Design Issues in a Prospective Randomized Double-Blinded Trial of Prophylaxis with Fluconazole versus Voriconazole after Allogeneic Hematopoietic Cell Transplantation

John R. Wingard
Department of Medicine, Blood and Marrow Transplant Program, and Clinical and Translational Research, Division of Hematology/Oncology, University of Florida Shands Cancer Center, Gainesville, Florida

Background. Aspergillus infections pose the toughest infectious challenges to the clinician caring for hematopoietic cell transplant recipients. About 15% of patients become infected, with a case fatality rate of ∼65%. To date, no effective prophylactic strategies have been developed.

Methods. Voriconazole, a recently licensed extended-spectrum azole, with demonstrated efficacy against aspergillus, is currently being tested as a potential prophylactic agent against aspergillus and other invasive fungal infections. Logistic issues—such as patient selection, choice of comparator, blinding of study drugs, duration of study drug administration, and how to handle empirical amphotericin B for possible invasive fungal infections—and analytic concerns, including choice and definition of the primary end point and the potential confounding effect of informative censoring (as a result of noninfectious events), were considered in the design of the clinical trial.

Results. The trial is now under way, with a projected 3-year enrollment period.

Conclusions. Each design decision shaped the trial in a way that permitted certain questions to be answered while not allowing others to be addressed. Once completed, the trial’s results must be interpreted in light of these design details.

Invasive fungal infections (IFIs) are frequent among hematopoietic cell transplant recipients and occur both during the several weeks before engraftment (mostly Candida) and during the several months after engraftment (mostly Aspergillus) [1, 2]. The current practice of fluconazole prophylaxis embraced by the consensus guidelines of the Centers for Disease Control and Prevention, the Infectious Diseases Society of America, and the American Society of Blood and Marrow Transplantation [3], addresses Candida, but there is no proven prophylaxis against Aspergillus infections. Aspergillus infections are occurring at increasing rates [4]. Mortality from Aspergillus infections is high, and, of all patient risk groups, HCT recipients have the highest mortality from Aspergillus [5]. Voriconazole, an extended-spectrum azole, available in both oral and intravenous formulations, is an effective treatment for aspergillosis, but even with voriconazole, the treatment results in patients undergoing allogeneic hematopoietic cell transplantation (HCT) still remain suboptimal, with only one-third of patients surviving infection [6].

These considerations indicate that current antifungal strategies for patients undergoing allogeneic HCT are at best suboptimal. These provide ample justification emphasizing the importance of a prophylaxis trial testing an agent with antimold activity.

SYNOPSIS OF THE TRIAL

A prospective, multicenter, randomized, double-blind trial to compare the effectiveness of voriconazole prophylaxis with fluconazole prophylaxis to improve survival free from probable or proven invasive fungal infection (“fungus-free survival”) in allogeneic HCT recipients has been designed by a protocol team of the newly created Blood and Marrow Transplant Clinical...
Trials Network, a consortium sponsored by the National Heart, Lung, and Blood Institute and the National Cancer Institute. The protocol team consists of John Wingard and Tom Walsh, cochairs, and Pat Brennecke, Shelly Carter, Dennis Confer, Iris Gersten, Joanne Kurtsberg, Brent Logan, Kieren Marr, and Trudy Small. The hypothesis to be tested is that prophylaxis with voriconazole, a mold-active azole proven to be an effective treatment for aspergillosis [6], will be associated with fewer IFIs and will reduce morbidity and mortality due to IFI. This will result in an improvement in fungus-free survival. Patients who are eligible must be aged ≥2 years, be recipients of both a myeloablative conditioning regimen and a stem cell graft (from blood, marrow, or, in children <12 years old, cord blood) from a 6 of 6 antigen donor (donor and recipient must be identical at HLA-A, HLA-B, and HLA-DR), and be undergoing transplantaion because of acute or chronic leukemia in remission or because of myelodysplastic syndrome in a low or intermediate prognostic category. Study drug will be administered from day 0 to day 100 blindly to all subjects, and for those who are receiving steroids at a dose of at least 1.0 mg/kg/day at day 100 to day 0, fluconazole at 400 mg once daily and voriconazole at 200 mg twice daily for adults. Lower doses will be given to children <12 years old. IFI will be scored by means of the consensus criteria of the European Organization for Research on Treatment of Cancer (EORTC) and the Mycoses Study Group (MSG) of the National Institute for Allergy and Infectious Diseases [7].

Galactomannan testing [8] will be done twice weekly from baseline to day 60 for all subjects; after day 60, for those who have experienced acute graft-versus-host disease, twice-weekly screening will continue until day 100, whereas others will be tested once weekly until day 100. The results will be used along with host and clinical criteria as an adjunct to the diagnosis of probable Aspergillus infection. The primary end point will be fungus-free survival at 180 days. A blinded data review board will review the end point data without knowledge of toxicity and will independently adjudicate the end points. A total of 600 subjects will be enrolled, a sample sufficient to detect a 12% difference in fungus-free survival with an α of 0.05 and a β of 0.80. The protocol in its entirety can be viewed at the Blood and Marrow Transplant Clinical Trials Network Web site (http://bmtctn.net).

**DESIGN ISSUES**

**Which patients should be enrolled in the trial?** For efficiency, it was recognized that the patients at highest risk for IFI should be of prime consideration. The problem faced by the protocol team was that the same factors that put patients at risk for IFI also put them at risk for other serious complications associated with transplantation, such as severe graft-versus-host disease and serious organ toxicity. If the nonfungal complications lead to early death, the possibility exists that confounding events could negate the ability to determine the end points of interest and could confound being able to detect a benefit in either arm. The guiding principle of the protocol team was to reduce heterogeneity of potentially confounding factors and to achieve the best balance between minimizing death from nonfungal events and choosing patients with sufficient risk for infection.

We excluded uncommon diseases, “high-risk” disease states, and mismatched grafts. Nonmyeloablative conditioning regimens were also excluded because this type of transplantation was still being unevenly applied in clinical practice, and no reliable data were available to judge risks of either fungal or nonfungal events. Using diagnoses for which patients commonly underwent transplantation, such as diseases in an early remission status, and donors well matched with recipients as the general patient selection criteria, we surveyed 2 large data sets from the Fred Hutchinson Cancer Research Center and the International Bone Marrow Transplant Registry to assess the risk of IFI and nonfungal mortality. The rates of IFI at 180 days for the 2 data sets were 14% and 15%, respectively. The nonfungal mortality rates at 180 days were 13% and 19%, respectively. Exploratory analyses demonstrated differences in patient mix in the 2 data sets, with a higher proportion of patients with chronic myelogenous leukemia in the Fred Hutchinson Cancer Research Center data set (a group with a lower risk for nonfungal mortality), which likely explain much of the difference in nonfungal mortality. Because fewer transplantations for chronic myelogenous leukemia are being done since the introduction of imatinib, it was believed that this group would constitute only a small proportion of expected enrollees.

The protocol team decided that these eligibility criteria represented a suitable compromise: the signal (the rate of IFI) was sufficiently loud and the background noise (the rate of nonfungal mortality) low enough not to drown out the signal. With continued increases in IFI rates, changes in the case mix, a decrease in the proportion of patients at low risk for IFI, and the use of galactomannan testing as an adjunctive diagnostic—which would convert some determinations of “possible” infection to “probable” infection—it was thought that reasonable estimates of IFI rates were 15%–20% among the patients entered in the trial.

**What should the comparator be?** The reference standard is fluconazole, shown in several studies to provide excellent protection against Candida infections, and its use outside the trial is regarded as appropriate by consensus guidelines of the Centers for Disease Control and Prevention, the Infectious Disease Society of America, and the American Society of Blood and Marrow Transplantation [3]. The concern was whether its lack of antifungal activity would be a fair comparison and whether clinicians...
would embrace a study in which one arm offered antimold coverage and the other arm did not.

The team thought that 2 features of the trial offered considerable protection to subjects. First, incorporation of galactomannan assay for subjects would allow detection of *Aspergillus* infection earlier than by traditional clinical diagnostic measures, as suggested from earlier studies in this patient population [8], and would allow prompt institution of treatment, thereby potentially improving outcome. This, incidentally, adds an important dimension to the study: it is entirely possible that voriconazole would prevent more infections than fluconazole, as hypothesized, yet early therapy prompted by the galactomannan assay for fluconazole recipients could reduce mortality from *Aspergillus* infection, and although more infections might be seen, an excess of deaths could be avoided. Another way to view this would be a comparison of antimold prophylaxis (the voriconazole arm) versus vigorous monitoring and early antimold therapy (the fluconazole arm). Second, empirical trials of amphotericin B (the type of formulation would be left to the discretion of the investigator) would be permitted for patients with possible IFI, with criteria adapted from the consensus criteria of EORTC and MSG [7]. To avoid the possibility that excessive use of empirical amphotericin B might dampen the ability to detect a difference between study arms, limits were placed (no more than 2 weeks of empirical therapy); the team believed that this would provide for patient safety and would allow sufficient time for the clinician to perform the needed diagnostic testing to determine the cause of fever, pulmonary infiltrates, and other signs and symptoms that suggested IFI but might be due to another infectious pathogen or other noninfectious cause. Another consideration led the protocol team to believe that equipoise is present: it is possible that a reduction in IFIs by voriconazole may be offset by an increase in toxicity.

A second comparator considered for the study was itraconazole. Two small trials have tested this agent in the HCT setting [9, 10]. In both studies, prolonged courses of either itraconazole or fluconazole were given for ~3 months to patients undergoing allogeneic HCT, similar in duration to this proposed trial. In a trial by Winston et al. [9], a lower rate of IFI was noted in patients receiving itraconazole (9% vs. 25%). Because of the trial’s small size, however, the rates of *Aspergillus* infections were not significantly different, although there was a trend in favor of itraconazole; breakthrough *Aspergillus* infections tended to occur in patients with lower blood levels of itraconazole. A concern raised by this trial was a higher rate in the itraconazole arm of study drug discontinuation due to death (30% vs. 18%) or adverse reactions (6% vs. 1%). The second, much larger, trial by Marr et al. [10] dealt with the variable bioavailability of itraconazole by increasing the dose and the number of doses per day. Unfortunately, substantially greater toxicity was encountered in the itraconazole arm, with an excess of patients in that arm having the study drug discontinued or because of a greater rate of renal and hepatic toxicity, and these safety issues led to premature stopping of the trial. Overall, there was no significant reduction in rate of IFI in the itraconazole arm, although a subset analysis showed that among patients who could tolerate the drug at this dose schedule there was a reduction in IFIs. With the uncertain benefit and potential risks of itraconazole seen in these data (none yet published at the time of consideration), the team thought that fluconazole was the most appropriate comparator.

A third comparator that could have been used is one of the amphotericin B formulations. The deoxycholate formulation is poorly tolerated by patients concomitantly taking calcineurin inhibitors. The lipid formulations are better tolerated, but the lack of an oral formulation was thought to pose an enormous practical disadvantage for long-term use.

**Can and should the study drugs be blinded?** Both voriconazole and fluconazole affect cytochrome P-450 activity and potentiate levels of cyclosporine and tacrolimus, but voriconazole has a greater degree of interaction with cyclosporine and tacrolimus [11–13]. The question arose as to whether safety might be a concern or whether high levels of the immunosuppressive drugs in one arm could unmask the identity of the study drug. It was thought that close monitoring of the immunosuppressive drugs would optimize safety. Moreover, the considerable heterogeneity of cyclosporine and tacrolimus levels at any given dose schedule, even in the absence of the study drugs, means that close monitoring of the immunosuppressive drugs and dose adjustment are needed anyway, and variability of immunosuppressive drug levels was unlikely to pose a substantive incremental safety issue or to negate the blinding. Another concern regarding blinding is the unique visual toxicity of voriconazole, photopsia, which does not occur with fluconazole. This concern was less easily disposed of, but similar differential toxicities have been faced by other blinded antifungal trials, such as comparisons of different amphotericin B formulations, without serious impediment. Fortunately, only a minority of subjects experience this nonserious side effect. Blinding was judged important scientifically to minimize bias, because empirical amphotericin B trials for “possible cases of IFI” were permitted, as noted above, and the protocol team feared that clinicians might more readily start empirical treatment with amphotericin B or continue it longer for patients on the fluconazole arm, which could potentially confound the ability to detect differences between the study drugs. Thus, it was decided that blinding both could and should be done. To eliminate bias, a blinded data review board will review the end.
point data without knowledge of toxicity and will independently adjudicate the end points.

What is the proper duration of study therapy? Some antifungal prophylaxis trials administered the test drug, generally fluconazole or itraconazole, during the neutropenic episode and stopped the trial at time of neutropenic recovery. However, it was recognized that for prevention of *Aspergillus* infection as a major target, this would not be an adequate duration, because most *Aspergillus* infections occur after engraftment [14, 15]. The 2 above-mentioned studies of itraconazole in patients undergoing HCT performed the prophylaxis for 100 days after HCT, in recognition that *Aspergillus* infections now typically occur 2–3 months after bone marrow transplantation and even later in some patients [14, 15]. Accordingly, it was decided that study drug would be continued routinely for all patients until day 100. At day 100, a decision would be made as to the likelihood of continuing risk for IFI. For patients with acute graft-versus-host disease still receiving at least 1.0 mg/kg/day prednisone equivalent or for recipients of T cell–depleted donor graft-versus-host disease still receiving at least 1.0 mg/kg/day prednisone equivalent or for recipients of T cell–depleted donor grafts in whom CD4+ cell counts were still <200 cells/μL, study drug would be continued until day 180, the end of the study period. It was recognized that some cases of chronic graft-versus-host disease could occur after day 100, requiring additional immunosuppression that might place some patients at late risk for invasive aspergillosis; the practical hurdles of restarting study drugs beyond day 100, when most patients are no longer being seen at regular intervals at the study sites, led us to believe that the difficulties in adherence to protocol would be insurmountable.

Should an empirical trial of amphotericin B be scored as a failure of prophylaxis? It could be argued that the most rigorous test of a prophylaxis strategy would be to not allow any empirical trials of amphotericin B at all. The protocol team agreed with this proposition. This was not deemed practical, however, because there are gaps in the spectrum of activity of both agents being tested and because delays in treatment are associated with poor outcome, patient safety considerations led the team to decide that empirical amphotericin B had to be permitted, but the indications for both initiation and duration of use would be restricted. Given that decision, we could not escape the dilemma of how to analyze these amphotericin B trials.

Cogent arguments can be made for and against including amphotericin B trials in the definition of failure. On the one hand, a truly effective preventive measure should not require rescue, and empirical therapy with another agent should not be necessary. On the other hand, the causes of fever, pulmonary infiltrates, and the other clinical signs and symptoms that lead a clinician to prescribe empirical antifungal therapy for “possible” fungal infection are manifold; fungal infections can be documented in only a minority of such situations. In fact, the unexplained syndromes and those with nonfungal causes are more frequent than the syndromes found to be due to fungi. For example, in a study of fluconazole prophylaxis for patients undergoing HCT [16], a fungal etiology could be established in only 1% of all the empirical amphotericin B trials. Thus, counting these cases of “possible” IFI as failures would be an overestimate of events and could be misleading or, worse, could obscure a benefit of one arm of the study over the other. In the end, the protocol team was persuaded that fungal events should meet the EORTC and MSG criteria [7] for probable or proven IFI, and “possible” infections or empirical trials of amphotericin B would not. This decision conformed with the trials of prolonged itraconazole prophylaxis by Winston et al. [9] and Marr et al. [10], described above.

Ideally, it would have been nice to expand the sample size to have sufficient power to examine efficacy in subsets with and without empirical therapy (to ensure that the empirical trials did not contaminate the rates of the primary end points); however, the cost proved to be prohibitive. Notwithstanding, the rates of empirical therapy will be examined in secondary analyses with respect to effects on IFI and fungus-free survival.

What should the primary end point be? Clearly, the foremost question of interest is whether one of the study drugs can reduce the rate of IFI compared with that seen with the other. Thus, the straightforward approach would be to use the rate of IFI as the primary end point. However, the possibility was recognized that an interaction of one of the study drugs with immunosuppressive drugs or some other aspect of the transplantation process could lead to a difference in toxicity or mortality, which could result in informative censoring. Such an occurrence could lead to an erroneous interpretation. Therefore, it was thought that the more rigorous method to test the relative merits of the 2 study drugs was to use fungus-free survival (freedom from both fungal infection and death) as the primary end point. The validity of this end point relies on the randomization process to ensure equality of death from causes unrelated to the study drugs in the 2 groups. This hinges on the trial’s vital initial consideration: Who should be eligible for the trial? With reduced heterogeneity, it is hoped that an accidental imbalance between the 2 groups will be unlikely.

**CONCLUSIONS**

This trial posed a number of practical and analytic issues. The investigative team faced concerns about logistics (patient selection, choice of comparator, blinding of study drugs, duration of study drug, and use of empirical amphotericin B) and analysis concerns (choice and definition of the primary end point, informative censoring). Each design decision shaped the trial in a way that permitted certain questions to be answered and others not. Attention to the details of the design is necessary...
to interpret the results appropriately and to translate the results into practice.

Acknowledgments

Financial support. Grant U01HL69301 from the National Institutes of Health, and grants from Pfizer, Merck, and Schering.

Potential conflicts of interest. J.R.W. has received lecture fees and has been a consultant for Pfizer, Merck, and Enzo.

APPENDIX A

THE BLOOD AND MARROW TRANSPLANT CLINICAL TRIALS NETWORK

The Blood and Marrow Transplant Clinical Trials Network is funded by the National Heart, Lung, and Blood Institute and by the National Cancer Institute. The protocol team members are John Wingard and Thomas Walsh, cochairs, and Pat Brennecki, Shelly Carter, Dennis Conver, Iris Gersten, Joanne Kertsberg, Brent Logan, Kieren Marr, and Trudy Small, members. Participating centers are below.

Core Study participants: Case Western Reserve University Consortium (Billard Lazarus); Oregon Health and Science University; St. Louis Children’s Hospital; University Hospitals of Cleveland; Washington University; City of Hope National Medical Center (Stephen Forman); Dana-Farber/Partners Cancer Center (Joseph Antin); Duke University Medical Center (Joanne Kertsberg); Fred Hutchinson Cancer Research Center (Frederick Appelbaum); Johns Hopkins University (Richard Jones); Memorial Sloan-Kettering Cancer Center (Richard O’Reilly); Pediatric Blood and Marrow Transplant Consortium (Alan Gamis); Cardinal Glennon Children’s Hospital, St. Louis University; Children’s Mercy Hospital, University of Missouri; Children’s Memorial Hospital, Chicago; Children’s National Medical Center, Washington, DC; Hackensack University Medical Center (pediatric service), Primary Children’s Medical Center, University of Utah; Texas Transplant Institute; University of Rochester Medical Center; Vanderbilt University Medical Center; Stanford Hospital and Clinics (Robert Negrin); University of California, San Diego/Scripps School of Medicine (Edward Ball); University of Florida College of Medicine (John Wingard); University of Michigan Medical Center (James Ferrara); University of Minnesota (Daniel Weisdorf); University of Nebraska Medical Center (Julie Vose); University of Pennsylvania Hospital (Edward Stadtmauer); and University of Texas, M. D. Anderson Cancer Research Center (Sergio Giralt).

Non–Core Study participants: Indiana University Medical Center; Kansas City Cancer Centers, Central Bone Marrow Transplant; Hackensack University Medical Center (adult medical services); Utah Bone Marrow Transplant, University of Utah; Mount Sinai Medical Center; University of Alabama at Birmingham; and Wake Forest University Baptist Medical Center.

References