Substantial Evidence: When Is a Single Trial Sufficient for Approval and Promotion?

Article *in* Therapeutic Innovation and Regulatory Science · May 2011

DOI: 10.1177/009286151104500306

3 authors, including:

David E Coutant
Eli Lilly and Company
26 PUBLICATIONS **176** CITATIONS

SEE PROFILE

Substantial Evidence: When Is a Single Trial Sufficient for Approval and Promotion?

David Coutant, PhD, RAC
Principal Research Scientist, Drug Disposition, Eli Lilly and Company, Indianapolis, Indiana

David Riggs, PharmD
Director, US Regulatory Affairs,Eli Lilly and Company, Indianapolis, Indiana

Elizabeth Van Sant Hoffman, PharmD
Associate Manager, US Regulatory Affairs, Eli Lilly and Company, Indianapolis, Indiana

The Federal Food, Drug, and Cosmetic Act provides FDA and sponsors with a statutory definition of "substantial evidence." Most often drug effectiveness is established through the conduct of two adequate and well-controlled clinical trials (AWCTs), but there are situations where effectiveness can be sufficiently established through the conduct of a single AWCT. Some of these situations were detailed in FDA's 1998 guidance document, but there remains some uncertainty for sponsors in determining what cases might merit a single AWCT to meet the substantial evidence standard (depending on indication, drug class, etc). This article is meant as an aid for drug development teams, and regulatory professionals in particular, to navigate successfully through such questions of substantial evidence. We review some of the prerequisites and hurdles for drug approval based on FDA interpretation of substantial evidence standards. Finally, we will provide some suggestions to sponsors, in the context of case examples demonstrating FDA's interpretation of substantial evidence to support promotional claims.

Key Words
Substantial evidence; Clinical effectiveness; Promotion

Correspondence Address
David Coutant, PhD, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285 (email: cout@lilly.com).

INTRODUCTION

The Federal Food, Drug, and Cosmetic Act (FD&C Act) provides the Food and Drug Administration (FDA) and sponsors with a statutory definition of "substantial evidence," as follows:

> The term substantial evidence means evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof. If the Secretary determines, based on relevant science, that data from one adequate and well-controlled clinical investigation and confirmatory evidence (obtained prior to or after such investigation) are sufficient to establish effectiveness, the Secretary may consider such data and evidence to constitute substantial evidence for purposes of the preceding sentence. (1)

In 1998, the FDA Guidance "Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products" (2) provided many useful recommendations for establishing substantial evidence of clinical effectiveness of drugs and biologics in the United States. In particular, this guidance on evidence provided FDA guidelines to sponsors to facilitate understanding when clinical effectiveness could be adequately assessed without fulfilling the approval requirement of multiple phase 3 clinical studies. The evidence guidance (2) further clarified the FDA interpretation of the 1997 FDA Modernization Act that modified Section 505(d) of the FD&C Act to make it clear that the FDA may consider "data from one adequate and well-controlled clinical investigation and confirmatory evidence" (1) to constitute sufficient substantial evidence of clinical effectiveness. As stakeholders want to bring innovative therapies to patients as efficiently as possible, the evidence guidance was welcomed by many. However, even with careful reading, several questions remained as to when less than two adequate and well-controlled clinical trials (AWCTs) might be part of a successful registration strategy or be adequate support for promotional claims.

This article does not revisit the evidence guidance (2) page by page, nor discuss the FDA's additional thoughts related to substantial

Drug Information Journal, *Vol. 45, pp. 253–263, 2011 • 0092-8615/2011*

Submitted for Publication: November 4, 2009
Accepted for Publication: October 14, 2010

evidence put forth in more recent guidances (3,4). This article does outline examples of when a single AWCT was sufficient for effectiveness and when a single AWCT was not sufficient for product promotion, and highlights suggestions for sponsors based on these examples and FDA guidances.

EFFICACY

BACKGROUND

For the approval of a New Drug Application (NDA), Biologics License Application (BLA), or supplemental application, it is critical that sufficient evidence of effectiveness is available so that both the sponsor and the FDA can adequately complete the benefit/risk (B/R) assessment of the new molecular entity (NME). Section 505(d) of the FD&C Act, as well as Section 351 of the Public Health Service Act, indicate that new drugs and biologics should establish substantial evidence of clinical effectiveness through means of "adequate and well-controlled studies." The base assumption is that since the term *studies* is plural, two or more AWCTs are required to establish efficacy. Although this is the base case, there exist many scenarios (and many past examples) where the FDA has interpreted the "substantial evidence" standard differently. For further US regulatory guidance on this topic, in particular, the reader is referred to three guidances that provide details on establishing clinical evidence of effectiveness (2), discuss how proof of effectiveness should preferably be presented in NDAs or BLAs (3), and discuss how cancer and lifesaving medicines may merit approval on the basis of less than two adequate and well-controlled trials (4).

There are many good reasons for basing the evaluation of many NMEs on more than a single AWCT. The 1998 FDA evidence guidance (2) cites four reasons of particular concern when drug development is based on a single AWCT: bias, chance, site-specific results, and fraud. For instance, as a consequence of the inherent variability of clinical drug development, the FDA points out (2) that 1 in 40 clinical trials (CTs) will be erroneously positive if all these CTs were found to be statistically significant using a two-tailed P value of 0.05. The use of multiple AWCTs has proven to help mitigate against approval of unsafe or ineffective drugs based on chance positive results. However, there are many situations where clinical effectiveness and safety can be adequately characterized utilizing less than two AWCTs. Table 1 summarizes the many cases in which substantial evidence of less than two AWCTs may be possible (2,5–10). It is acknowledged that due to the body of evidence previously established, supplemental applications (sNDAs, sBLAs) for approved drugs and biologics have an increased likelihood of acceptance based upon less than two AWCTs compared to NMEs. Approvals today for new drugs based on a single phase 3 AWCT are more likely to occur when the drug is for an important unmet medical need, such as for certain oncology or high-risk cardiovascular indications. For instance, temsirolimus (approved May 2007), everolimus (approved May 2009), and pazopanib (approved October 2009) are all kinase inhibitors approved for treatment of advanced renal cell carcinoma based on a single phase 3 AWCT.

NEEDS

By necessity, the substantial evidence requirement must be uniquely pursued for each NME. In other words, the nature of the set of studies that comprise substantial evidence varies among applications. Even if the sponsor follows all the road signs (eg, guidances, past history), consults a navigation aid (eg, fruitful FDA interactions), and starts with a clear destination in mind (eg, the targeted desired draft launch label or the target product profile), there remains the possibility that the sponsor will not reach the destination (eg, timely approval achieved on less than two AWCTs). We submit the following considerations to provide the sponsor with previous examples of implementing a successful substantial evidence strategy based on a single AWCT, or a single AWCT of a new use with independent substantiation from related clinical study data.

TABLE 1

Establishing Substantial Evidence of Clinical Efficacy Using Less Than Two New Adequate and Well-Controlled Clinical Trials

Scenarios	Comments	Prior Example
Case 1: Substantial evidence proven via extrapolation from existing studies or from new comparative PK bioavailability trial. No new AWCT required.		
Pediatric use labeling, or a different salt, formulation, dose regimen, modified release, or dosage strength of an NME may not need a new AWCT to establish safety and efficacy. A bioequivalence study may provide sufficient PK data to allow bridging.	These types of changes for antihypertensive or antianginal therapies generally still need at least one AWCT. A modified dose form may or may not require a clinical trial.	Pediatric use labeling for ibuprofen extrapolated from adult efficacy data (initial approval May 1984) without need for a pediatric clinical trial.
Case 2: A single AWCT with independent substantiation from related clinical data.		
Studies of different doses, regimens, or dosage forms may be approved in this manner.	If the PK/PD relationship is not well characterized, a single AWCT may be needed to bridge doses or dosage forms.	Because the PK/PD relationship of risperidone (initial approval December 1993) was not well understood at the time, a single AWCT was needed between the q.d. and b.i.d. dose regimens (second dose regimen approved October 1997).
Studies in other populations may be approved in this manner.	Consult with FDA to robustly define the NME's target population, especially for novel indications or populations.	Initially approved for use in females in December 1977, tamoxifen was approved years later for the treatment of breast cancer in males based on a single AWCT.
Combination use and monotherapy use; one AWCT for each may support the other in this manner.	This approach does not always meet with success. For instance, in the late 1990s, the FDA concluded that a single favorable CT was insufficient evidence to support a monotherapy indication for Neurotonin (gabapentin), although it had already shown effectiveness as an antiepileptic drug used as part of combination therapy.	Victoza (liraglutide) was approved in January 2010 as a treatment for type 2 diabetes mellitus, without being restricted to combination use. Monotherapy label language was approved on the basis of a single (746 patient) monotherapy AWCT and four supporting combination-use AWCTs.
Studies in other phases of the same disease.	In many cases, drugs that are effective in one phase of a disease will be effective in another phase of the disease, albeit the magnitude of response and benefit/risk ratio may differ.	The approval of timolol for the reduction of postinfarction mortality was based a single AWCT with a low *P* value that showed a major effect on reinfarction rate and mortality. Patients in the trial were randomized into three strata of disease severity—each stratum demonstrated efficacy.
Studies in two closely related diseases or in two pathologically related conditions. One AWCT for each disease may support the other in this manner.	It appears the same principle can be applied to an NME studied against two separate comparators. That is, one AWCT vs active comparator 1 and one AWCT vs active comparator 2, with each AWCT supporting the claim of the other, has been a successful strategy for certain antidiabetics (eg, liraglutide).	Eptifibatide was approved as add-on to aspirin for treatment of acute coronary syndrome (ACS), and for the treatment of patients undergoing percutaneous intervention (PCI). The PURSUIT trial evaluated eptifibatide as add-on to aspirin in ACS patients while the IMPACT II trial evaluated PCI patients. Together the two trials provided adequate evidence of effectiveness for both indications (approval May 1998).
Less-closely related diseases, similar purpose of therapy. (Example: effectiveness in one tumor might suggest reliance on a single study in a second tumor, depending on tumor type.)	Establishing substantial evidence in this situation may be difficult.	Taxotere was initially approved (May 1996) for treatment of patients with breast cancer after failure of prior chemotherapy. It now has several approved indications. Taxotere was approved in December 1999 for the treatment of

TABLE 1

Continued

Scenarios	Comments	Prior Example
		non-small-cell lung cancer and for treatment of hormone-refractory prostate cancer in May 2004, each based on a single AWCT.
Studies with two different, but related, clinical endpoints.	As a general note, strength of data will determine if an outcome claim can be made on the label based on results of a surrogate endpoint (eg, a lipid-lowering outcome).	Enalapril (initial approval 1985) was approved for the treatment of heart failure on the basis of one trial that showed improved survival, and another complementary trial that showed symptom improvement over several months.
Support by pharmacologic or pathophysiologic endpoints. Note that when "the pathophysiology of a disease and the mechanism of action (MoA) of a therapy are very well understood, it may be possible to link specific pharmacologic effects to a strong likelihood of clinical effectiveness." (2)	This is a particularly difficult scenario, as the sponsor will need to establish on a case-by-case basis if the MoA is sufficiently "very well understood." The sponsor must weigh how much epidemiologic proof or prior history is needed to support an outcome claim based on a single AWCT.	Vaccines are a class of drugs in which one AWCT plus supporting animal challenge data (showing protection against specific pharmacological effects) can be considered sufficient weight of evidence for approval according to the 1998 FDA evidence guidance (2).
One AWCT supported by extensive prior clinical safety data and efficacy from a closely related drug. One example would be reference to a closely related drug with extensive clinical safety data. Another would be registration of a metabolite of a previously approved prodrug.	In general, it is expected that rarely will two drugs be sufficiently closely related to allow for the approval of the second drug for a specific indication based on a single AWCT.	The ESSENCE trial compared enoxaparin, a low molecular weight (LMW) heparin, with unfractionated heparin in patients. Despite missing the targeted reduction in the composite endpoint, the advisory committee recommended approval based on the extensive prior clinical safety record of LMW heparin, superiority to placebo, and published results from another LMW heparin (dalteparin). Initial approval was March 1993.
Case 3: Reliance on a single multicenter study, without supporting information. Single AWCT.		
Large multicenter trial.	All investigators must follow protocols in the same manner: for recruitment, enrollment, randomization, and all aspects of data collection.	The Beta-blocker Heart Attack Trial (BHAT) showed a significant reduction in all-cause mortality in propranolol group compared to the placebo group (7.2% vs 9.8%) during the average 24-month follow-up period (BHAT stopped early in October 1981 due to demonstrated benefit).
Consistency across subsets or groups.	Sponsor must not selectively present only the most favorable data. Proof of effectiveness needs to be established on prespecified endpoints, not post hoc analysis.	Firmagon (degarelix) was approved in December 2008 for the treatment of advanced prostate cancer on the basis of a single phase 3 study. Efficacy (testosterone suppression and medical castration) was observed in both degarelix dose arms, and in patients with varying degrees of disease severity.
Multiple studies within a single study.	Sponsors needs to make appropriate adjustments for multiple comparisons.	The Second International Study of Infarct Survival (ISIS-2) examined the effectiveness of i.v. streptokinase, oral aspirin, or both among 17,187 cases of suspected acute myocardial infarction (ISIS-2 completed in 1988). This is an example of how a proper factorial design and a series of pairwise comparisons can show, in a single AWCT, efficacy as a monotherapy and efficacy in combination with another drug.

TABLE 1

Continued

Scenarios	Comments	Prior Example
Multiple endpoints involving different events.	Betaseron is a successful example of this category. However, in other cases failing significance for a primary endpoint, further analyses (subpopulations or secondary endpoint analysis) will be considered exploratory only, and generally will be considered insufficient proof of substantial evidence.	Betaseron (Interferon beta Ib) was approved in July 1993 for the prevention of exacerbations of relapsing forms of multiple sclerosis because the trial showed significant improvement in two different, but logically related primary endpoints, namely decrease in MRI-demonstrated disease activity and decreased rate of exacerbation.
Statistically very persuasive. Extreme *P* value.	The requirement for low *P* values is sometimes relaxed for drugs for important treatments.	The initial approval (November 1997) for clopidogrel was based on a single large AWCT with a *P* value of only 0.04 on its primary endpoint. Clinical results that suggested equivalence of clopidogrel and aspirin, superiority to historical controls, and an absence of major clinical toxicities likely aided in ultimate regulatory approval based on a single AWCT.

The 21 CFR 314.126 rule (11) describes the five types of control groups in AWCTs:

1. Placebo concurrent controls
2. Exposure response concurrent controls
3. No treatment controls
4. Active concurrent controls
5. External (historical controls)

Most commonly the AWCT will be a placebo-controlled or active-controlled trial, looking at one or more dose levels of the investigational drug (12). The AWCT will be either designed to show superiority to active control or placebo, or noninferiority to active control. For superiority trials, the choice of active control versus placebo control should be considered on a trial-by-trial basis (13). When a noninferiority design is merited, the important study design features (including active control dose level) should be the same as in the previously conducted trials in which the active control demonstrated clinically relevant efficacy (13,14). In both superiority and noninferiority trials with active controls, it is important to choose an appropriate dose and dose regimen of the control and the test drugs and to this end it may be necessary to study several doses of the control and perhaps several doses of the test treatment (15).

Ideally, there should be baseline comparability of patients in AWCTs (16). In some cases, a lead-in period may help compare patients at baseline. The appropriate patient population must be selected. All primary endpoints should be reliable and prespecified, as prespecification of study objectives helps prevent the well-intentioned sponsor from data fishing. One memorable case of data fishing occurred in the early 1980s, when a retrospective meta-analysis linked coffee drinking to an increased risk of pancreatic cancer (17). Thankfully for those of us who are coffee drinkers, this conclusion was later refuted by additional epidemiological and statistical data, but it serves as an example of a type of error that can be caused by post hoc analysis. The method of statistical data analysis should be prespecified. The targeted level of statistical significance should be clearly stated a priori and achieved for the intended efficacy claim (typically $P \leq 0.05$ for two studies; $P \leq 0.01$ to 0.001 for one study). Interim looks should be described (12). Adjustments need to be made for multiple comparisons and for co-primary endpoints. Sensitivity analysis is recommended, with particular focus on the magnitude of the treatment effect (rather than focusing on the presence or absence of statistical

significance). Finally, patients should be accounted for at the end of the analysis. It was the Anturane Reinfarction Trial (ART) of 1980 that first showed the importance of not dropping plausible outliers (18). Although Anturane seemed to help prevent sudden death after myocardial infarction, and ART was a study that was double-blinded, randomized, and placebo-controlled with 1,600 participants, serious analysis bias was later discovered. In brief, nine participants who had died (eight on Anturane and one on placebo) were excluded from the sponsor analysis. When these exclusions were put back, no clinical effectiveness could be shown.

BENEFIT/RISK

In cases where a sponsor considers a regulatory application pathway involving less than two AWCTs, the sponsor must remember that the application package must thoroughly describe clinical safety in addition to clinical efficacy. The B/R ratio must be established. Ideally, the clinical safety picture will be generated not only by considering the single phase 3 AWCT, but also the phase 2 clinical data. If one or more of the phase 2 studies can be continued via the use of appropriate extensions, this will add to the robustness of the overall clinical safety database. Another way to build the safety database prior to regulatory submission includes running one or more acute or chronic safety trials in parallel with the single AWCT. These safety CTs could be designed to investigate higher doses than those explored in the main efficacy trial, or they could be enriched (19) with the appropriate subpopulations (eg, gender, age, severity and incidence of disease, coadministered medicines, etc).

DOSE RANGE ANALYSIS

We suggest that sponsors should consider extending dose range analysis into phase 3, in particular when planning on a single AWCT. Phase 2 studies are not statistically powered to give accurate and precise measures of adverse drug reactions at each dose level; phase 2 studies are not long enough to detect delayed safety effects. Also, in the shorter duration, more tightly controlled phase 2 studies patients might remain on doses that they would not tolerate in real life or in a phase 3 AWCT, thus giving a false sense of tolerability at a given dose. For instance, in February 2000 Lotronex (alosetron) was approved as a treatment for diarrhea in women at a dose of 1 mg twice daily. The occurrence of constipation was approximately 30% at this dose, which was manageable in early clinical trials, but was problematic in subsequent phase 3 and phase 4 testing (20). The Lotronex label now indicates that patients should begin at 0.5 mg twice daily. Regardless of the path taken to establish substantial evidence of safety, at the end of the journey the sponsor must be able to present a compelling, well-organized, and well-substantiated summary of efficacy and safety. It is not recommended to rely on a single active NME dose level in phase 3, particularly when relying on a single AWCT. In March 2008 Bob Temple (now Deputy Center Director for Clinical Science, CDER, FDA) stated, "recent examples suggest need for more attention to [dose-response exploration] in phase 3; i.e., don't settle on dose too soon" (21). A longer duration of dosing, more heterogeneous population, and larger patient population in phase 3 may point to a different optimal dose than indicated by phase 2 results (ie, different B/R ratios may be uncovered in larger studies that may also have a longer duration of dosing).

DISPROVING A NEGATIVE NOT EASY

An increasingly common pitfall involves the subject of failed primary endpoints or negative trials. For instance, the sponsor may have planned to submit an application based on two AWCTs, but one trial fails its primary endpoint. Subsequently, the sponsor may request that the regulatory agency approves the NME based on a single positive AWCT. A few of the possible arguments include the following: (a) retrospective pooled analysis of the two trials suggests efficacy in a subpopulation or based on secondary endpoints; (b) the negative study should be discounted because it was flawed in trial design; or (c) the positive AWCT has high statistical significance (low *P* value). However, proving that

one positive AWCT constitutes substantial evidence in light of a negative AWCT has been an uphill battle, and sponsors should routinely expect that replication of positive effect in a second AWCT will be needed before desired approval is granted. When Pfizer asked the FDA to approve Neurontin (gabapentin) for the treatment of diabetic peripheral neuropathy (DPN), in 2001 the FDA replied that one positive study for DPN and one negative study for DPN was insufficient, and that another positive AWCT in DPN patients would be needed (22). Ultimately, Pfizer did not complete a second positive AWCT in DPN patients, and as a consequence Neurontin is not indicated for treatment of DPN.

PROMOTION

Although considerations of substantial evidence are often most prominently discussed during the NDA or registration phase of compound development, they should be considered throughout the product's life cycle. Indeed, when evaluating promotional claims that fall outside of FDA-approved product labeling, sponsors must evaluate whether clinical data of approved drug products meet the requirements of substantial evidence, as noted in 21 CFR Part 202 (23).

Sponsors would be remiss to assume that FDA's Division of Drug Marketing, Advertising, and Communication (DDMAC) subscribes to a different definition of substantial evidence than FDA Review Divisions. In fact, from a regulatory perspective, the substantial evidence definition used to support approval of the NDA equally applies to the substantiation of product promotional claims, and evidence from DDMAC enforcement actions corroborates this basic premise. In simple terms, if there were important, relevant reasons for the FDA Review Division to grant drug approval based on a single phase 3 AWCT, then DDMAC should be equally accepting of promotional claims based on that single AWCT. The sponsor should not, however, expect that a single AWCT will be substantial enough to support all promotional claims.

Sponsors may wonder if there are exceptions to the rule that seem to allow a lower standard to be applied for product promotional claims. If sponsors seek insight into DDMAC thinking on this subject, the authors recommend monitoring publicly available DDMAC enforcement letters and applying these opinions when evaluating their own evidence. If after FDA approval the sponsor concludes that a second AWCT must be conducted to support promotion of a particular claim, the sponsor must weigh the cost, medical necessity, and benefits of conducting clinical trials purely for product promotion, rather than for new indications or labeling updates.

Recent enforcement actions by DDMAC indicate that substantial evidence requirements are a key consideration when evaluating the appropriateness of promotional material. The author's analysis of DDMAC publicly available enforcement letters (24) indicates that inadequate or absent substantial evidence is cited in 74 out of 118 (63%) warning and "untitled" letters from 2005 through 2009. However, it is relatively uncommon for DDMAC to specifically mention when data was inadequate simply due to lack of replication. Our analysis revealed that lack of replication was cited in only 8% (10/118) of letters reviewed.

Table 2 provides examples of enforcement actions (between 2005 and 2009) that have included violations related to failure to meet substantial evidence standards which specifically mention that a single AWCT was unable to meet the substantial evidence standard (24). For example, in April 2009, DDMAC took enforcement action against Sanofi-Aventis regarding promotional activity related to its oncology product Taxotere (docetaxel). Specifically, DDMAC objected to the use of a professional reprint carrier describing the results of a study designed to demonstrate the superiority of docetaxel versus paclitaxel. DDMAC asserted in their enforcement letter that the study failed to meet the substantial evidence requirements based on the following: first, the study failed to meet its primary endpoint, thereby invalidating the secondary endpoints used to support the product promotional claim; second, the study results were not replicated in an additional

TABLE 2

Select FDA Enforcement Actions Related to Promotional Materials and Substantial Evidence

Product and Date of Enforcement	Type of Letter	Proposed Claim	Type of Citation, Proposed Evidence From Sponsor, and DDMAC Rationale for Their Objections to Sponsor's Claims
Zyvox, July 2005	Warning letter	Zyvox is superior to vancomycin for the treatment of nosocomial MRSA.	Implied superiority claims: This claim was not based on two adequate, well-controlled studies, but rather on one study that was conducted prior to marketing and continued as a postmarketing study. Furthermore, the claim was based on a post hoc subgroup analysis; *P* values are meaningless in this case.
Survanta, July 2005	"Untitled" letter	Although Survanta infants weigh significantly less at entry, survival is better in the <600 g infants; 17 of 23 Survanta-treated infants survived, compared to 11 of 30 Infasurf-treated infants ($P = 0.007$); other similar statements made.	Unsubstantiated effectiveness claims: Study cited as supporting claim was an analysis of a patient subgroup and involved unplanned subset analysis. Also, data presented were at odds with other evidence relating to the survival rate. Because of such conflicting evidence, the result from the study needed to be replicated by prospective, randomized, controlled clinical studies designed to measure survival rate to be credible.
Loprox, January 2006	"Untitled" letter	Loprox is effective for long-term maintenance treatment of seborrheic dermatitis.	Broadening of indication: Medicis had submitted a supplemental NDA seeking approval for this indication, but this indication was determined to be not approvable by the Review Division because only one of the two trials showed a reduction of relapse rate for patients treated with Loprox. DDMAC applied the same standard of substantial evidence for these promotional claims.
Benicar, January 2006	Warning letter	Benicar is superior to Diovan, Cozaar, Norvasc, Plendil, and Avapro (suggestive language to this effect).	Unsubstantiated superiority claims: Studies supporting claim did not generate valid data due to being (a) open-label and uncontrolled, (b) involving meta-analyses and (c) titration-to-effect comparisons, and (d) did not compare treatments at maximum dosages. DDMAC asserted that in general a claim of superiority should be based upon the maximum dose of two drug products in two adequate, well-designed head-to-head clinical trials.
Solaraze, July 2007	"Untitled" letter	Solaraze is approved for use in the treatment of actinic keratoses when used in combination with cryotherapy.	Broadening of indication: The reference provided for the claim was a single, open-label, pilot trial that included 10 patients in each of two treatment arms. This did not constitute substantial evidence for efficacy information.
Geodon, July 2007	"Untitled" letter	Geodon has proven advantages over Haloperidol IM, namely twice the improvement as measured on the BPRS.	Unsubstantiated superiority claims: Study cited was a single, open-label study. This was not an appropriate study design to assess a subjective endpoint.
Sanctura, January 2009	"Untitled" letter	Sanctura provides "day 1 relief."	Overstatement of efficacy: This claim was based on a post hoc analysis of efficacy data, which was taken from one of the clinical studies described in the PI, in which patients with overactive bladders were randomized to placebo or Sanctura. The onset of action was analyzed utilizing a reverse stepwise method. The claim was based on only one of two studies identified as pivotal that were submitted to the NDA.
Taxotere, April 2009	"Untitled" letter	Taxotere was superior to paclitaxel in terms of TTP and response duration, and	Unsubstantiated superiority claims and overstatement of efficacy: The claims referenced an open-label, randomized study. The reference cited in support of these claims did not constitute substantial evidence or substantial clinical

TABLE 2

Continued

Product and Date of Enforcement	Type of Letter	Proposed Claim	Type of Citation, Proposed Evidence From Sponsor, and DDMAC Rationale for Their Objections to Sponsor's Claims
		Taxotere was also superior to paclitaxel in the treatment of metastatic breast cancer.	experience to support these claims and representations because, among other factors, the study failed to demonstrate statistical significance on the primary endpoint and was not replicated.
TriLuma, August 2009	Warning letter	TriLuma cream, when used in sequence with glycolic acid peels, is well tolerated and may yield enhanced results in the treatment of melasma.	Promotion of unapproved uses or broadening of indication: Claim was based on a single nonrandomized open-label pilot study. According to DDMAC, this approach was not appropriate for an assessment of efficacy and was based on only a single study. Overstatement of efficacy: Claim was based on an open-label uncontrolled clinical study. DDMAC stated that results from a single open-label clinical trial with no control group did not constitute substantial evidence or substantial clinical experience to support this, or any other, efficacy claim.
Nalfon, August 2009	Warning letter	Claims made suggesting that Nalfon can treat pain associated with plantar fasciitis, albeit such an indication is not approved in the PI.	Unapproved new use: No specific mention of supporting AWCT data was included in the direct mailer ad. DDMAC noted that to promote the drug for the pain associated with plantar fasciitis, one or more adequate and well-controlled clinical trials evaluating the drug in the treatment of plantar fasciitis would need to be conducted and the PI would need to be updated to reflect this information.

study. In particular, DDMAC stated that "a claim of superiority generally must be supported by two well-designed, head-to-head clinical trials comparing appropriate doses and dose regimens of your drug and the comparator drug" (24). This example suggests that DDMAC utilizes enforcement activity to apply the same substantial evidence standard that would typically be expected for product registration requirements even though, in most cases, DDMAC does not approve promotional materials prior to their dissemination. Granted, in this instance the desired promotional claim for Taxotere could not have been made on the basis of a single trial, as the trial failed its primary endpoint. However, this example does highlight where a lack of replication was specifically cited as a reason not to promote this data and an example of how DDMAC monitors and communicates compliance with the substantial evidence standards.

By definition, data contained in the approved product labeling should reasonably be considered substantial evidence; as such, these data are widely used to support product promotion. If the data supporting a product claim are not contained in the product labeling, the substantial evidence test must be applied. Importantly, it must be determined that promotional claims are considered generally consistent with the approved labeling prior to evaluating whether the available data can be considered substantial evidence. Otherwise, sponsors may, in fact, meet the requirements for substantial evidence only to find that the promotional claims are considered violative because the claims are considered inconsistent with the approved product labeling. To avoid this unfortunate circumstance, it may be advisable to leverage new study data to seek changes to the product labeling rather than simply pursuing a promotional strategy without corresponding data in the product labeling. Sponsors, however, are urged to carefully consider the consequences of submitting a supplemental NDA to seek a change in labeling. A possible consequence could be to open the product to unanticipated label change

requests by the FDA, or to open the enforcement of new requirements by the FDA (eg, new pediatric use labeling, new PMCs, REMS, etc). If sponsors are uncertain about whether a single AWCT is sufficient for promotion, sponsors may consider seeking input from DDMAC or the Review Division via special protocol assessment prior to beginning the trial or seeking DDMAC advisory comments prior to promoting the data from a single AWCT.

CONCLUSION

Reliance on substantial evidence of clinical efficacy and safety involving less than two AWCTs is sometimes appropriate. The evidence suggests that the sponsor should carefully look at confirmatory or supporting evidence when the NDA, BLA, or supplemental application contains less than two AWCTs. Productive discussion is recommended with the FDA at multiple stages of development (eg, EOP1, EOP2, pre-NDA, advisory committees) with regard to the sponsor's substantial evidence strategy. The sponsor should work with the FDA such that each NME-specific substantial evidence strategy evolves appropriately over time—as new scientific advances are made, as new FDA guidances are written, and as new data (nonclinical and clinical) on the NME become available. Evolving science and patient safety must always come first, superseding past agreements with the FDA about substantial evidence study designs. Discussions between sponsor and FDA are a guide, not an ironclad contract that guarantees the FDA will approve the sponsor's drug based on prior discussions and negotiations. In all, substantial evidence of effectiveness and safety based on a single AWCT, and the ability to make promotional claims backed by this evidence, are most likely to occur as an outcome of meaningful discussions with FDA and as an outcome of lengthy, careful, and reasoned consideration by the sponsor.

REFERENCES

1. US FDA. Federal Food, Drug, and Cosmetic (FD&C) Act, chapter V: drugs and devices. http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/FDCActChapterVDrugsandDevices/ucm108125.htm (accessed May 7, 2010).
2. US FDA. Guidance for industry: providing clinical evidence of effectiveness for human drug and biological products. May 1998. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM078749.pdf (accessed May 7, 2010).
3. US FDA. Draft guidance for industry: integrated summary of effectiveness. August 2008. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM079803.pdf (accessed May 7, 2010).
4. US FDA. Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. May 2007. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071590.pdf (accessed May 7, 2010).
5. Gould AL. Substantial evidence of effect. *J Biopharm Stat*. 2002;12:53–77.
6. US FDA. Taxotere (docetaxel) label. http://www.accessdata.fda.gov/drugsatfda_docs/label/2010/020449s044lbl.pdf (accessed May 7, 2010).
7. A randomized trial of propranolol in patients with acute myocardial infarction. I. Mortality results. *JAMA*. 1982;247:1707–1714.
8. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet*. 1988;2(8607):349–360.
9. US FDA. Victoza (liraglutide) label. http://www.accessdata.fda.gov/drugsatfda_docs/label/2010/022341lbl.pdf (accessed May 7, 2010).
10. US FDA. Drugs @ FDA. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/ (accessed August 10, 2010).
11. Code of Federal Regulations. Title 21 Food and Drugs. Part 314.126. http://edocket.access.gpo.gov/cfr_2008/aprqtr/pdf/21cfr314.126.pdf (accessed May 7, 2010).
12. Temple RJ. NDA/demonstrating product effectiveness. Presentation at Unapproved Drugs Workshop, January 9, 1977. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegula

toryInformation/EnforcementActivitiesbyFDA/SelectedEnforcementActionsonUnapprovedDrugs/ucm120000.pdf (accessed May 7, 2010).
13. ICH. ICH E9, statistical principles for clinical trials. http://www.ich.org/LOB/media/MEDIA485.pdf (accessed May 7, 2010).
14. US FDA. Draft guidance for industry: non-inferiority clinical trials. March 2010. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM202140.pdf (accessed August 11, 2010).
15. ICH. ICH E10, choice of control group and related issues in clinical trials. http://www.ich.org/LOB/media/MEDIA486.pdf (accessed May 7, 2010).
16. Powers JH. Increasing the efficiency of clinical trials of antimicrobials: the scientific basis of substantial evidence of effectiveness of drugs. *Clin Trials Antimicrobials*. 2007;45(Suppl 2):S153–S161.
17. MacMahon B, Yen S, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med*. 1981; 304:630–633.
18. Temple R, Pledger G. The FDA's critique of the anturane reinfarction trial. *N Engl J Med*. 1980; 303:1488–1492.
19. Peck CC, Wechsler J. Report of a workshop on confirmatory evidence to support a single clinical trial as a basis for new drug approval. *Drug Inf J*. 2002;36:517–534.
20. Temple RJ. Examples of when protocol design became a critical factor in the success or failure of a product. Presentation to the Drug Information Association, June 21, 2006.
21. Temple RJ. FDA drug approval process, potential efficiencies and active control trials. Presentation to UNC School of Public Health, March 6, 2008.
22. Kessler DA. Expert report, July 31, 2008. http://www.pharmalot.com/wp-content/uploads/2008/10/neurontin-kessler.pdf (accessed May 7, 2010).
23. Code of Federal Regulations. Title 21 Food and Drugs. Part 202. http://edocket.access.gpo.gov/cfr_2008/aprqtr/pdf/21cfr202.1.pdf (accessed May 7, 2010).
24. US FDA. Warning letters and notice of violation letters to pharmaceutical companies. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/EnforcementActivitiesbyFDA/WarningLettersandNoticeofViolationLetterstoPharmaceuticalCompanies/default.htm (accessed May 7, 2010).

The authors report no relevant relationships to disclose.