.

We have demonstrated the power calculations to detect a 15% difference (in absolute terms) between the 3-year PFS of the auto-allo group and the auto-auto group. These calculations assume the two treatment groups would be compared according to the Intention-to-Treat (ITT) principle in which patient data are analyzed according to the treatment group to which they were assigned. The 15% difference takes into account that some of the HLA matched subjects might cross-over to the auto-auto group and that some patients might not receive the complete treatment course (treatment “dropouts”). An alternative approach could have postulated cross-over and dropout rates and corresponding attenuations to the 15% difference. We chose to avoid such an approach since we felt it difficult to estimate these unknown parameters.

Generating enthusiasm among physicians and patients for a trial

In some circumstances, physician and patient preconceptions that HLA-matched sibling HCT is optimal (and therefore should not be denied to patients who have such a donor) may further slow accrual to a randomized study or introduce selection bias in referrals for enrollment. Gray and Wheatley [2] argue that biologic assignment trials provide an ethical way to conduct a trial comparing allogeneic transplantation with a non-allogeneic transplant therapy even when the physicians do not have equipoise with regard to the efficacy of the treatment.

POTENTIAL SOURCES OF BIAS IN A BIOLOGICAL ASSIGNMENT TRIAL AND PROPOSED ADJUSTMENTS

Prognostic factor imbalance

The first type of bias is treatment group imbalance with respect to prognostic factors. There are several statistical methods to adjust for imbalances in known prognostic factors. These include analysis of covariance for continuous data, Cox proportional hazards regression [30] for time-to-event data, and logistic regression for dichotomous outcome data. One can also construct strata based on the various prognostic factors and conduct a stratified analysis, e.g., using a stratified log-rank test for time-to-event data or a Mantel-Haenszel test for dichotomous data. However, this stratification can be cumbersome and possibly statistically unstable for large numbers of stratification variables. A propensity score analysis may be a useful alternative when correcting for large numbers of prognostic factors [31,32]. Propensity scores can also be used to assess the amount of enrollment bias, discussed in the next section, by comparing the patient-by-patient probability of having a matched donor with whether such a donor was identified or not. Since biological assignment is implicitly related to family size (number of

siblings) as well as age (availability of siblings for donation), adjustment for these factors may be advisable [33]. Nevertheless, while statistical methods may adjust for known imbalances, they cannot account for unknown imbalances which are more likely when biological assignment is used as opposed to randomization.

Enrollment bias

In an HCT trial, physician or patient knowledge of HLA-matched sibling donor availability may differentially affect enrollment to the study based on disease stage or risk classification [34,35]. For example, in some cases, patients who are at high risk of relapse may be more likely to be enrolled if they have a donor and less likely to be enrolled if they do not because of a perceived lack of efficacy of the non-allotransplant therapy in high risk patients. Conversely, patients who are at low risk of relapse, older, and/or frail may be less likely to enroll on the study if they do have a donor because of the perceived toxicity burden of an allogeneic transplant. Such differential enrollment could lead to more high risk patients on the allogeneic transplant arm potentially reducing its apparent efficacy relative to the autologous arm. Similarly, younger and otherwise healthy patients may be enrolled on the allotransplant arm potentially reducing its apparent relative toxicity. Also, novel non-transplant (off-trial) therapies might be considered to offer enhanced therapeutic potential to subjects without donors while they may be less frequently considered for those with an available donor.

Ideally, to avoid these issues, one should enroll patients without knowledge of whether a donor is available. However, excluding patients whose donor determination, via tissue typing, is done prior to referral to a transplant center might limit the pool of eligible patients. In protocol 0102 we allowed patients to enroll regardless of whether the typing had been done previously. Patients were assigned to the auto-allo arm or auto-auto arm according to the availability of a matched sibling donor and had their outcome data analyzed as such regardless of whether they received their assigned treatment. This allowed for analyses to be undertaken following the ITT principle.

Note that if prior tissue typing is allowed, enrollment bias can also occur prior to referral of a patient to the transplant center. It is important to keep the point of study enrollment as close to the tissue typing process as possible to reduce this potential for referral bias. In the multiple myeloma population, most candidates for this trial would be referred to a transplant center to receive at least an autologous transplant. The HLA typing would most likely be done by the transplant center, so the transplant center was considered an appropriate point of enrollment. In the setting of AML, where non-transplant therapies are more widely used, HLA typing may be done by the oncologist prior to referral to the transplant center. Thus for an AML trial, trial enrollment should take place at the primary oncologist stage.

Careful consideration of possibilities of enrollment bias may lead one to modify the target population for the primary analysis. For example, in protocol 0102, the possibility was raised that patients with high risk of relapse of their myeloma based on known clinical prognostic factors may be preferentially enrolled to the allogeneic transplant arm due to physician and patient preferences. Consequently, while such patients were enrolled (expected to be 20% of the total population), the protocol specified that they would be excluded from the primary analysis.

Monitoring for enrollment bias

Acknowledging the potential for enrollment bias, in protocol 0102 we implemented a monitoring plan to assess enrollment on the study. To monitor for enrollment bias, we compared the characteristics of all transplanted patients from the CIBMTR database with those who were transplanted in the trial. In protocol 0102, each clinical center was required to register

consecutive transplant recipients with the CIBMTR. Data were collected on all transplanted patients regardless of whether they were enrolled in the protocol. These data were used to determine if the demographic composition, performance status, and disease risk of patients screened and entered in the trial were representative of the population of transplant patients treated at that center who were potentially eligible for trial entry. Comparisons between CIBMTR-registered patients and trial-enrolled patients were performed periodically to assess potential biases.

Additionally, each clinical center was monitored separately to ensure consistency with the expected proportions of patients with HLA-matched donors (20-30%). As HCT trial data are typically analyzed by treating centers as strata, a reasonable representation of each treatment arm at each center was needed. For the 0102 trial, the trial investigators were concerned about low enrollment to the auto-allo arm since the auto-auto arm was considered standard of care for myeloma patients without a donor. That is, the trial investigators were more concerned that a patient with a matched sibling donor would decline enrollment to the trial than a patient without a matched sibling donor. In other studies, it may be more appropriate to consider enrollment bias in either direction.

Consequently, some important safeguards were used to ensure adequate enrollment to the auto-allo arm. First, a minimum center enrollment requirement was used. Only centers expecting to enroll at least 20 standard risk patients were allowed to participate in protocol 0102. This ensured a high probability ($\geq 93\%$, based on estimated rates of HLA-matched donors) of having at least two patients at a particular center with an HLA-matched sibling donor, which in fact occurred at each center. Second, if a clinical center did not use an HLA-matched donor because they were “not eligible,” the donor file was reviewed by the Data and Coordinating Center (DCC) Protocol Officer to verify that the decision was within the guidelines of the protocol. Finally, a truncated Sequential Probability Ratio Test (SPRT) [36] was used for statistical monitoring of each center's accrual. The monitoring plan used for protocol 0102 is given in Table 3. This monitoring plan was based on an SPRT testing the null hypothesis that the proportion of enrolled patients with a matched sibling donor would be 25% against the alternative hypothesis that the proportion would be 10%, using a 10% type I error rate and a 20% type II error rate. The null hypothesis proportion was obtained by noting that if the auto-allo and auto-auto arms have the same true enrollment rates (i.e. the probability that a patient being treated at that center will enroll on the study), then the proportion of enrolled patients who have a sibling donor should be approximately the same as the biological assignment proportion in the population (assumed to be 25%). Only the lower stopping boundary for the number of patients with a matched sibling donor was used to focus on identifying lower enrollment rates among patients with a matched sibling donor. The expression for the lower stopping boundary is $S_n = -1.893 + 0.166 * n$, where n is the number of patients accrued by the center at a monitoring time. The operating characteristics of this monitoring rule for a given center depend on the number of patients enrolled at that center and can be determined by simulation. If a center was enrolling its patients at a true (i.e., asymptotic) rate of one auto-allo patient for every three auto-auto patients (corresponding to the null hypothesis), there would be a 5.5% chance of reviewing the center during the first 20 patients accrued by the center and a 10% chance of reviewing during the first 50 patients. If the center was biasing its enrollment so that its true enrollment was one auto-allo patient for every nine auto-auto patients (the alternative hypothesis), there would be a 48% chance of reviewing the center during the first 20 patients and a 90% chance of reviewing the center during the first 50 patients.

This alternative hypothesis corresponds to when the true enrollment rate for auto-allo patients out of all eligible patients at that center with a matched sibling donor is one-third the true enrollment rate of auto-auto patients out of all eligible patients at that center with no donor

available. It is important to note that because of the limited sample sizes at many centers, this monitoring rule is powered only to detect large deviations in a center's enrollment rates.

INTENTION TO TREAT ANALYSIS

For an intention to treat analysis, a randomized clinical trial defines the starting time as the time of randomization. However, when the patients are biologically assigned, the starting time is not unique and an appropriate choice is important to minimize bias.

Protocol 0102 compared strategies after an initial autologous transplant. The time of the initial autologous transplant was chosen as the starting time so that the intent to treat analysis would include patients irrespective of whether they received the assigned second transplant. With this starting time, an intention to treat analysis will be performed using the biologically assigned treatment arms. That is, if a patient enrolled on the study and had an eligible donor, they were regarded as being assigned to the auto-allo arm even if they did not receive an allogeneic transplant. This choice of starting time avoids the potential bias that could arise if there were differential acceptance of the second transplant procedure or if the time elapsed between the first and second transplants differed between the two treatment groups.

ETHICAL ISSUES IN BIOLOGIC ASSIGNMENT TRIALS

An ethical concern about a biological assignment trial is that the results may not be relevant to all patients in the trial. For example, if in protocol 0102 it were found that HLA-matched sibling HCT was more effective than autologous HCT, this would not have been directly relevant to the patients who did not have an HLA-matched sibling donor. We do note that allogeneic HCT is generally first tested using HLA-matched sibling donors who are historically associated with the lowest rates of transplant-related morbidity and mortality. Demonstration of efficacy with this type of donor gives impetus for subsequent testing with alternative donors, i.e., HLA-mismatched relatives or unrelated donors which are available for most patients. If, conversely, allogeneic transplantation were not found to be superior, there would be no impetus for such testing and auto-transplantation would be the appropriate therapy for all patients.

To further address the ethical concern that results might not be relevant to all patients in the trial, protocol 0102 randomized the auto-auto patients in a one-to-one ratio to observation or maintenance therapy after their second autologous transplant (Figure 1). Other earlier studies also randomized patients without sibling donors into two different autologous HCT treatments or autologous versus non-transplant treatment, thus, including the allogeneic arm, making a total of three treatment arms [17,20,23]. The advantage of a three arm design is that it satisfies the ethical concern about testing treatments which are directly applicable to all patients on the study, both with and without a matched sibling donor. However, it places an additional burden on the sample size. Because of the need to adjust for multiple pairwise comparisons, a three arm design would require a larger total sample size to detect differences between the arms than a two arm design. Other three-arm designs may be considered. For example, if there were no statistically significant difference between the arms without a sibling donor, one could combine those arms in a statistical test against the arm which had a sibling donor. Otherwise, if there were a statistically significant difference between the arms without a sibling donor, the better non-sibling donor arm could be compared to the sibling donor arm. This statistical strategy was taken in protocol 0102, but for the sake of brevity we omit the details. Whether an additional randomization is appropriate depends on the circumstances of the study including the ability to enroll subjects, the treatment being studied, the nature of the standard therapy, and the availability of an appropriate second question to test.

In protocol 0102, the randomization of the auto-auto arm took place prior to the initial autologous transplant. This is because some patients could drop out or die prior to the second

transplant or the post-second transplant chemotherapy treatment. Not including such patients in the analysis would make the results applicable only to patients who received the two autologous transplants. Randomization prior to the initial therapy makes the randomized comparison applicable to those initially eligible for the first of the two transplants. Wheatley [3] reports on two examples of studies [9,15] in which patients without matched sibling donors who were not randomized (33% and 24% respectively) were not included in the comparison to the patients with matched sibling donors. This may have biased the results.

CONCLUSIONS

A biological assignment trial can be a reasonable way to compare treatments which are biologically based, such as HLA-matched sibling allogeneic transplants, when the gold-standard randomized trial design is impractical or impossible. Implementing such a trial requires careful consideration of the ethical issues and potential biases yet gives the opportunity to answer questions that may otherwise not be feasible to answer. Such trials also provide opportunities for innovative designs, monitoring, and analyses to address these issues.

ACKNOWLEDGEMENTS

This work was supported by grant U01-HL-69294 from the National Heart, Lung, and Blood Institute and the National Cancer Institute (BRL, CB, MMH, ME, SC).

Grant support: U01-HL-69294 from the National Heart, Lung, and Blood Institute and the National Cancer Institute (BRL, CB, MMH, ME, SC).

REFERENCES

1. CIBMTR 2005 Summary Slides Part 1. <http://www.cibmtr.org/ABOUT/NEWS/2006May.pdf>, Accessed August 1, 2007
2. Gray RG, Wheatley K. How to avoid bias when comparing bone marrow transplant with chemotherapy. *Bone Marrow Transplantation* 1991;7:9–12. [PubMed: 1855097]
3. Wheatley K. Current Controversies: Which patients with acute myeloid leukaemia should receive a bone marrow transplantation?—A statistician's view. *British Journal of Haematology* 2002;118:351–356. [PubMed: 12139718]
4. Klein, JP.; Zhang, M-J. Statistical challenges in comparing chemotherapy and bone marrow transplantation as a treatment for leukemia. In: Jewell, NP., editor. *Lifetime Data: Models in Reliability and Survival Analysis*. Kluwer; the Netherlands: 1996. p. 175-185.
5. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology* 2003;32:1–22. [PubMed: 12689998]
6. Thomas DC, Conti DV. Commentary: The concept of 'Mendelian Randomization.'. *International Journal of Epidemiology* 2004;33:21–25. [PubMed: 15075141]
7. Clayton D, McKeigue PM. *Lancet* 2001;358:1356–1360. [PubMed: 11684236]
8. Burnett AK, Goldstone AH, Stevens RMF, et al. Randomized comparison of addition of autologous bone-marrow transplantation to intensive chemotherapy for acute myeloid leukaemia in first remission: results of MRC AML 10 trial. *Lancet* 1998;351:700–708. [PubMed: 9504514]
9. Cassileth PA, Harrington DP, Appelbaum FR, et al. Chemotherapy compared with autologous or allogeneic bone marrow transplantation in the management of acute myeloid leukemia in first remission. *NEJM* 1998;339:1649–1656. [PubMed: 9834301]
10. Gamberi B, Bandini G, Visani G, et al. Acute myeloid leukemia from diagnosis to bone marrow transplantation: experience from a single centre. *Bone marrow transplantation* 1994;14:69–72. [PubMed: 7951122]
11. Keating S, De Witte T, Suci S, et al. The influence of HLA-matched sibling donor availability on treatment outcome for patients with AML: an analysis of the AML 8A study of the EORTC

- Leukaemia Cooperative Group and GIMEMA. *British Journal of Haematology* 1998;102:1344–1353. [PubMed: 9753069]
12. Ljungman P, De Witte T, Verdonck L, et al. Bone marrow transplantation for acute myeloblastic leukaemia: an EBMT Leukaemia Working Party prospective analysis from HLA-Typing. *British Journal of Haematology* 1993;84:61–66. [PubMed: 8338779]
 13. Reiffers J, Stoppa AM, Attal M, et al. Allogeneic vs. autologous stem cell transplantation vs chemotherapy in patients with acute myeloid leukemia in first remission: the BGMT 87 study. *Leukemia* 1996;10:1874–1882. [PubMed: 8946925]
 14. Wahlin A, Brinch L, Hornsten P, et al. Outcome of a multicenter treatment program including autologous or allogeneic bone marrow transplantation for de novo acute myeloid leukemia. *European Journal of Haematology* 1997;58:233–340. [PubMed: 9186533]
 15. Woods WG, Neudorf S, Gold S, et al. A comparison of allogeneic bone marrow transplantation, autologous bone marrow transplantation, and aggressive chemotherapy in children with acute myeloid leukemia in remission: a report from the Children's Cancer Group. *Blood* 2001;97:56–62. [PubMed: 11133742]
 16. Zittoun RA, Mandelli F, Willemze R, et al. Autologous or allogeneic bone marrow transplantation compared with intensive chemotherapy in acute myelogenous leukemia. *NEJM* 1995;332:217–223. [PubMed: 7808487]
 17. Burnett AK, Wheatley K, Goldstone AH, et al. The value of allogeneic bone marrow transplant in patients with acute myeloid leukaemia at differing risk of relapse: results of the UK MRC AML 10 trial. *British Journal of Haematology* 2002;118:385–400. [PubMed: 12139722]
 18. Mohty M, de Lavallade H, Ladaïque P, et al. The role of reduced intensity conditioning allogeneic stem cell transplantation in patients with acute myeloid leukemia : a donor vs. no donor comparison. *Leukemia* 2005;19:916–920. [PubMed: 15843817]
 19. Oosterveld M, Suci S, Verhoef G, et al. The presence of an HLA identical sibling donor has no impact on outcome of patients with high-risk MDS or secondary AML (sAML) treated with intensive chemotherapy followed by transplantation: results of a prospective study of the EORTC, EBMT, SAKK, and GIMEMA Leukemia Groups (EORTC study 06921). *Leukemia* 2003;17:859–868. [PubMed: 12750698]
 20. Rowe JM, Buck G, Burnett AK, et al. Induction therapy for adults with acute lymphoblastic leukemia: results of more than 1500 patients from the international ALL trial: MRC UKALL XII/ECOG E2993. *Blood* 2005;106:3760–3767. [PubMed: 16105981]
 21. Hehlmann R, Berger U, Pfirrmann M, et al. Drug treatment is superior to allografting as first-line therapy in chronic myeloid leukemia. *Blood* 2007;109:4686–4692. [PubMed: 17317858]
 22. Bruno B, Rotta M, Patriarca F, et al. A comparison of allografting with autografting for newly diagnosed myeloma. *NEJM* 2007;356:1110–1120. [PubMed: 17360989]
 23. Garban F, Attal M, Michallet M, et al. Prospective comparison of autologous stem cell transplantation followed by dose-reduced allograft (IFM99-03 trial) with tandem autologous stem cell transplantation (IFM99-04 trial) in high-risk de novo multiple myeloma. *Blood* 2006;107:3474–3480. [PubMed: 16397129]
 24. <https://web.emmes.com/study/bmt/protocol/0102_protocol/0102_protocol.html>, last accessed Aug. 1, 2007
 25. <http://www.seer.cancer.gov>; Last accessed Dec. 3, 2007
 26. CIBMTR 2005 Summary Slides Part 2. <http://www.cibmtr.org/ABOUT/NEWS/Newsletter2006v12-2.pdf>, accessed August 1, 2007
 27. National Cancer Institute Cancer Clinical Trials: The Basic Workbook. Available from: <http://www.cancer.gov/clinicaltrials/resources/basicworkbook/>
 28. Baquet CR, et al. Recruitment and participation in clinical trials: sociodemographic, rural/urban, and health care access predictors. *Cancer Detection and Prevention* 2006;30(1):24–33. [PubMed: 16495020]
 29. Peppercorn JM, et al. Comparison of outcomes in cancer patients treated within and outside clinical trials: conceptual framework and structured review. *Lancet* 2004;363(9405):263–270. [PubMed: 14751698]
 30. Cox DR. Regression models and life tables (with discussion). *JRSS-B* 1972;34:187–220.

31. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
32. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *JASA* 1984;79:516–524.
33. Serna DS, Lee SJ, Zhang M-J, Baker KS, Eapen M, Horowitz MM, Klein JP, Rizzo JD, Loberiza FL. Trends in survival rates after allogeneic hematopoietic stem-cell transplantation for acute and chronic leukemia by ethnicity in the United States and Canada. *Journal of Clinical Oncology* 2003;21:3754–3760. [PubMed: 14551294]
34. Ellenberg JH. Selection bias in observational and experimental studies. *Statistics in Medicine* 1994;13:557–567. [PubMed: 8023035]
35. Berger VW, Exner DV. Detecting selection bias in randomized clinical trials. *Controlled clinical trials* 1999;20:319–327. [PubMed: 10440559]
36. Wald, A. *Sequential Analysis*. Wiley; New York: 1947.

Figure 1.
Illustration of study design for BMT CTN protocol 0102.

Table 1

Anticipated annual enrollment of multiple myeloma patients to BMT CTN protocol 0102, required sample size, and total years of accrual for a biologic assignment trial compared to a randomized trial which can only randomize patients with a matched-sibling donor. Figures are given for various proportions of patients with an HLA-matched sibling donor. The sample sizes required correspond to 90% power to detect a difference between the auto-allo arm and the auto-auto arm with respect to three year progression-free survival (PFS) rates of 60% (auto-allo) vs. 45% (auto-auto). A two-sided type I error rate of 5% is used with no correction for interim analyses.

% with HLA-Matched Sibling Donor	Trial Design	Annual enrollment	Sample Size required	Years of Accrual
20%	Biological Assignment (all patients)	254	720	2.8
	Randomized trial (matched sibling donors only)	51	462	9.1
25%	Biological Assignment (all patients)	254	615	2.4
	Randomized trial (matched sibling donors only)	63	462	7.3
30%	Biological Assignment (all patients)	254	550	2.2
	Randomized trial (matched sibling donors only)	76	462	6.1

Table 2

Total sample size and power with respect to the proportion of patients with a matched sibling donor (i.e., patients assigned to the auto-allo arm). The number of auto-allo patients is fixed at 150. Power is calculated to detect a difference between the auto-allo arm and the auto-auto arm with respect to three year progression-free survival (PFS) rates of 60% (auto-allo) vs. 45% (auto-auto). A two-sided type I error rate of 5% is used with no correction for interim analyses.

Proportion with Matched Sibling Donor	Number on auto-allo arm	Number on auto-auto arm	Total n	Power
0.20	150	600	750	91%
0.25	150	450	600	89%
0.30	150	350	500	87%

Table 3

Monitoring plan for the proportion of patients enrolled to the auto-allo arm at a particular transplant center. Monitoring plan is based on a Sequential Probability Ratio Test testing the null hypothesis that the true (i.e., asymptotic) proportion of enrolled patients with a matched sibling donor is 25% against the alternative hypothesis that the proportion is 10%. A 10% type I error rate and a 20% type II error rate are used.

<u>Total Enrolled</u>	<u>Enrolled to the auto-allo arm</u>
0 - 11	No review
12 - 17	Review if 0 patients enrolled to the auto-allo arm
18 - 23	Review if 1 or fewer patients enrolled to the auto-allo arm
24 - 29	Review if 2 or fewer patients enrolled to the auto-allo arm
30 - 35	Review if 3 or fewer patients enrolled to the auto-allo arm
36 - 41	Review if 4 or fewer patients enrolled to the auto-allo arm
42 - 47	Review if 5 or fewer patients enrolled to the auto-allo arm
48 - 53	Review if 6 or fewer patients enrolled to the auto-allo arm