1503

Public Citizen

Buyers Up • Congress Watch • Critical Mass • Global Trade Watch • Health Research Group • Litigation Group
Joan Claybrook, President

December 23, 1999

Dockets Management Branch (HFA-305)
Food and Drug Administration
5600 Fishers Lane
Room 1061
Rockville, MD 20852
Fax: (301)594-6197

To whom it may concern:

The Draft Guidance on Choice of Control Group in Clinical Trials, prepared as part of the International Conference on Harmonisation (ICH), is a clear attempt by the Food and Drug Administration (FDA) to spread its pro-placebo-controlled trial ideology globally. This proselytizing intent was made clear at an FDA meeting on the use of placebos in clinical trials in which we participated in April of this year. Dr. Robert Temple, Director of the FDA's Office of Medical Policy, stated at the meeting, "And people do active control trials in Europe all the time. Europe is finally getting the idea that they need to add a placebo group to make them informative."[1] This sometimes unethical ideology has been laid out in a series of publications by an FDA employee[2,3,4,5] and would take on added force if this poorly thought-out Guidance were finalized and adopted by other ICH countries.

The zeal to expand the use of placebos in clinical trials has resulted in a document that is so unbalanced that its credibility is undermined. The structure of the document reflects that bias:

●An entire section (section 1.5) is devoted to attacking active-controlled trials; there is nothing similar for any of the other study designs, even clearly weaker designs such as historical controls.

●This section attacking active-controlled trials actually precedes the detailed descriptions of the types of controls, so the reader is poisoned against active-controlled trials before he or she even learns fully about them.

●The purported weaknesses of active-controlled trials are mentioned repeatedly, leading to an extremely redundant and tedious document.

●Ethical considerations are treated as subordinate to supposed data collection needs; ethics does not even appear in the critical Table 1, which describes the attributes of the different trial

Ralph Nader, Founder
1600 20th Street NW • Washington, DC 20009-1001 • (202) 588-1000

designs. The question confronting researchers is not and should not be: "How do I get the most easily interpreted data?" It is: "Given the available study designs, which one will provide the most useful data while maximizing the protection of patients?" When ethical concerns are quite literally out of the picture, researchers will be led to the first question instead. The Draft Guidance is a transparent attempt to legitimize evasions of the clear requirements of the Declaration of Helsinki, which requires that, "In any medical study, every patient - including those of a control group, if any - should be assured of the best proven diagnostic and therapeutic method. This does not exclude the use of inert placebo in studies where no proven diagnostic or therapeutic method exists."[6]

In addition to its attempts to water down the existing ethical codes, the document places undue emphasis on the supposed needs of regulators and pharmaceutical companies (who together make up the ICH) and places these above the needs of patients or physicians. Most patients and physicians have little need for information addressing whether a new drug for a disease for which there already is an effective therapy is better than nothing; they would like to know whether the new drug is better than the existing drug. But the proposed Guidance would drive clinical trials in the opposite direction. While this may make things easier for regulatory bodies, which can approve drugs simply on the basis of superiority to placebo, and to the pharmaceutical industry, which can more easily prove a new drug superior to placebo than approximately equivalent to a known effective treatment, patients will often not receive optimal medical treatment during the trial.

The remainder of our comments address particularly problematic aspects of the Draft Guidance.

Section 1.4.1: Evidence of Efficacy

In the discussion of how to demonstrate efficacy in clinical trials, only the active-controlled trial is singled out for criticism, although all studies can be plagued by difficulties in concluding whether or not a treatment is efficacious. These points criticizing active-controlled trials are repeated in sections 1.4.2, 1.5, 2.4.1, 2.4.4 and 2.4.7.1. There is no similar repetition of possibly disadvantageous aspects of other clinical trial designs.

Section 1.4.2: Comparative Efficacy and Safety

The Draft Guidance states: "The active comparator(s) should be acceptable to the region for which the data are meant." This sentence is extremely vague and opens the door to potential abuses of study participants. Is the document saying that if a comparator is "acceptable" to the region (whatever the region means), that it can be used no matter how weak the science to justify its use? Or that it is acceptable to do a clinical trial if the elites who conduct the trial (and who themselves probably have access to the best medical care) deem the comparator acceptable because participants are too poor to afford better, even if the treatment is second-rate from a scientific perspective? This is the kind of ethical relativism that those intent on watering down

the existing ethical guidelines have been attempting to formalize.[7] (Our comments in this section apply equally to the promotion of ethical relativism in sections 2.1.3 and 2.4.1.)

Section 1.4.2.2: Patient population

This discussion of the dangers of "enriched" study populations and the consequent implications for generalizability makes some legitimate points. But many of these points are relevant to placebo-controlled trials as well. Run-in periods in placebo-controlled trials very often select for patients who are more likely to respond to a drug, by eliminating sicker and less adherent patients. In addition, it is unclear why this particular assault on active-controlled trials appears in section 1.4, which is supposed to be about all clinical trials, not only active-controlled ones.

Section 1.4.2.3: Selection and timing of endpoints

The statement that a treatment may appear effective at a particular point in time and with a particular endpoint, but not at another point in time or with a different endpoint, is valid. The question is why the document implies that this is only a problem for active-controlled studies, when it clearly can occur in placebo-controlled (or other) studies as well.

Section 1.5: Sensitivity-to-Drug-Effects and Assay Sensitivity

In many ways, this section is the centerpiece of the document. It represents the fullest elaboration of the authors' gripes against active-controlled trials. This is true even though section 1 is supposed to be about all clinical trials (not active-controlled trials in particular), most of the criticisms are reiterated later in the document, and there is no similar section to criticize other study designs. The intent of this section seems to be to convince the reader of the possible inadequacies of active-controlled trials before they are discussed more comprehensively in section 2.4, where the reader can learn of some of their advantages.

The issue of assay sensitivity is presented in a manner that seeks to denigrate all active-controlled trials. The authors of the Draft Guidance should answer the following question: Which are the conditions for which consistently effective treatments have been identified? Then, rather than attacking active-controlled trials with a broad brush, the ICH should be putting forth an official list of drugs and doses known to be effective in particular conditions and which the regulatory agencies will accept as positive controls. This would take the onus off the investigator or pharmaceutical company wishing to conduct an active-controlled trial; in the present circumstance they have to make a case for conducting an active-controlled trial to the very regulatory authorities whose hostility toward active-controlled studies is reflected in this document.

Section 1.5.1: Need to Ensure Assay Sensitivity in Noninferiority (Equivalence) Trials

Assay sensitivity is a valid concern, but, like most other criticisms of active-controlled trials mentioned in this document, the criticisms are overstated. Assay sensitivity is a much more legitimate issue for treatments of marginal effectiveness; in some cases, the effectiveness may be so marginal that another placebo-controlled trial may be advisable. But in many cases, there can be little question that the treatment is consistently effective; in that case, identifying a delta such that one could conclude with confidence that the new treatment is better than nothing presents few challenges. Placebo-controlled studies of prophylaxis against recurrent genital herpes, for example, have consistently been strongly positive. In such cases, a delta that could with confidence preclude an erroneous conclusion of treatment effectiveness would be easy to identify.

Indeed, as the authors of the Draft Guidance finally acknowledge in this section, placebo-controlled trials can also fail to yield the correct result due to being underpowered and a consequent lack of sensitivity. Reviews of clinical trials have shown that many placebo-controlled studies are underpowered.[8] But this is only reluctantly acknowledged in this section, whereas the assay sensitivity problems of active-controlled trials are described at length and then reiterated in sections 1.5.2, 1.5.4, 2.4.4, 2.4.7 and 2.1.6.1.

The Draft Guidance describes the addition of a placebo arm to an active-controlled study as a "straightforward" solution to the problems of assay sensitivity. (This point is made again in section 2.1.5.1.1.) This fails to take into account the fact that the very impetus for doing an active-controlled study is often that a placebo-controlled study would be unethical. It is no less unethical to assign a smaller fraction of study participants to an unethical control group than it is to assign a larger fraction of study participants to that group. For the individual who receives a placebo, the number of study arms is irrelevant.

Section 1.5.2: Choosing the Noninferiority Margin

The Draft Guidance states that "there is little published experience on how to" calculate delta. But the FDA itself has guidelines on what delta to use in particular situations in equivalence studies for antimicrobials. Assigning delta is a decision based on experience from prior clinical studies (usually placebo-controlled ones) and clinical judgement. This is not substantially different from deciding what difference one will power a placebo-controlled trial to detect. These decisions always contain an element of arbitrariness. Simply because investigators are, by force of repetition, more comfortable with one form of arbitrariness compared to another, is no excuse for resisting change to a study design with some undeniable advantages. Statistical methods for sample size calculations in active-controlled trials have been published.[9]

The statement that past experience guides the choice of delta and therefore "gives the noninferiority trial an element in common with a historically controlled study" represents the substitution of epithet for argument. Because most researchers harbor a justifiable suspicion of

many historically controlled studies, this is in effect a slur on active-controlled studies. Moreover, since placebo-controlled studies also depend on historical information in their design (see previous paragraph), to make that point only in association with active-controlled studies seems unbalanced.

Section 1.5.4: Assay Sensitivity and Study Quality in Noninferiority Designs

This section is distinguished primarily by identifying a series of problems that exist for both placebo-controlled and active-controlled studies (and in some cases all studies), but then focusing on them as if they were problems only in active-controlled studies. Indeed, the very placement of these critiques in section 1.5, the Draft Guidance's assault on active-controlled trials, instead of in a section that reviews quality characteristics in all clinical trials, emphasizes this point. The following points, listed in this section as disadvantages for active-controlled studies, are true for placebo-controlled studies as well:

- The eight points that reduce assay sensitivity listed in the Draft Guidance;
- The need to pay attention to sensitivity-to-drug-effects;
- The need to scrutinize entry criteria, characteristics of the study population, specific endpoints measured, and timing of assessments;
- The need to examine adherence to therapy, concomitant therapies and dropout rates; and
- The need to pay particular attention to atypical results.

The Draft Guidance also reiterates the allegation made in some of the papers authored by an FDA employee mentioned above and repeated in section 2.1.4 that there is less of an incentive for conducting a high quality trial in an active-controlled study compared to a placebo-controlled one. This contention cannot be supported. It is true that in a placebo-controlled study, in which the object is to prove the treatment superior to placebo, non-differential misclassification (the result of nonspecific "noise" due to a poorly conducted trial) will bias the results toward the null. But in active-controlled studies the hypotheses are reversed:[10,11] the null hypothesis is not that the two study arms are the same (as it would be in a placebo-controlled trial); rather, the null hypothesis is that they are different by at least delta. This null hypothesis must be rejected if the study is to yield statistically significant results and the researchers are to conclude that the treatments under study are equivalent. As in the placebo-controlled trial, non-differential misclassification biases results toward the null. Thus "noise" in the study does not bias the results toward equivalence, and the researcher interested in proving equivalence (the usual intent in active-controlled studies) must strive as hard to maintain study quality as his or her counterpart conducting a placebo-controlled study. (The Draft Guidance makes the same mistake in section 2.4.2, this time in reference to blinding.)

Section 2.1.3: Ethical Issues

The Draft Guidance seeks to create a loophole in what should be an ironclad rule: there should be *no* placebo-controlled trials of conditions in which withholding effective therapy causes serious

harm such as death or irreversible morbidity. The loophole that the Draft Guidance would create would be for situations where "standard therapy has toxicity so severe that many patients will refuse therapy." We do not see the justification for this exception. If patients would refuse to be randomized to the standard therapy, they should not be included in the study, which should randomize participants to the standard and new therapies (not to placebo). There is no need for a new placebo-controlled study in those who refuse the standard treatment; the risks and benefits of the standard treatment should already have been adequately established in the previous placebo-controlled trial(s).

In the following paragraph, the Draft Guidance seeks to expand the use of placebo controls (at least relative to their current use in Europe) in a different area -- the Draft Guidance would permit such controls as long as "there is no major health risk associated with withholding or delay of effective therapy." (This perspective is reiterated in section 2.3.3.) The Draft Guidance is silent on just what constitutes a major health risk. Rather than promoting practices that place patients unnecessarily at risk, the FDA and ICH should formulate a placebo policy that stipulates for specific diseases of particular severities whether and for how long effective therapy can be denied. The FDA and ICH should establish a threshold of acceptable increases in the absolute incidence of adverse events due to the withholding of effective therapy for at least the major clinical syndromes. If, based on previous placebo-controlled trials, withholding therapy can be expected to lead to an excess incidence above that threshold, the placebo would be precluded. (We would agree that trials of treatments for mild, self-limited conditions such as mild pain or seasonal allergies can use a placebo because the risk to patients is minimal.) Rather than providing such direction, the document is mired in nebulous phrases which are intended to expand placebo use. The clause "it is considered ethical ..." is particularly egregious as it lays claim to summarizing existing thinking on this issue, when in fact there is great controversy.[12] What is lacking in this section is a clear statement of the researcher's obligation to provide the patient with either the most effective therapy or an alternative therapy he or she has reason to believe may be about as effective. Instead, the section is concerned with creating as many exceptions to that obligation as possible.

Section 2.1.4: Usefulness of Placebo-controlled Trials

This section extols the benefits of placebo-controlled trials, but fails to acknowledge that efficacy compared with placebo is very often not the most important clinical question. For example, the Draft Guidance insists that placebos are necessary to properly ascribe an adverse effect to a drug. This may at times be true (although not if the adverse event in question is rare in the general population), but if there is already a known effective medication, the more relevant issue is to describe the relative risks and benefits of the standard and new therapies, a description that will not be provided by a placebo-controlled trial. Moreover, the careful selection of doses in a dose-response study should make the placebo unnecessary in most cases (see also section 2.1.5.1.2).

Section 2.1.5.1.3: Factorial/combination studies

While the factorial design is in many respects an advantageous one, from an ethical perspective it brings no advantage over conventional placebo-controlled trials. If both portions of the study involve placebo controls in diseases with known effective treatments, three-quarters of the participants will be receiving at least one placebo and one-quarter will be receiving two placebos.

Section 2.1.5.2.3: Limited placebo period

While this study design does, at times, have merit (subject to our concerns expressed in section 2.1.3), it is incumbent upon the FDA and ICH to list specific examples of diseases and durations for which they believe such a study design would be appropriate.

Section 2.1.5.2.4: Randomized withdrawal

The major disadvantages of this study design are inadequately emphasized. As noted at the very end of this section, study participants who are stabilized on a particular drug have disproportionate numbers of people who are adherent to treatment, respond well to the drug or have few adverse effects. It would be expected that such a person, when taken off an effective drug, will do worse. In general, regulatory approval should never be based solely on a withdrawal trial.

Section 2.1.6.3: Efficiency

The claim is made that "Placebo-controlled trials are efficient in that they can detect treatment effects with a smaller sample size than any other concurrently controlled study." This claim is greatly overstated. While placebo-controlled trials do at times lead to lower sample sizes than active-controlled trials (and therefore a potentially faster, more "efficient" trial), this is dependent upon a series of decisions made by the investigator or by the particular circumstances of the trial. These include the selection of alpha and beta, the expected event rates in the study arms, the difference one would like to detect (in the case of placebo-controlled trials) and the margin within which one is willing to declare two treatments equivalent (i.e., delta, for active-controlled trials). The attached figure makes this point graphically. In this example of perinatal HIV transmission prevention, we assume that the placebo HIV transmission rate is 25% and that the transmission rate is 10% in the standard (long) therapy group. On the x-axis, are possible transmission rates in the alternative (short) therapy group. A value on the x-axis of 15%, for example, corresponds to the alternative therapy transmission rate in a placebo-controlled trial as well as to a delta of 5% in an active-controlled trial (10% in standard therapy group plus delta of 5%). (In this example, we have used one-sided testing for the active-controlled study, an accepted practice in a noninferiority study.[13]) As the figure demonstrates, there are combinations of event rates (to the left of the point of intersection of the two lines in the graph) where the active-controlled trial design leads to higher sample size estimates compared to a placebo-controlled design and other combinations of event rates (to the right of the point of intersection)

where placebo-controlled trials lead to higher sample size estimates. Other authors have come to similar conclusions.[14] The point is that blanket statements of the superior efficiency of placebo-controlled trials are inappropriate (see also sections 2.3.7 and 2.4.6.2).

In most cases, the differences between the sample sizes required for placebo- and active-controlled trials are modest. One can compensate for this supposed benefit of placebo-controlled trial by recruiting more actively or opening additional study sites, a relatively straightforward practice in this era of multicenter studies and human experimentation corporations (HECs). There should be few, if any, situations where this efficiency argument can legitimately be used to justify the unethical research contemplated in the Draft Guidance. This is particularly true when we are, by definition, discussing studies of conditions where a known effective therapy exists and so the efficiency arguments should carry less weight. (Efficiency may be relevant to pharmaceutical companies seeking to market their products as soon as possible, but should not be emphasized at the expense of patient protection.)

Section 2.1.7.1: Ethical concerns

The Draft Guidance needs to back up its assertion that "it would be difficult to conclude that a noninferiority trial [of beta-blockers in postinfarction patients] would have sensitivity-to-drug-effects."

Section 2.1.7.4: No Comparative Information

Seemingly reluctantly, at the very end of the section on placebo-controlled trials, comes the acknowledgment that placebo-controlled trials do not provide the information that most patients and physicians need: "information about comparative effectiveness, information that is of interest and importance in many circumstances." This kind of information is increasingly important as the number of conditions without effective therapy declines and the number of drug choices for a given condition increases. In this sense, the FDA's continued attacks on active-controlled trials are bucking the historical trend in drug development. This is of particular concern because FDA regulations do permit the licensing of new drugs based on active-controlled trials[15] and in whole categories of drugs (e.g., antibiotics, oncologic drugs) active-controlled trials are routinely submitted and approved as parts of New Drug Applications. Because they provide clinically relevant information, the FDA and ICH should be encouraging more of these studies in additional therapeutic categories, rather than encouraging studies that the industry prefers (it is easier to show that your me-too drug is better than nothing than that it is about as good as another drug in the same therapeutic category) and which leave patients at unnecessary risk.

Section 2.4.3: Ethical Issues

This section attempts to create uncertainty over the ethical superiority of active-controlled trials when a known effective therapy exists by fretting that the new drug may be ineffective and that patients in that arm may be at a disadvantage compared to those treated with the established

therapy. What a strange concern, given that this Draft Guidance encourages the use of study arms (placebo) that investigators are *certain* are inferior to available treatment for almost all of the diseases being studied! This section ignores the ethical concept of equipoise: it is ethical to randomize participants to a new, as yet unproven, drug, if there exists sufficient uncertainty among the relevant experts as to whether the new drug is effective/equivalent/noninferior to the comparator. If the researchers are convinced before the trial that the new drug (or placebo) is inferior, the study is unethical and should not take place.

Section 2.4.6: Advantages of Active-Control Trials

Not surprisingly, this section is short, and even then qualified by phrases like "if properly designed." These qualifiers are not generally applied to placebo-controlled trials, even though they should be in many instances.

Section 3.0: Choosing the Control Group

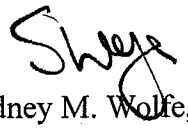
A discussion of adaptive and crossover study designs should be added.

In summary, this Draft Guidance is a remarkably biased description of the advantages and disadvantages of various clinical trial designs. The document continues the FDA's longstanding assault on active-controlled trials and does so at a time where there is less clinical and ethical justification for such trials than ever. Rather than challenging investigators to obtain the best possible data using an ethical design, the Draft Guidance subordinates these ethical concerns to the reflexive tendency of some researchers to prefer placebo-controlled studies, to the short-sighted interpretations of drug regulatory authorities bent on approving any drug as long as it is somewhat better than nothing, and to the concerns of the pharmaceutical industry.

Yours sincerely,

Peter Lurie, MD, MPH
Deputy Director
Public Citizen's Health Research Group

Curt Furberg, MD, PhD
Professor
Wake Forest University School of Medicine



Sidney M. Wolfe, MD
Director
Public Citizen's Health Research Group

1. "The Use of Placebos in Clinical Trials and the Ethics of the Use of Placebos." U.S. Food and Drug Administration, Committee for Advanced Scientific Education Seminar Series, Rockville, MD, April 21, 1999.

2. Temple RJ. When are clinical trials of a given agent vs. placebo no longer appropriate or feasible? Controlled Clinical Trials 1997;18:613-20.

3. Temple RJ. Special study designs: early escape, enrichment, studies in non-responders. Communications in Statistics 1994;23:499-531.

4. Temple RJ. Problems in interpreting active control equivalence trials. Accountability in Research 1996;4:267-75.

5. Temple R. Difficulties in evaluating positive control trials. Proceedings of the American Statistical Association, Biopharmaceutical Section. 1983:1-7.

6. World Medical Association Declaration of Helsinki: recommendations guiding physicians in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, 1964 and revised by the 29th World Medical Assembly, Tokyo, 1975, the 35th World Medical Assembly, Venice, 1983 and the 41st World Medical Assembly, Hong Kong, 1989.

7. Lurie P, Wolfe SM. Proposed revisions to the Declaration of Helsinki: paving the way for globalization in research. Western Journal of Medicine 1999;171:6.

8. Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. New England Journal of Medicine 1978;299:690-4.

9. Makuch R, Simon R. Sample size requirements for evaluating a conservative therapy. Cancer Treatment Reports 1978:62:1037-40.

10. Senn S. Statistical Issues in Drug Development. Chichester: John Wiley & Sons, 1997, pp. 207-17.

11. Hauck WW, Anderson S. Some issues in the design and analysis of equivalence trials. Drug Information Journal 1999;33:109-18.

12. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. New England Journal of Medicine 1994;331:394-8.

13. Machin D, Campbell MJ. Statistical Tables for the Design of Clinical Trials. Oxford: Blackwell Publications, 1987, pp. 35-7.

14. Blackwelder WC. "Proving the null hypothesis" in clinical trials. Controlled Clinical Trials 1982;3:345-53.

15. 21 CFR 314.126(b)(2)

Sample Sizes for Placebo-controlled and Equivalency Studies when Event Rate in Standard Treatment Group is 10%