Progress on Developing Endpoints for Registrational Clinical Trials of Community-Acquired Bacterial Pneumonia and Acute Bacterial Skin and Skin Structure Infections: Update From the Biomarkers Consortium of the Foundation for the National Institutes of Health

George H. Talbot,1 John H. Powers,2 Thomas R. Fleming,4 Judith A. Siuciak,3 John Bradley,5,6 Helen Boucher,7,8 on behalf of the CABP-ABSSSI Project Team

1Talbot Advisors, Anna Maria, Florida; 2SAIC-Frederick, in support of the National Institute of Allergy and Infectious Diseases, and 3Foundation for the National Institutes of Health, Bethesda, Maryland; 4Department of Biostatistics, School of Public Health, University of Washington, Seattle; 5Department of Pediatrics, School of Medicine, University of California–San Diego, and 6Department of Pediatrics, Rady Children’s Hospital San Diego, California; and 7Tufts Medical Center and 8Tufts University School of Medicine, Boston, Massachusetts

(See the Editorial by Toerner et al, on pages 1122–3.)

Efficacy endpoints for previous registrational trials of antimicrobials for acute bacterial skin and skin structure infections (ABSSSIs) and community-acquired bacterial pneumonia (CABP) were based on nonstandardized, clinician-based observations and decisions, as well as on patient reports. More quantifiable, reproducible, and externally verifiable endpoints could improve the design of future noninferiority trials. At the request of the Food and Drug Administration, the Foundation for the National Institutes of Health convened a broadly representative scientific project team to evaluate potential endpoints for such registrational trials. Review of historical and modern data led to the conclusion that antimicrobial treatment effects are most apparent early in therapy; later outcomes provide important supportive information. Although evidence is incomplete, early response endpoints can anchor noninferiority hypotheses in ABSSSI and CABP registrational trials, thereby allowing evidence-based drug development to continue. Further research is underway to establish which short- and long-term outcomes are well-defined, reliable, and reflective of how patients feel, function, or survive.

Acute bacterial skin and skin structure infections (ABSSSIs) and community-acquired bacterial pneumonia (CABP) are common infections, successful treatment of which requires administration of safe and efficacious antimicrobials. Regulatory requirements for design and conduct of registrational trials for these indications must be scientifically sound and clearly articulated to facilitate development of new antimicrobials.

In past decades, efficacy endpoints for ABSSSI and CABP registrational trials were based on resolution/improvement of signs and symptoms of infection at a time point after completion of therapy. By design, these endpoints included assessments at earlier time points as an element of outcome. As the science advanced, the Food and Drug Administration (FDA) and others realized that the design of noninferiority...
trials evaluating antimicrobials could be improved by defining more reliable outcome measures that reduced dependence on subjective elements and by evaluating outcomes at time points for which prior evidence had demonstrated reliable and reproducible drug effects. The ability to measure known treatment effects on well-defined and reliable outcome measures is essential for noninferiority trial designs. In addition, unambiguous standards for measurement are needed to ensure constancy of the effect of antimicrobials from prior trials to current trials. Finally, the FDA stipulates that outcome measures for studies that support drug registration should be direct measures or established surrogates of how patients feel, function, or survive.

These considerations led to publication of new FDA guidance for ABSSSI and CABP trials [1, 2], both of which focus on assessment of efficacy at an earlier time point than previously recommended. For example, the ABSSSI guidance notes historical data demonstrating treatment effects at 48–72 hours after initiation of antimicrobials. However, the data are incomplete and cannot be audited; the outcomes included body temperature, pulse, respiratory rate, and other measures that are biomarkers (ie, not direct measures of how a patient feels, functions, or survives); and there are significant questions regarding the applicability of these historical data to current circumstances (ie, is the constancy assumption met, which is required for noninferiority designs).

The FDA asked the Biomarkers Consortium of the Foundation for the National Institutes of Health (FNIH; available at: http://www.biomarkersconsortium.org/) to convene a project team (PT) to evaluate past evidence on endpoints for these 2 indications and to develop and evaluate proposed endpoints for future ABSSSI and CABP registrational trials. Team membership included broad participation from the National Institutes of Health (NIH), the FDA, the academic research community, and biopharmaceutical companies. Considering the FDA's standards for drug approval, the PT evaluated historical evidence for treatment effects; evaluated outcomes in recent clinical trials; outlined research gaps; proposed interim, bridging endpoints; and began developing future research on outcome assessment in ABSSSI and CABP trials. The PT conclusions were submitted to the relevant FDA dockets [3, 4]. This article summarizes the process, major findings, and planned future research of this intensive, multidisciplinary effort.

**REGULATORY STANDARDS FOR APPROVAL OF NEW DRUGS**

The US regulatory standard for drug approval is "substantial evidence" from "adequate and well-controlled trials" [5]. One requirement is that methods of assessment of subjects’ response must be “well-defined and reliable.” A proposed study should explain (1) what to measure, (2) how the outcome was measured, and (3) how the data were analyzed to define a meaningful outcome; that is, what defines a “responder” [6].

US regulations, as well as a recent Institute of Medicine recommendation [7], define a “clinically meaningful” outcome measure as a direct measure of how patients feel, function, or survive. However, appropriately developed and validated surrogate endpoints can be used as outcome measures to support approval if the “surrogate endpoint … is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit” [8] and if follow-up studies show that benefit on the surrogate reflects benefit on direct patient-centered outcomes. “Surrogate endpoints” are biomarkers (eg, signs of disease, results of radiological tests, cultures, and laboratory values) that act as substitute indirect measures for how patients feel, function, or survive [9]. In trials of antimicrobials, acute infection symptoms (eg, pain) are direct measures of patient benefit, and signs of disease (eg, erythema) and selected markers of systemic response (eg, leukocytosis) are biomarkers having mechanistic links to infection.

An important consideration for the design of registrational trials for these indications is that they are conducted globally and submitted for marketing authorization globally. Accordingly, harmonization of clinical trial design with requirements of non-US regulatory authorities, such as the European Medicines Agency, is crucial [10].

**HISTORICAL EVIDENCE**

Two trials published in 1937 compared sulfa drugs to ultraviolet light in the treatment of erysipelas [11, 12]. The studies used alternate assignment and were not blinded. The investigators evaluated outcome measures that included cessation of erysipelas skin lesion spread, durations of pyrexia and toxemia, death, relapse, and complications. The greatest treatment effect on which to base noninferiority hypotheses was cessation of lesion spread at 48 hours (Table 1). In a meta-analysis of the 2 studies, the difference was 24% in favor of sulfa drugs (95% confidence interval [CI], 18.2%–30.0%) (communication from FDA to FNIH, 5 May 2010). If deaths and failures during initial therapy are included in the analysis, the difference was 26% in favor of sulfa drugs (95% CI, 19.0%–33.5%). There was no treatment effect on death, relapse, or complications on which to base future noninferiority hypotheses. The PT agreed that data on control of lesion spread were a useful—albeit imperfect—starting point for further investigation. The PT performed a review of the medical literature evaluating current techniques for measuring skin lesions. No study specifically addressing the measurement of acute cellulitis lesion size was identified; therefore, the characteristics, reliability, and reproducibility of various measurement techniques in the
measurement of ABSSSI lesions are unknown. Optimal measurement methods for acute bacterial skin lesions still need to be determined.

In CABS, reviews of work from the preantibiotic era [13, 14] provided illustrations of the typical course of symptoms associated with acute bacterial pneumonia in adults, including cough, dyspnea, chest pain (especially pleuritic pain), and sputum expectoration. In general, steady deterioration of untreated patients’ respiratory symptoms continued through days 3–4. These results contrast with those in the early antibiotic era, when the treatment effect was both rapid and large [15–20]: a quantitative estimate of treatment effect for symptom resolution at 48–72 hours was 29% (95% CI, 21%–37%) [18], a quantitative estimate of treatment effect for clinical recovery at day 3 was 72%–77% [14–16], and quantitative estimates of treatment effect for mean days to clinical improvement, fall in temperature, and clinical recovery were 2.5, 3.4, and 4.2 days, respectively [17].

A review of historical data on antimicrobial treatment effect in CABS was presented at a 2009 FDA advisory committee meeting [21]. Figure 1 is illustrative of the FDA’s findings.

These data do have limitations: they mostly derive from observational or small studies, cross-study comparisons were used to determine treatment effect, and the endpoints were not clearly defined, although they were clinically reasonable. However, collectively, these preantibiotic and early antibiotic

![Figure 1. Duration of symptoms in patients with pneumococcal pneumonia [21].](https://academic.oup.com/cid/article-abstract/55/8/1114/339262)
era data indicated a significant treatment effect at days 3–4 after initiation of therapy.

**RETROSPECTIVE DATA ANALYSES**

The PT performed retrospective analyses of data sets from modern clinical studies to (1) refine the FDA-proposed outcome measures by evaluating their operational characteristics, changes over time, and responsiveness to change at specific time point and (2) identify additional relevant endpoints or biomarkers. For ABSSSI, data from registrational trials of ceftaroline fosamil, dalbavancin, and linezolid were used [22–24]. For CABP, data from trials of tigecycline, daptomycin, and ceftaroline fosamil were used [25–27]. Analyses were performed per statistical analysis plans drafted prospectively by qualified PT biostatisticians.

The PT analyzed these data sets in a post hoc subgroup of all randomized patients defined by criteria similar to those used for inclusion in current clinical trials. These analyses aimed to provide descriptive data of potential early response criteria for ABSSSI and CABP. Analyses provided descriptive statistics and graphical depictions of summary data. No inferential statistical analyses (ie, hypothesis tests) were completed. The PT noted that these analyses are relevant to the specific context of use evaluated in these studies and have the potential to confirm observations from the earliest antimicrobial studies. A detailed description of the results, which is beyond the scope of this article, can be found in the FNHI submissions to the ABSSSI and CABP FDA dockets [3, 4] and also in the Supplementary Materials.

**INTERIM RECOMMENDATIONS**

On the basis of these analyses, the PT proposed outcome measures for use by sponsors for registrational trials involving ABSSSI and CABP. Most active-controlled trials of antimicrobials for these indications use a noninferiority design because of ethical and feasibility concerns. The PT’s discussion therefore focused on the outcome measures used in ABSSSI noninferiority trials in adults. These outcome measures are proposed as “bridging” measures for registrational clinical trials planned for submission to the FDA.

In ABSSSI, response should be determined 48–72 hours after randomization, with baseline lesion size measurement and administration of the first dose of study drugs occurring as close to randomization as possible. For adults, success should be defined as control of infection lesion spread (ie, a ≥20% decrease in lesion area [calculated as the longest head-to-toe length times the longest perpendicular width] vs baseline). Absence of elevated body temperature (fever) should not be a component of the primary outcome measure, particularly because it is not on the causal pathway of the disease; furthermore, a requirement for fever as an inclusion criterion would exclude important patient populations (eg, elderly individuals), and frequent temperature measurements cannot be obtained reliably in many clinical trial settings, as confirmed in a subsequent phase 3 study [28]. However, resolution of elevated body temperature is important in clinical practice. Thus, resolution of fever and achievement of stability in other physiological parameters should be included as sensitivity analyses and/or as part of late assessment endpoints.

The proposed early endpoint serves as the link to the historical evidence of drug effect in previous studies, as required in the design of noninferiority trials [29]. The classification of response or failure is also independent of study events other than death; institution of alternative therapy, unplanned surgical drainage, and adverse events are not counted as part of the early response classification, but these measures can be captured as part of supportive information.

Patients may continue to receive the study drug even if they are classified as an early nonresponder, or they may have therapy withdrawn even if they are classified as an early responder. This element of trial conduct is especially important to enable analysis of a traditional later endpoint for both FDA and non-US regulatory authorities. Longer-term clinical outcome measures at a prespecified time point, such as the end of therapy and thereafter, are of relevance to clinicians and patients. These can be measured as secondary outcome measures. Evaluation of sustained control of lesion spread is critical, as are other clearly defined assessments of how patients feel or function from treatment initiation to infection resolution. However, defining components of later time points requires further research. In the meantime, for later endpoints, sponsors should propose to the FDA what they plan to measure and how and when they will measure it. Performing early and late assessments will allow separate analyses that meet differing requirements of FDA and non-US regulatory agencies, thereby allowing global drug development [10].

In CABP, assessment of symptom improvement at treatment day 4 (approximately 72 hours after randomization and initiation of therapy) provides relevant data on how patients feel and function and can provide evidence of a strong treatment effect for antibiotics via its link to assessments of symptom improvement in historical studies. Four biologically relevant symptoms (cough, pleuritic chest pain, dyspnea, and sputum production, scored as points for absent, mild, moderate, and severe) are recommended for adults. The proposed endpoint measure comprises (1) a 1-point improvement in at least 2 symptoms, (2) no worsening of any other symptoms, and (3) assessment on study day 4. Evaluations of the comprehensiveness of the symptoms in the definition should be a focus for future research. The timing of resolution of early...
symptom endpoints has not been well documented for infants and children.

As for ABSSSI, absence of elevated body temperature is not recommended as part of the early CABP endpoint, in part since it may be confounded by antipyretic therapy.

**ALTERNATIVE VIEWPOINTS**

Alternative viewpoints were voiced regarding the conclusion that the interim, bridging primary endpoints for registrational trials of ABSSSI and CABP should be based on the proposed early outcome measures. Although team members agreed that these early measurements provide important information, concerns included the lack of rigorous justification of a biomarker (control of lesion spread) as a registrational endpoint, the insufficiency of control of lesion spread as a comprehensive measure of overall improvement and ability to return to full functional status, the implicit inclusion of early response measures in test-of-cure endpoints, and the inability to determine whether the use of test-of-cure endpoints has led to approval of inactive therapies for life-threatening infections such as ABSSSI. Members agreed that well-defined, reliable, later endpoints would provide an important overall perspective. Furthermore, minimal lesion size requirements at baseline, as required by the FDA, exclude certain target populations (eg, infants and children). Finally, recent nonrandomized pharmacometric analyses show a correlation between drug exposure and traditional clinical and microbiological endpoints, although some observed that such evidence does not establish causality. The PT agreed that these issues should be addressed in future research.

**ADDITIONAL INFORMATION NEEDED TO ADVANCE TO FINAL RECOMMENDATIONS**

Analysis of current data reveals that the evidence base for determining outcomes in ABSSSI and CABP remains incomplete, particularly for certain populations (eg, pediatric patients). Future studies must address research gaps such as determining which outcomes are most important to patients and how to reliably and reproducibly assess them across various disease states, populations, and anatomic sites. The FNIH has initiated studies to address these questions.

**DISCUSSION AND CONCLUSIONS**

Clinical studies serve many purposes. For registrational trials, FDA regulations state that “the purpose of conducting clinical investigations of a drug is to distinguish the effect of a drug from other influences, such as spontaneous change in the course of the disease, placebo effect, or biased observation” [6]. For valid inferences from noninferiority trials, historical evidence of drug effects under the conditions of prior studies must apply to the active comparator’s effect in the noninferiority trial [30], leading to need for consistency in design features, including definitions and timing of outcome measures.

In the process outlined above, members of academia, industry, and governmental agencies proposed interim, bridging outcome measures for future trials involving the ABSSSI and CABP indications. The PT evaluated historical evidence for drug effects—that is, the difference in outcomes between the test and control groups for antimicrobial and nonantimicrobial interventions for ABSSSI and CABP—and found that the largest treatment effects were documented early in the course of disease. This observation aligns with clinician assessment in that therapy is often changed in the absence of response after a few days of treatment. The evidence shows that, as the duration of follow-up increases for patients in the test and control groups, the proportion of patients who appear to achieve a response increases. However, the difference in outcomes between the antimicrobial and nonantimicrobial interventions often decreases over this period, which suggests that patients’ responses later in treatment may be due in part to the disease’s natural history.

Previous studies have not evaluated response at early time points in a standardized way. Carrying early nonresponses forward and combining them with later nonresponses may lower success rates, making it more challenging to demonstrate noninferiority while also not accurately measuring drug effect.

During the interim bridging period, early time points should be used in primary hypotheses of registrational noninferiority trials, because the choice of outcome measures is constrained by that design’s reliance on estimates of the active control’s effects from prior studies. However, the PT noted that measurement of effects early on the basis of noninferiority hypotheses does not preclude superiority hypotheses for later effects. This approach conforms to clinical concerns that some drugs may differ on later outcomes. Superiority studies can evaluate other endpoints and timing of outcomes as long as they are clinically meaningful to patients, well-defined, and reliable. Of importance, failure to show superiority at later time point should not preclude approval.

The importance of a late time point is greatest in diseases in which a lack of sustained cure is part of the natural history of the disease. Evaluation of the historical evidence and current trials showed low rates of late failure in both ABSSSI and CABP. This situation is in contrast to conditions such as *Candida difficile* infection, in which up to a quarter of patients may not maintain an initial response [31].

The measurement of early effects addresses clinical practice as well as clinical trials. The historical evidence shows that a
few doses of an antimicrobial can have a profound effect on patient outcome. The historical evidence, as well as results of more recent clinical trials, also shows that shorter therapy du-

rations may be appropriate for both ABSSSI and CABP.

The PT recognized the limitations of the proposed interim endpoints. Therefore, the FNIH is proceeding with prospective research to evaluate the content validity and measurement properties of endpoints in both ABSSSI and CABP for adults. The studies on ABSSSI are anticipated to begin in the summer of 2012, with work on CABP to begin in the subsequent winter pneumonia season.

On the basis of these deliberations and input from the FDA, new antimicrobials for ABSSSI have advanced into phase 2 and 3 development [18, 32–36]. In addition, retro-

spective analyses of other ABSSSI and CABP clinical trial data sets have produced results concordant with those of the PT analyses [37–40].

In summary, these interim, evidence-based outcome me-

asures are proposed on the basis of an analysis of the historical literature, which showed an early treatment effect of anti-

microbials for ABSSSI and CABP. A core focus of registra-

tional studies is to determine the presence or absence of a treatment effect. These interim outcome measures allow noninferiority registra-

tional studies to proceed in the United States while future qualitative and quantitative research studies are being con-

ducted. These future studies are critical in addressing knowledge gaps and would help address unanswered questions related to designing noninferiority trials involving these 2 im-

portant indications.

Supplementary Data

Supplementary materials are available at Clinical Infectious Diseases online (http://cid.oxfordjournals.org). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copystyled. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. We thank the following members of the CABCABSSSI Project Team for their contributions: Joseph Toerner, MD, MPH (cochair, FDA; nonvoting); George H. Talbot, MD (cochair, Infectious Diseases Society of America [IDSA]); Jeff Alder, PhD (Bayer); Paul G. Ambrose, PharmD (Institute for Clinical Pharmacodynamics, University of Oxford); Helen Boucher, MD (Tufts University); John Bradley, MD (University of California, San Diego); Laurie Burke, RPh, MPH (FDA; nonvoting); Edward M. Cox (FDA; nonvoting); Aaron Dane, MSc (Astra-Zeneca); Anita Das, PhD (AssiStat); Dennis Dixon, PhD (NIH/NIADD); Michael Dunne, MD (Durata Therapeutics); Barry Eisenstein, MD (Cubit Pharmaceuticals); Thomas R. Fleming, PhD (University of Washington); Dean Follmann, PhD (NIH/NIADD); David Friedland, MD (Cerexa); Ian Friedland, MD (Cubist Pharmaceuticals); Nicholas Karlsonis, MD (Merck); Achim Kaufhold, MD, PhD (Basilea Pharmaceutica International); Scott Komo DrPH (FDA; nonvoting); Mike Kurilla, MD, PhD (NIH/NIADD); Kim Lindfield, PhD (Cubist Pharmaceuticals); Lily Llorens, PhD (Cerexa); Susan Moriarty, MD (Cempra Pharmaceuticals); Sumati Nambiar, MD (FDA observer; nonvoting); Gary J. Noel, MD (Parexel Pharmaceuticals); David Oldach (Cempra Pharmaceuticals); Elekta Papa-
dopoulous, MD (FDA; nonvoting); Roger Pomerantz, MD (Merck); John Powers, MD (SAIC, in support of NIH/NIADD); Philippe Prokocimer, MD (Trius Therapeutics); John Quinn, MD (Pfizer); John Rex, MD (AstraZe-

neca); Denise Russo (NIH/NINR); Jennifer Schranz, MD (Cempra Pharmaceuticals); Judith A. Siuciak, PhD (FNIH); William Stubbs, PhD (Basilea Pharmaceutica International); and David Wholley (FNIH).

Disclaimer. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Ser-

vices, nor does mention of trade names, commercial products, or organiza-

tions imply endorsement by the US Government. The conclusions described within this document represent the work of the FNIH Biomark-

ers Consortium Project “Developing Endpoints for Clinical Trials of Drugs for Treatment of Acute Bacterial Skin and Skin Structure Infections and Community-Acquired Bacterial Pneumonia (CABP ABSSSI Project).”

Financial support. This work was supported by the National Cancer Institute (contract HHSN261200800001E) and the National Institute of Allergy and Infectious Diseases, National Institutes of Health; Actelion Pharmaceuticals; Basilea Pharmaceutica International; Cempra Pharmaceuticals; Cerexa, a wholly owned subsidiary of Forest Laboratories; Cubist Pharmaceuticals; Merck; Nabriva Therapeutics; and Trius Therapeutics. Clinical trial data were also contributed to the project by Cerexa, Cubist Pharmaceuticals, Durata Therapeutics, and Pfizer.

Potential conflicts of interest. Within the past 36 months, G. H. T., through Talbot Advisors, has received board compensation and/or consult-

ancy fees from Actelion, Basilea, Bayer, Cempra, Cerexa, Durata, Cubist-

Calixa, FAB Pharma, J&J, Kalidex, Meiji, Merada, Merck, Nabriva, and Wyeth/Pfizer (data safety monitoring board). G.H.T. also has equity inter-

ests in Durata, Calixa, Cerexa, Kalidex, and Nabriva; he did not receive compensation for work on the PT, including manuscript preparation, but travel expenses to PT meetings were reimbursed by the Infectious Diseases Society of America and the FNIH. In past 36 months, J. P. served as con-

sultant for Advanced Life Sciences, Astra-Zeneca, Cardeas, Contrafect, Gilead, Johnson and Johnson, LEO, Methylenyl, Novartis, Optimer, Pfizer, Trius, Warner Chalcott, and Wyeth. J.S. is an employee of the FNIH. In the last 3 years, H. B. provided advice or consultation to Basilea, Cerexa, Durata, Merck (adjudication committee), J&J, Rib-X, Tar-

ganta/The Medicines Company, and Wyeth/Pfizer (data safety monitoring board). All other authors: No reported con-

flicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References


2. US Food and Drug Administration, Center for Drug Evaluation. Guidance for industry: community-acquired bacterial pneumonia: de-


3. Foundation for the National Institutes of Health Biomarkers Consor-

tium Project Team. Recommendations to FDA for interim endpoints for clinical trials in acute bacterial skin and skin structure infections. Available at: http://www.regulations.gov/#/docketDetailrpp=10;po=0;D=FDA-2010-D-0433T. Accessed 4 March 2012.

4. Foundation for the National Institutes of Health Biomarkers Consor-

tium Project Team. Recommendations to FDA for interim endpoints for clinical trials in community-acquired bacterial pneumonia.
1120 • CID 2012:55 (15 October) • REVIEWS OF ANTI-INFECTIVE AGENTS


39. Hait H, Arbeit R, Molnar D, Noel GJ, Tanaka SK. In a phase 2 complicated skin and soft tissue infections trial, outcomes assessed early in the course of therapy were consistent with outcomes 1017 days after completing therapy with either omadacycline (OMC; PTK796) or linezolid [abstract P1528]. In: Abstracts of the 21st ECCMID/27th ICC, Milan, Italy, 7–10 May 2011.