

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

## **Clinical Trial Design: Considering Usual Medical Care in the Design of Comparison Groups**

### **I. Introduction**

#### **The problem of comparison groups in clinical trials**

Randomized controlled trials, as the evidentiary gold standard for testing health interventions, are prospective, systematic comparisons between groups of patients or subjects, randomized to two or more different arms of a trial. In some cases the selection of comparison arms for a particular trial is challenging. Choices of control or comparator conditions can become bound up in debates about the adequacy of current medical practice, how much scientific evidence exists to support existing interventions, or the value of competing interventions from the point of view of trade-offs among multiple outcomes, patient preferences, cost-effectiveness, or other criteria. This analysis aims to provide guidance on how to determine an appropriate comparator group when background conditions of medical practice are complex.

#### **What kind of clinical trials are included in the scope of this analysis?**

This analysis is focused only on research trials with human subjects that a) involve prospective assignment to an intervention; b) in an area of medical practice that has at least one existing intervention; c) have two or more comparison arms, d) and have health related outcomes. Therefore retrospective studies, uncontrolled trials, and prospective observational studies are not addressed.

In addition, this document principally addresses trials that are designed to be relevant to clinical practice, rather than those designed to measure basic biological processes or disease mechanisms. The latter distinction is made along a continuum of trials that often have some characteristics of both of these two broad categories. Trials may measure a handful of primary and secondary outcomes, some of which are directly relevant to clinical decision-making and others of which inform more long-term research goals related to disease mechanisms or biological effects of interventions.

### **II. Reasons to use a comparison group based on current medical practice.**

There are a number of scientific, ethical and practical considerations that drive decisions to use some form of usual medical care as a comparison group in a trial.

Scientifically, the motivation usually is the need to establish the superiority or non-inferiority of either a new intervention, or a competing existing intervention, to current medical practice, through direct comparison in the setting of a randomized trial. In

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

contrast, a trial testing an intervention to placebo or no treatment provides direct evidence about absolute efficacy of an intervention. However, it is often more relevant to clinical practice to determine which intervention or approach is best, among two or more competing alternatives, based on outcomes measured in the trial.

Furthermore, the existence of effective interventions often creates an ethical imperative to provide at least one of the effective methods to participants in a clinical trial, while recognizing that randomization to different arms in a trial means that, at the outset, no participant is guaranteed to receive one particular intervention. There is active debate about whether researchers ought to be obliged to provide effective interventions that are not routinely used in clinical care in the setting of the research. A particular subset of these situations will be relevant to this discussion, in which a trial might be designed with a usual care arm that mimics typical clinical practice, and does not incorporate effective interventions when they are not part of usual medical practice.

Practically speaking, use of current medical practice in one arm of a trial simply involves less disruption of everyday clinical care for the group of patients who are randomized to that arm. Practical considerations become especially critical in situations where a validated treatment is not widely or consistently used in everyday practice, and where implementing it in a trial context is logistically challenging, prohibitively costly, or viewed unfavorably by the community of physicians. The difficulty in implementing a validated intervention may be part of the impetus for conducting a trial but may also confound attempts to determine the most ethical and scientifically valid comparison group.

### **III. Determining the purpose of an individual clinical trial**

The purpose of the trial determines the range of possible trial designs. In deciding on the main purpose of a trial, a thorough understanding of previous evidence on the topic is needed.<sup>1</sup> Since the social value of the research depends heavily on the context, the need to review and analyze the existing sources of evidence becomes not only a scientific but also an ethical priority.<sup>2</sup> New information needed for clinical practice may not be obvious from review of previous trials; additional sources of information may be needed. Barriers to implementation of previously validated interventions may not be obvious without further research.<sup>3</sup> Background conditions may inform trial design on at least two levels: they can provide impetus for conducting the trial in the first place, as well as being used to formulate a comparison group in the trial.

Two different kinds of purposes can be described in broad terms: a) testing a new intervention against some existing practice, or b) testing existing practices against each other. The practical implication of trials in category b) is that by definition, some of the

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

*Oct 19<sup>th</sup>, 2005*

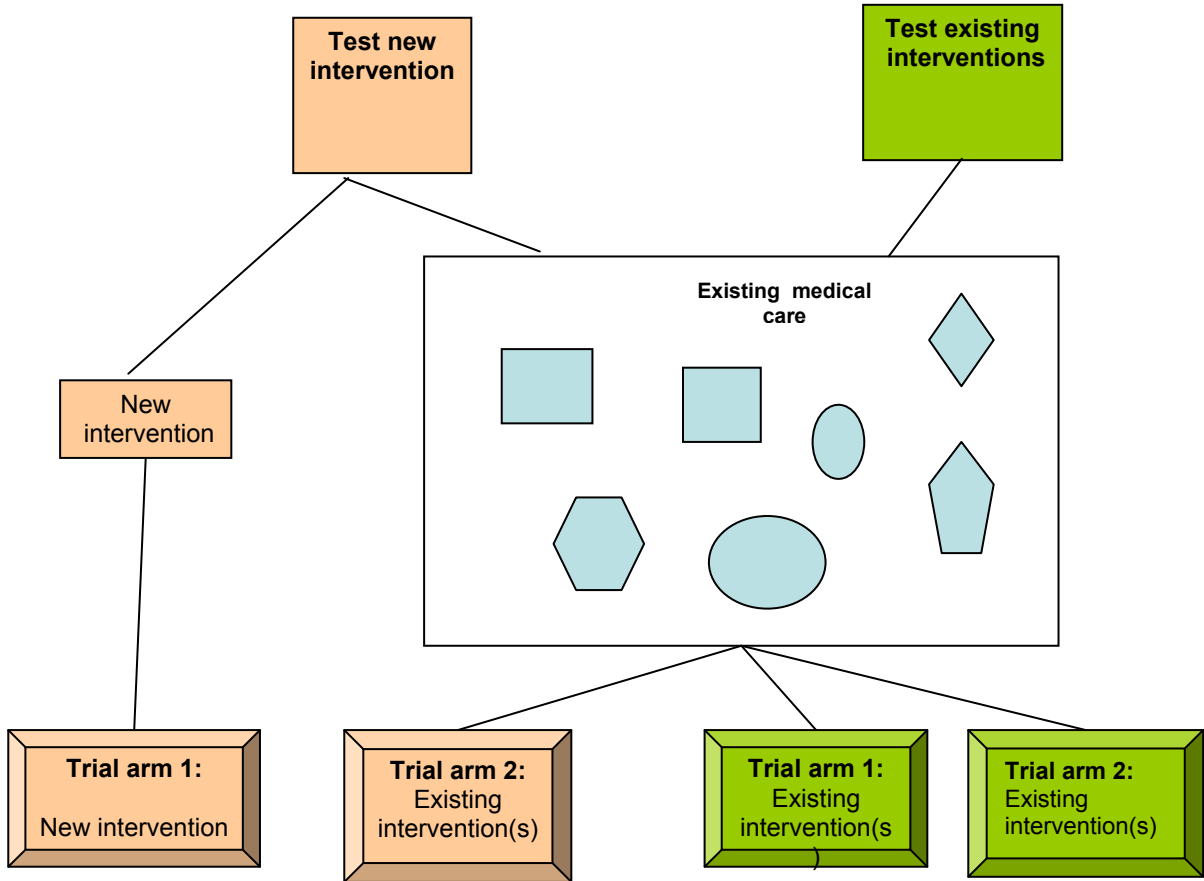
components or modalities used in current practice must be separated from each other in distinct arms of a trial. Within this very general description, however, there is a wide range of possibilities about how any particular trial should dissect the individual interventions or components of care. Trials testing new interventions might be constructed as trials comparing a new intervention to a single method or treatment currently used, or might use a comparator arm encompassing a range of usual care practices, or some combination or variation on these. In add-on trials, existing and new interventions are used in combination, for example, a comparison group consisting of existing intervention A plus new intervention B, versus an arm receiving intervention A plus placebo. A huge number of variations of these design elements are possible. The following are a few examples of the many different kinds of possible comparison groups that may represent some elements of usual care:

- 1) Two or more comparison groups are constructed, each representing different components of usual care;
- 2) A comparison group is a protocolized version of usual care;
- 3) A comparison group is a flexible usual care group, consisting of heterogeneous procedures;
- 4) A trial consists of usual care provided to both groups, plus an add-on intervention to one arm only;
- 5) A comparison group is protocolized version of care and does not represent current practice.

The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.

Oct 19<sup>th</sup>, 2005

### Purpose of trial: two examples:



*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

#### **IV. Issues to consider in formulating comparison groups representing current medical care.**

There are four kinds of factors, occurring individually or coexisting, that often pose challenges to defining comparison groups based on current medical care.<sup>4</sup>

##### **1. Disputes about interpretations of evidence.**

Frequently there is disagreement and debate in the expert community about interpretation of currently available evidence. This can lead to further disputes about whether current treatments have been validated by research, or not, as well as disagreements about which of several treatments should be the current best standard.

Disagreements about interpretation of existing evidence can lead to dispute about what research question is the most relevant to current needs.<sup>5</sup> When there is lack of agreement as to which set of practices ought to be held up as the gold standard, the structure of a standard comparison group becomes problematic as well.

Designs that directly address the source of the controversy are valuable to the research and clinical care communities. However sometimes it is impossible to even design a study to address the disagreements, since the views on the scientific validity of any particular design always depend upon interpretations of prior evidence. Disagreements in about existing evidence also lead to disputes about the risk-benefit profile of a given design. If clinicians believe that one intervention or another is inferior based on their interpretation of the evidence, trials which randomize subjects to one group or the other will be considered ethically dubious. In contrast, those who believe that there is no clear evidence supporting one intervention over another may advocate for a trial to rigorously test the competing interventions. Therefore, in the case of these fundamental disagreements about evidence, there may not be any single trial design that can satisfy all critics, precisely because of the divergence of views on the evidence base.

In these situations, the most important first step is to correctly identify the source of the disagreement about evidence. Then at least the problem of competing interpretations of existing evidence can be the object of discussion, rather than the problem of competing trial designs, which is in fact, a consequence of the more fundamental issue.

##### **2. Lack of adherence to, or translation of, evidence-based recommendations or practice guidelines.**

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

There may be a proven intervention that is not widely used.<sup>6</sup> There is general agreement among experts about interpretation of evidence, but there are other barriers or factors that lead to lack of implementation or translation of these research results into clinical practice, such as:

- lack of physician belief in the treatment
- difficulty in implementing;
- cost;
- side effects;
- lack of physician expertise;
- patient heterogeneity;

The choice of research question and overall study design may depend on an analysis of the specific factors underlying the lack of widespread use of an existing treatment.

In these situations, usual medical practice, used as a comparator arm, may expose subjects to less than optimal medical care in the context of the trial. This situation is problematic on several levels. There is often disagreement about whether this kind of trial design can be justified on the basis of the overall societal benefit to be gained from the research.

Research conducted in these background conditions might be designed to develop and test a new intervention that could supplant the currently used suboptimal intervention with an alternative that is effective but that also is cheaper, easier, or in some other way more accessible. If this goal is within reach, it would be reasonable to use the best known method as a comparator, and the scientific goals and ethical mandates do not conflict. However, if there is a substantial likelihood that the new method will be inferior, a quandary remains.

Generally, non-inferiority trials require greater numbers of subjects than superiority trials, for statistical reasons. If a new intervention is compared to best methods, the feasibility of conducting the non-inferiority trial might be a limiting factor in getting the research off the ground. A superiority trial comparing a new method to an inferior method in current use might be more feasible, raising the question of whether it is preferable to forgo the research altogether or forge ahead to testing an intervention using a comparison group that experts recognize represent inferior medical care, based on the premise that this mimics the care actually delivered in current practice.

One goal of a clinical trial might be to test an intervention delivered rigorously according to best practices (a “Cadillac version”) to the same intervention as used in current clinical practice. Testing interventions delivered according to rigorous standards against a usual

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

care comparison group can help clarify whether outcomes are measurably improved using the best practices approach. The acceptability of this design might depend on whether the best practices are already universally recognized as more effective, or whether this is still an open question.

Ethical questions arise when sub-optimal usual care is provided in the context of a trial. Some argue that as long as none of the comparison arms is known to be inferior to the care usually rendered to the subject in his home community, the trial is ethically acceptable. However, our sense of justice is predisposed against the idea that certain subjects are eligible to be exposed to greater risk because they usually receive a lower quality of medical care. The acceptability of exposure to sub-optimal care will depend in part on whether serious or irreversible harm could occur in the patients receiving less than best care, when that harm could have been prevented by receiving best known care during the trial.

Where the prevailing, sub-optimal intervention is effectively nothing, investigators face a very serious question of whether to use a placebo as a control or the evidence-based intervention. Clearly, the risk-benefit profile for subjects favors the latter over the former. However, there are existing guidelines for the use of placebos<sup>7</sup> and these, at the very least, show us cases where placebo use might be countenanced.<sup>8</sup>

### **3. There is no single “best” intervention: different interventions have trade-offs in terms of different outcomes or side-effects.**

There may be two or more interventions for a single condition, involving multiple outcome measures. The interventions are characterized by different profiles of performance across the different outcome measures, or different side effects, and thus there are trade-offs between or among interventions.

Choices among interventions may be made on the basis of disease or patient characteristics, or on the basis of physician or patient preferences, or all of these. For example, more aggressive chemotherapy, versus less aggressive, may have a greater likelihood of producing severe side-effects, with only minimal gain in survival, or with a small of chance of gain for an individual patient. Choices between regimens would involve balancing the importance of traditional measures of effectiveness with concerns relating to quality of life.<sup>9</sup>

For some medical conditions, physician assessment of individual patient characteristics is the common paradigm for determining best treatment.<sup>10</sup> There is an inherent scientific challenge in measuring the effectiveness of physician selection of treatments, when multiple patient characteristics are involved in this decision-making. One obstacle is that

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

it would require an impossibly large trial to encompass all the stratified patient subgroups needed to rigorously test all the factors that go into treatment decisions. Physician decision-making could be evaluated in a usual care group that allows choice of treatment, compared to a group with pre-determined treatment plan. When there is variation among physicians in decision-making about individual patient treatments, it also will be difficult to know which set of criteria actually used by physicians in a usual care arm is the best possible set, given that all the sets of criteria used by all the physicians in a usual care arm are tested as one “procedure.”

Generally, when “packages” of care, or groups consisting of heterogeneous practices, are included as one arm of a trial, inferences can more reliably be made about the effectiveness of the “package,” or the trial arm as a whole, than about its individual components. In trials that separate individual components or approaches into separate arms, while effectiveness of each component is measured, the degree to which other contextual factors affect outcomes is not directly tested in the trial.

When available interventions present trade-offs of differential performance across multiple outcomes, patient preferences may be particularly relevant, as well as physician choices. Patient choice of treatment can affect overall effectiveness of the intervention when quality of life measures are involved. Allowing patient preferences to play a role is problematic in randomized trial design, but has been addressed by allowing patients or doctors to choose their treatment within a usual care arm. Randomization to the arm as a whole then allows comparison of a usual care approach to an alternative intervention. For example, a randomized trial of high versus low intensity weight training for clinical depression in older adults included a “general practitioner care” group.<sup>11</sup> In the latter arm of the trial, approximately half the patients received an intervention, either pharmacotherapy or counseling. The advantage of this flexibility is that preferences are incorporated without disrupting the overall randomization; however, inferences can only be reliably made about the usual care arm as a whole, not its subgroups. In some cases use of a usual care group may increase relevance of trial data to actual practice, and may increase patient and provider willingness to participate and satisfaction with a trial.

Some of the inferential problems relating to heterogeneity of interventions within an individual arm of a trial may be addressed with post-hoc subgroup analysis, although there are limitations to this approach.<sup>12</sup> Post-hoc subgroup differences are often spurious associations driven by chance that fail to be confirmed in subsequent research designed to test them directly.<sup>13</sup>

If usual care practices change during the trial period, this may complicate data analyses. There is potential for contamination of the usual care arm with the experimental arm of the trial, thus effectively reducing the ability to detect differences between arms. Another

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

methodologic problem may be compromise of blinding in usual care arms, which would be particularly relevant to interventions sensitive to patient behavioral factors, or to bias in assessment, or other sources of bias.

#### **4. Lack of, or insufficient, evidence base for existing interventions.**

In this scenario, interventions are commonly used and there is little, or insufficient, or outdated, rigorous evaluation of them through clinical trials. An issue related to lack of evidence is that existing clinical trial data may be of poor quality; for example, studies may be underpowered,<sup>14</sup> randomization inadequate, blinding incomplete, or other problems. Also, some studies may not be relevant or generalizable to current needs. Frequently the patient population of clinical trials is narrower and distinctly different from the broader population that may experience the intervention in clinical practice. The net effect of insufficient information from well-designed and relevant trials is that it may not be at all clear which treatment is preferable among the choices available, or even if a given treatment is better, or worse, than nothing.

In situations where there is a fundamental lack of evidence, or where the evidence lacks applicability to the current uses, a high priority may exist for addressing this evidence gap. Among the many possibilities are:

- 1) Trials with multiple arms comparing existing interventions
- 2) Trials comparing a single intervention to a heterogeneous mix of existing interventions (usual care in the broadest sense)

In the absence of a solid evidence base, there are likely to be a multitude of existing interventions. If there is a genuine lack of conviction for any particular intervention, trials with a single interventions in each arm and, where permissible, a placebo, are able to provide information on relative and absolute efficacy for a range of interventions about which there is uncertainty.

For example, the NIMH-funded CATIE trial tested 8 antipsychotic drugs currently used in schizophrenia treatment in 1600 patients over a period for 18 months, and will provide data not only on first line use of the drugs, but on outcomes for patients who need to switch drugs due to non-response. Each of these drugs has been tested to some extent in placebo-controlled trials, but there is insufficient information about the relative merits of the drugs from head-to-head trials, and little to guide choices among them for first line or second line treatment.

An alternative design in an area with an insufficient evidence base would be to compare a new intervention (or one from the plethora of existing interventions) to a usual care arm

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

where clinicians are given choice of treatment from among most common interventions. The principal problem with this design is the limitations on inferences that may be drawn from trial results. In non-inferiority trials, inferences could be problematic in situations where evidence for the overall effectiveness of the usual care group is not solid enough.<sup>15</sup> Also, different levels of evidence may exist for different interventions used in the usual care group. Absent clear evidence about each of the usual care interventions, if results from a trial comparing a single intervention to a usual care mix do not show that the single intervention is clearly superior or inferior to the mix, the trial may reveal little information which is crucial for clinical decision making.

## **V. Conclusion**

Comparison groups that represent some form of current medical practice can be constructed a number of ways. A comparison group can consist of a single intervention which is used in clinical practice, multiple arms consisting of different existing care practices, a usual care group reflecting a range of practices, or a variation of these. The most central principle is that the design should fit the purpose of the trial. Even this most basic assertion may lead to conflict, however, when there is no consensus on what the purpose of the trial ought to be.

The purpose of a given trial itself is often driven in part by complex background conditions of medical practice. Therefore, an assessment of current medical care—its pattern of usage, and evidence base—is critical to clinical trial design on several levels: it helps determine the highest priority for filling information gaps; it often is relevant to attempting to reduce barriers to translation of research results into practice; and it forms part of the structure of the trial itself, in terms of comparison groups. Disagreements about evidence relating to current practice often lead to disputes about what, in fact, the most useful and relevant research question really is, in a given situation.

Four kinds of complicating factors often make formulation of a comparison group particularly difficult: a) disputes about current evidence; b) wide variations in utilization of best methods, or widespread failure to use these methods; c) trade-offs relating to physician and patient preferences for different treatments; and d) insufficient pre-existing evidence base to guide rational treatment selection.

Trial design choices reflect how much emphasis is placed on factors such as external validity; measurement of care delivered in typical community care situations versus more controlled trial settings; measurement of a package of services or treatment algorithm versus individual components of care; testing a single treatment versus a complex mix of choices in a single arm; and assessment of a new intervention or validation of existing

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

*Oct 19<sup>th</sup>, 2005*

methods. Selection of scientific priorities is related to the overall purpose of the trial and its relation to the background conditions of medical practice.

When debates about trial design arise, it is important to localize the source of the fundamental disagreement. Differences of opinion can arise as a result of disputes about interpretation of prior evidence or about what research question is most important, both of which then lead to disputes about trial design. There can also be differences of opinion about the evidentiary or ethical strengths and weaknesses of different trials. In many situations more than one scientifically and ethically valid trial design is possible and experts may disagree about which one is preferable. Because of the multitude of scientific concerns and ethical priorities that go into any clinical trial, there is no one formula or approach that can be generally applied to specifying a trial design.

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

Oct 19<sup>th</sup>, 2005

---

<sup>1</sup> Working Group. Users' guides to the medical literature. IX. A method for grading health care recommendations. *JAMA* 1995;274:1800-04. Guyatt GH, Sackett DL, Sinclair JC, Hayward RC, Cook DJ, Cook RJ; Chalmers; Ioannidis et al *JAMA* 2001;286:821-830. Comparison of RCTs and observational studies; Norris SL and Atkins D. Challenges in Using Nonrandomized Studies in Systematic Reviews of Treatment Interventions. *Annals of Internal Medicine* 2005;142:1112-1119; Ioannidis et al *JAMA* 1998

<sup>2</sup> Ferguson D, K Glass, B Hutton, S Shapiro. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? *Clinical Trials* 2005; 2: 218–232.

<sup>3</sup> Irwig L, M Zwarenstein, A Zwi, I Chlamers. A flow diagram to facilitate selection of interventions and research for health care. *Bull World Health Organ.* 1998;76(1):17-24.

<sup>4</sup> In some specific cases the state of medical practice can be characterized with all four of these factors. For example, drug-eluting stents (DES) have been shown in clinical trials to reduce re-stenosis and target lesion revascularization after percutaneous coronary intervention (PCI). In one study the use of DES reduced target lesion revascularization by 73%. Referrals for stenting have increased and bypass surgery rates have declined since FDA approval of the first DES in 2003. However, DES are expensive (\$2400 to \$3000 per stent), and in one report, usage varied significantly depending upon public versus private coverage (25% versus 75%, respectively). In terms of effectiveness, while they reduce revascularization, they have no overall effect on mortality or MI rates, and they are now being used in patients unlike the patient population tested in pivotal clinical trials. Cost effectiveness analysis from one trial indicated a cost of \$27540 per QALY gained. Therefore, there is evidence of effectiveness, and they are being used in some, but not all cases; it is not clear how broadly their use should be supported nor if their cost is justified. References: TAXUS IV trial. Teirstein PS. A chicken in every pot and a drug-eluting stent in every lesion. *Circulation* 2004;109:1906-1910. Chew DPB. Cost-effectiveness of drug-eluting stents: if only all things were equal. *Medical Journal of Australia* 2005;182:376-377.

<sup>5</sup> See Eichacker PQ et al, Meta-analysis of acute lung injury and acute respiratory distress syndrome trials testing low tidal volumes. *Am J Respir Crit Care Med* 2002;166:1510-1514; and commentary by Brower RG et al. Meta-Analysis of Acute Lung Injury and Acute Respiratory Distress Syndrome Trials—letter. *Am J Respir Crit Care Med* 2002;166;1515-6.

<sup>6</sup> Tcheng Je, Madan M, et al. 2003. Ethics and Equipoise: Rationale for a Placebo-Controlled Study Design of Platelet Glycoprotein IIb/IIIa Inhibition in Coronary Intervention. *J. Interventional Cardiology* 16(2):97-105; Mann H and AJ London. Equipoise in the Enhanced Suppression of the Platelet IIb/IIIa Receptor with Integrilin Trial (ESPRIT): a critical appraisal. *Clinical Trials* 2005; 2: 233–243.

<sup>7</sup> “The Use of Placebo Controls in Clinical Trials” <http://www.ama-assn.org/ama/pub/category/8424.html> last accessed 11 Aug 2005

<sup>8</sup> ICH E10

<sup>9</sup> Johnson N et al. Methods of hysterectomy: systematic review and meta-analysis of randomized controlled trials. *BMJ* 2005;330:1478.

<sup>10</sup> Reference criticism of SPORT trial from surgeons who argue that good surgical candidates were excluded from surgery arm.

<sup>11</sup> Singh NA et al. A Randomized Controlled Trial of High Versus Low Intensity Weight Training Versus General Practitioner Care for Clinical Depression in Older Adults. *J. Gerontology.* 2005;60A(6):768-776.

<sup>12</sup> For example, subgroup analysis in National Emphysema Treatment Trial (NETT); overall results comparing lung reduction surgery to medical management for emphysema patients were negative, but some subgroups had positive results; some subgroups had early negative results and were discontinued in trial. However, subgroup analysis was not embraced as necessarily robust enough for the policy decisions based upon it; high costs of conducting the trial precluded follow-up trials to test the subgroup findings directly.

*Considering Usual Medical Care in the Design of Clinical Trials: Scientific and Ethical Issues*  
*Draft discussion document for NIH clinical trials meeting, November 14-15, 2005.*

*The findings, opinions, and conclusions are those of the authors. They do not reflect any position or policy of any federal agency. This draft is made available to you for comment and discussion; no final document or position on this issue has been disseminated by any Federal agency. Please do not cite or distribute.*

*Oct 19<sup>th</sup>, 2005*

---

<sup>13</sup> DeMets and Califf. Lessons learned from recent cardiovascular clinical trials I. *Circulation* 2002;106:746-751.

<sup>14</sup> Jadad AL, et al. Treatment of Attention-Deficit/Hyperactivity Disorder. AHRQ evidence Report/Technology Assessment Number 11. November 1999. AHRQ Publication No. 00-E005

<sup>15</sup> McAlister FA and DL Sackett. Active-control Equivalence Trials and Antihypertensive Agents. *American Journal of Medicine*. 2001;111:553-558.