
John E. Bennett, 1 John Powers, 2 Thomas Walsh, 3 Claudio Viscoli, 9 Ben de Pauw, 4 William Dismukes, 4 John Galgiani, 1 Michel Glauser, 10 Raoul Herbrecht, 11 Carol Kauffman, 6 Jeannette Lee, 4 Peter Pappas, 4 John Rex, 7 and Paul Verweij 8

1 Clinical Mycology Section, Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, and 2 National Cancer Institute, National Institutes of Health, Bethesda, Maryland; 3 Division of Special Pathogen and Immunologic Drug Products, US Food and Drug Administration, Rockville, Maryland; 4 University of Alabama at Birmingham; 5 VA Medical Center and University of Arizona, Tucson; 6VA Medical Center and University of Michigan, Ann Arbor; 7 University of Texas Medical School, Houston; 8 University Hospital St. Radboud, Nijmegen, The Netherlands; 9 University of Genoa, National Institute for Cancer Research, Genoa, Italy; 10 University Hospital, Lausanne, Switzerland; and 11 Hôpital de Hautepierre, Strasbourg, France

There is inferential evidence that some patients with prolonged neutropenia and fever not responding to antibacterial agents are at sufficient risk of deep mycoses to warrant empirical therapy, although superiority of an antifungal agent over placebo has not been conclusively demonstrated. Amphotericin B deoxycholate, liposomal amphotericin B, and intravenous itraconazole followed by oral itraconazole solution are licensed in the United States for this indication. Fluconazole and voriconazole have given favorable results in clinical trials of patients with low and high risk of deep mold infections, respectively. Design features that can profoundly influence outcome of empirical trials are (1) inclusion of low-risk patients, (2) failure to blind the study, (3) obscuration of antifungal effects by changing antibacterial antibiotics, (4) failure to balance both arms of the study in terms of patients with prior antifungal prophylaxis or with severe comorbidities, (5) the merging of end points evaluating safety with those of efficacy, and (6) choice of different criteria for resolution of fever.

BASIS FOR THE CONCEPT OF EMPIRICAL ANTIFUNGAL THERAPY

The practice of administering empirical antifungal therapy to persistently febrile neutropenic patients who are apparently not responding to antibacterial therapy has become a standard of care in many centers. Although this practice is based on sound theoretical principles, the actual data from clinical trials supporting the use of empirical antifungal therapy in the management of infections in neutropenic patients are sparse. The initial suggestions of the possible benefits of an empirical antifungal therapy approach came from autopsy and clinical studies showing an increasing incidence of deep fungal infections, many recognized only at autopsy, in patients with cancer [1]. Clinical observations also suggested the importance of early intervention in the successful treatment of fungal infections. The difficulty in making an antemortem diagnosis of invasive fungal infection, given the limitations of current diagnostic techniques, necessitates an empirical approach to drug therapy.

On the basis of these observations, Pizzo et al. [2], at the National Cancer Institute, performed the first randomized trial of empirical antifungal therapy in treating persistently febrile neutropenic patients not re-
sponding to antibacterial therapy. They studied 50 patients with fever and neutropenia persisting after 7 days of empirical antibacterial therapy and without documented infection. Patients were randomized to discontinue all antibiotic treatments (16 patients), to continue the initial antibacterial treatment (16 patients), or to add empirical amphotericin B deoxycholate (18 patients). When results for the 2 groups not receiving amphotericin B deoxycholate were combined, 6 patients developed fungal infections—4 with deep fungal infections and 2 with more-superficial disease. Among the 18 patients receiving amphotericin B deoxycholate, there was only one severe fungal infection, and it was caused by Pseudallescheria boydii, a species resistant to amphotericin B. The difference in breakthrough deep fungal infections did not reach statistical significance in this small study (6/32 vs. 1/18; \( P > .10 \)).

Subsequently, Françoise Meunier and the International Antimicrobial Therapy Cooperative Group of the European Organization for the Research and Treatment of Cancer (EORTC) performed another trial of empirical antifungal therapy among 132 persistently febrile and neutropenic cancer patients not responding to 4 days of treatment with empirical antibacterial therapy. The investigators randomized patients to receive empirical amphotericin B deoxycholate or to continue antibacterial therapy without modification [3]. The primary end point for this trial was resolution of fever during 5 days of amphotericin B deoxycholate treatment: 69% of patients in the amphotericin B deoxycholate arm and 53% in the placebo arm had resolution of fever, a difference that was not statistically significant. In a subset analysis, there were no deaths due to fungal infection among patients receiving empirical amphotericin B deoxycholate, compared with 4 deaths in the group not receiving antifungal therapy (\( P = .05 \)). In addition, the number of documented fungal infections was higher among patients not receiving amphotericin B deoxycholate (6 vs. 1; \( P = .056 \), Fisher’s exact test). Multivariate analysis with linear logistic regression showed that the addition of amphotericin B deoxycholate correlated with defervescence in adults with hematologic malignancies who were not receiving antifungal prophylaxis and who were severely neutropenic (\( P = .03 \)). These results led to the suggestion that empirical antifungal therapy probably should be reserved for selected groups of high-risk patients.

At the time these studies were conducted, they were considered to be at the leading edge of antifungal therapy research in cancer patients. Now it is apparent that several factors complicate the conclusions of these trials. Neither of the studies was blinded. Both trials were statistically underpowered to determine differences in both the primary end points and the subgroup analyses. Although the optimal time at which empirical antifungal therapy should be started remains undetermined, many experts recommend waiting until the fifth or seventh day of persistent fever and neutropenia before initiating treatment [4].

**DRUGS FOR EMPIRICAL ANTIFUNGAL THERAPY**

A drug for use for empirical antifungal therapy in treating febrile neutropenic patients should have demonstrable activity against the fungal pathogens most likely to infect these patients, namely *Candida* and *Aspergillus* species. Amphotericin B deoxycholate has been the drug of choice. However, nephrotoxicity and infusion-related reactions are important limiting factors with this drug. Concerns with the toxicity of amphotericin B deoxycholate have led to the study of 3 triazoles and 2 lipid formulations of amphotericin B for empirical antifungal therapy to treat febrile neutropenic patients.

In 1996, Viscoli et al. [5] reported the first randomized trial of fluconazole as empirical antifungal therapy. They investigated a selected population of 112 patients who were not receiving fluconazole prophylaxis and who were at low risk for invasive aspergillosis. Response to fluconazole was not more than 10% worse than that to amphotericin B deoxycholate. In a more recent study of 317 patients, fluconazole was not found to be inferior to amphotericin B deoxycholate as empirical antifungal therapy [6].

Investigators also have studied triazoles with greater activity against *Aspergillus* species in empirical antifungal therapy trials. Recent trials with these triazoles have compared itraconazole with amphotericin B deoxycholate [7] and voriconazole with liposomal amphotericin B [8]. Two randomized double-blind trials have compared efficacy of amphotericin B formulations as empirical therapy for persistent fever in neutropenic patients. A study by Walsh et al. [9] found liposomal amphotericin B to be comparable in efficacy but less toxic than amphotericin B deoxycholate. White et al. [10] found amphotericin B colloidal dispersion to have comparable efficacy to but different toxicity than amphotericin B deoxycholate. Although amphotericin B lipid complex (Abelcet) has been compared with liposomal amphotericin B as empirical therapy in a randomized double-blind trial, the study was designed to compare safety and not efficacy [11]. As a result of these studies, liposomal amphotericin B and itraconazole have been approved by the US Food and Drug Administration for empirical therapy in treating febrile neutropenic patients.

The recent efficacy trials of empirical therapy have important similarities and differences. All were randomized, but the itraconazole and voriconazole trials were not double-blind [9]. All used similar composite end points, as discussed below. However, small differences in the definitions of these end points resulted in profound differences in the success rates across these trials.
PATIENT SELECTION

Selection of high-risk patients. The rationale for empirical antifungal therapy in treating persistently febrile neutropenic patients is to treat clinically occult fungal infections and to prevent subsequent development of mycoses. The key to study design and clinical practice is to identify patients at high risk of a mycosis, to use the best diagnostic tests available for identifying an infectious cause for the fever and, failing diagnosis, to treat with an antifungal agent. If none of the patients in a trial actually have an occult mycosis, then the antifungal drugs being compared will be equivalent because they are equally unnecessary.

Empirical antifungal therapy is best targeted to the patients with prolonged neutropenia (usually >10 days' duration), such as those with acute leukemia and myeloablative allogeneic hematopoietic stem cell transplant recipients. Host factors are an important factor in assessing risk and may change over time. For example, some recipients of autologous hematopoietic stem cell transplants and patients with lymphoma were previously considered at risk for invasive mycoses. However, as the result of cytokine therapy and peripheral blood stem cells, the risk is now diminished. On the basis of the current practices of antineoplastic therapy, most febrile neutropenic patients with solid tumors or lymphomas and recipients of autologous hematopoietic stem cell transplants appear to be at low risk of an occult deep mycosis because the duration of neutropenia is usually short. Neutropenia (<500 neutrophils/mm$^3$) in these patients rarely lasts >8–10 days. Patients who present with fever on day 3 of neutropenia might be enrolled in an empirical antifungal therapy study on day 8 or 9, at which point their neutropenia would soon resolve.

Inclusion of patients who have antibacterial therapy modified prior to enrollment. Modification of the empirical antibacterial regimen on the basis of persistent fever in neutropenic patients is common in hematology/oncology centers. The most frequent modification is the addition of a glycopeptide, such as vancomycin. If the patient is enrolled in a clinical trial of empirical antifungal therapy 24–48 h after alteration of the antibacterial regimen, it is difficult (in absence of microbiological data) to assess the relative merits of either the new antibacterial agent or the antifungal drug. More discussion is necessary on how to analyze the efficacy of antifungal therapy for a patient whose antibacterial regimen is altered after randomization to empirical antifungal therapy.

Inclusion of patients with microbiologically documented infections. Patients with clinically documented bacterial, fungal, viral, or protozoan infections were excluded from the EORTC study, the fluconazole trial [5], and the itraconazole trial [7]. The trials of liposomal amphotericin B and voriconazole included an unspecified number of patients with “controlled” bacteremia; however, no definition of “controlled” bacteremia was provided in the reports. Patients with bacteremia may take >3–4 days to defervesc. If these patients are included in trials of empirical antifungal therapy, resolution of fever may be erroneously ascribed to treatment of an occult fungal infection and therefore may dilute the treatment effect of antifungal therapy.

NEED FOR BLINDING THE STUDIES

The more subjective the end point, the more there is a need for blinding the patient, the investigator, and the sponsor to the study drug. In clinical trials of new antifungal agents, there may be a potential bias against the efficacy of a new drug. Lack of confidence in a new drug may increase discontinuations because of perceived lack of efficacy in a nonblinded study. Similarly, anticipation of amphotericin B toxicity may introduce a bias in unblinded physicians toward early discontinuation of that drug. Documentation in the case report form to validate the reason for early discontinuation has often been scanty and would benefit from an explanation in text form, particularly in unblinded trials.

Problems with blinding may arise in comparing a drug available only as an intravenous formulation with a drug with both intravenous and oral formulations. In this situation, it is logical that both the sponsor and the investigator are interested in exploring the possibility of sequential intravenous to oral therapy. Blinding the study may make such sequential therapy more difficult. For example, in the studies evaluating voriconazole and itraconazole for empirical antifungal therapy, blinding the studies might have precluded an understanding of the benefits of oral therapy following initial intravenous azole treatment. However, blinding would have reduced potential investigator bias in early discontinuations of patients from the study. The studies of itraconazole and voriconazole for empirical therapy clearly demonstrated this problem [7, 8]. There were more early discontinuations due to persistence of fever, perceived as inefficacy, among patients receiving the experimental drug than among those receiving the comparator drug. In these nonblinded trials, the question remains as to the degree to which these discontinuations were influenced by potential investigator bias.

THE CONCEPT OF THE COMPOSITE END POINT

The EORTC empirical antifungal therapy trial used resolution of fever as the main end point of efficacy. Because fever was required for entry into the study, resolution of fever was a logical end point. Given the numerous causes of fever in neutropenic patients, this end point lacks specificity and is no longer used as the sole efficacy end point in empirical antifungal therapy trials. There is general agreement that the most clini-
cally relevant end point in these trials is the absence of breakthrough fungal infections. However, patients who discontinue therapy early because of lack of efficacy, or who die while undergoing therapy and do not undergo autopsy, would have to be considered to have occult breakthrough infections, given the lack of specific diagnostic tests for such infections. With these considerations in mind, recent trials in empirical antifungal therapy have used a 5-component composite end point, representing a significant advance in the design of trials of empirical antifungal therapy in treating persistently febrile neutropenic patients [8, 9]. The components of the composite end point include successful treatment of any baseline fungal infection, absence of breakthrough fungal infection, resolution of fever at some designated time point, survival at some designated time point, and no premature discontinuation of drug because of lack of efficacy or toxicity.

The main concern with a composite end point is that investigators have attached importance to analyses of individual end points, such as the number of breakthrough fungal infections, even though these studies are not usually powered adequately to determine differences in the individual components of the composite end point. Even when the individual end points are planned for analysis before the start of the study, calculation of statistical probabilities for secondary end points remains a dubious exercise. This approach is particularly questionable if the study fails to prove that the drugs are noninferior to each other on the basis of the overall success rates with the primary composite end point. Whether the trial succeeds or fails, secondary end points are best used to generate hypotheses for future studies, not for analysis of a completed study.

Successful treatment of baseline fungal infections. Patients with known, documented, invasive fungal infections are excluded from clinical trials of empirical antifungal therapy. However, for some patients, investigators become aware after randomization of the results of cultures and histopathology done around the time of enrollment. These studies may reveal that these patients did have a baseline fungal infection. Successful therapy for these baseline fungal infections is one measure of antifungal drug efficacy; however, this component of the composite end point usually does not have a major effect on the overall outcome of the trial because the number of baseline infections is usually small. The majority of these patients are excluded before enrollment. Nevertheless, measurement of the outcome for these patients by a blinded review panel provides important insight into the efficacy of the antifungal compound.

It is important to establish protocol-defined, standardized, diagnostic workups to screen for baseline fungal infections before study entry. There may be major differences across centers in the ability to diagnose invasive fungal infections, leading to possible imbalances between centers. An infection that would be detected and either excluded from the trial or categorized as a baseline infection may be detected later in another center and categorized as a breakthrough fungal infection.

Absence of breakthrough fungal infections. A breakthrough mycosis represents unsuccessful treatment of an occult mycosis or emergence of a new mycosis during empirical antifungal therapy. Incidence of breakthrough fungal infections is one of the most objective ways of measuring drug efficacy in empirical trials. However, there are several important considerations in evaluating breakthrough fungal infections.

Establishing common definitions of both baseline and breakthrough infections seems crucial. Because there are currently no accepted distinctions between baseline and breakthrough fungal infections, it may be difficult to separate fungal infections into one of these categories. Separation of baseline from breakthrough infections is an important distinction. The mere presence of a baseline fungal infection (rather than the cure of such an infection) is not a measure of drug efficacy, given that a baseline infection was present before initiation of antifungal therapy. On the other hand, the incidence of breakthrough infections in patients already receiving antifungal therapy may be an important measurement of difference in efficacy between 2 drugs.

A clinical trial of empirical antifungal therapy that makes use of prevention of breakthrough infections as the sole end point may require an impractically large number of patients to achieve adequate statistical power. The true incidence of baseline and breakthrough fungal infections in persistently febrile and neutropenic patients in the absence of empirical antifungal therapy is unknown. In the early EORTC trial, the incidence of fungal infections among patients randomized to the group that did not receive empirical antifungal therapy was 9.4% [3]. More recent empirical antifungal therapy trials in which all patients have received empirical antifungal therapy have shown incidences of baseline and breakthrough infections ranging from 2% to 9% [7–9]. A study with adequate statistical power to show a reduction in breakthrough fungal infections from 9% to 4% with a new drug compared with some standard therapy (a superiority trial design) would require a sample size of >1000 patients. A noninferiority trial designed to show that a new drug was not worse than the comparator by some defined amount would need a comparably large number of patients. The exact number of patients would depend on the noninferiority margin selected for the trial and the success rate of the drugs used in the trial. Determining the acceptable difference in breakthrough fungal infections between 2 regimens (the noninferiority margin or “delta”) in a noninferiority trial is problematic without knowing the incidence in patients receiving no antifungal therapy. If the incidence without empirical antifungal therapy were 9%, the delta would need to be somewhat lower than 9% to allow for some margin of benefit over no treatment.
and to allow for some margin of error in the trial. In the worstcase scenario, the study drug could be judged noninferior to the control agent whereas one or both of the drugs may have no benefit over no empirical antifungal therapy.

Resolution of fever as an end point. Resolution of fever has been a time-honored end point in early empirical antifungal therapy trials. This end point allows some continuity with previous trials such as the EORTC study, upon which the concept of empirical antifungal therapy is based. Patients considered to have occult fungal infections at the time of enrollment on the basis of the presence of fever may still harbor these infections if they remain febrile at the end of therapy. However, fever in neutropenic cancer patients or allogeneic bone marrow transplant recipients may have numerous causes, such as neoplasm, drug reaction, or bacterial or viral infection, other than occult fungal infection. Therefore, fever is a relatively nonspecific marker for both the presence of an occult fungal infection and the clinical response to an empirically administered antifungal drug.

In recent trials of empirical antifungal therapy, the resolution-of-fever component of the composite end point has exerted a large influence on the overall outcome of the trials. The timing of the assessment of resolution of fever has not been consistent across trials, accounting for some of the differences in the success rates for both the resolution-of-fever end points and the overall outcomes for these recent trials. For example, in the study evaluating liposomal amphotericin B [9], resolution of fever was assessed at the time the absolute neutrophil count reached 500/µL and was 58% for liposomal amphotericin B and 58% for amphotericin B deoxycholate. In the study evaluating itraconazole compared with amphotericin B deoxycholate [7], the time point for assessment of resolution of fever was up to 28 days after start of therapy. The success rates for the resolution-of-fever component of the composite end point were 73% for itraconazole and 70% for amphotericin B deoxycholate. In the study evaluating voriconazole compared with liposomal amphotericin B, the timing of assessment of the resolution of fever was 48 h before an increase in the absolute neutrophil count to >500/µL. This more stringent criterion and the short duration over which the study drugs were given (median of 7 days) resulted in lower success rates for resolution of fever than in the previous trials, with 33% for voriconazole and 36% for liposomal amphotericin B.

Compared with a dichotomous value of assessing resolution of fever at a certain time point, one could consider evaluating the time to defervescence by means of a Kaplan-Meier analysis. Including a nondichotomous end point such as time to resolution of fever may be more difficult when used as part of a composite end point, and this requires further discussion. Some experts have suggested that resolution of fever could be considered a secondary end point.

Survival as an end point. A patient who dies while receiving empirical antifungal therapy may have died as a result of an occult fungal infection or a comorbid condition, such as their underlying malignancy or complications of treatment. Blinded data review committees can review an investigator’s assessment of fungal-related deaths, but it is often difficult for investigators to assess the exact cause of death in these complicated patients, even after an autopsy. Therefore, all deaths are classified as treatment failures.

In most studies, >85% of patients have survived beyond the completion of empirical antifungal therapy; therefore, this end point has not been a major cause of failure in the composite end point. Timing of the measurement of survival as a part of the composite end point has not been consistent in empirical antifungal therapy trials. Some trials have assessed survival within 7 days of the end of therapy. The itraconazole study assessed survival between 3 and 28 days after the start of therapy [7]. When assessment of survival is used as an end point, there is a potential for bias from disparities in randomization in the severity of underlying disease, and especially in the rate of comorbidities between treatment arms. Stratification at time of randomization on the basis of severity of underlying disease may help reduce this potential confounder.

Discontinuation for toxicity or lack of efficacy as an end point. Recent trials of empirical antifungal therapy have classified patients as having experienced treatment failure if they discontinued the drug early because of either lack of efficacy or drug toxicity. This end point combines the measurement of drug efficacy, the intended end point, with drug safety and masks important differences between drugs. Although lack of efficacy may be considered an issue of safety as well, it is often more informative to evaluate these 2 aspects of a drug separately. Drugs with lower efficacy but few discontinuations due to toxicity may appear noninferior to drugs that are more effective but are more often discontinued for toxicity. Just such a situation occurred in the trial comparing itraconazole with amphotericin B deoxycholate [7].

Using a composite end point: summary. Used correctly, this 5-component composite end point seems appropriate for the overall assessment of drug efficacy in clinical trials of empirical antifungal therapy in treating high-risk patients. If breakthrough fungal infections and survival were the only end points, the sample size for such trials might be too large to be practical, particularly for lower-risk patients.

ANALYSIS OF ANTIFUNGAL PROPHYLAXIS PRIOR TO EMPIRICAL ANTIFUNGAL THERAPY

Stratification of patients before randomization with respect to prior antifungal prophylaxis is essential and commonly done. However, stratification of patients does not examine the effect
of antifungal prophylaxis on the efficacy of the drugs used as empirical antifungal therapy. In most of these studies, 40%–50% of patients have received prophylaxis before entering the trial. The effects of prophylaxis on outcome of empirical therapy have been difficult to analyze because the prophylactic drug and duration of prior therapy have been heterogeneous. If empirical antifungal studies are to focus on high-risk patients, it is most likely that an even larger percentage of study patients will be receiving antifungal prophylaxis. Future studies should include a detailed secondary analysis of the effect of antifungal prophylaxis on empirical antifungal drug efficacy. It may be useful to require some uniformity of prophylaxis in the period immediately prior to study entry. Important questions would include whether antifungal prophylaxis with a particular triazole affects the efficacy of another triazole or an amphotericin B product when used as empirical antifungal therapy.

**Additional Participants**

Alice Baruch (Pfizer Pharmaceuticals Group, New York), Helen Boucher (Pfizer Global Research and Development, New London, CT), Donald Buell (Fujisawa Healthcare, Northbrook, IL), John Edwards (Harbor/UCLA Medical Center, Torrance, CA), Christopher Hitchcock (Pfizer Global Research and Development, Sandwich, UK), Michael Hodges (Pfizer Global Research and Development, New London, CT), Ann Kolokathis (Pfizer, New York), Helene Panzer (Pfizer Pharmaceuticals Group, New York), Thomas Patterson (University of Texas Health Science Center, San Antonio), John Perfect (Duke University Medical Center, Durham, NC), Patricia Ribaud (Hôpital Saint-Louis, Paris), Jack Sobel (Wayne State University School of Medicine, Detroit), Tania Sorrell (Westmead Hospital, Westmead, Australia).

**References**