EFFECTIVENESS GUIDANCE DOCUMENT

Pragmatic Phase 3 Pharmaceutical Trials:

Recommendations for the design of clinical trials that are more informative for patients, clinicians, and payers

Release Date: September 14, 2010

EFFECTIVENESS GUIDANCE DOCUMENT WRITING TEAM

Seema Sonnad PhD; Jennifer Goldsack MChem; Penny Mohr MA; C. Daniel Mullins PhD; Danielle Whicher MHS

Acknowledgements: The writing team would like to thank the following individuals for the thoughtful comments and edits they provided on draft of this Effectiveness Guidance Document: Marc Berger MD; Jesse Berlin Sc.D.; Judith Kramer MD, MS; Michael Lauer MD; Eleanor Perfetto PhD, MS; Jodi Segal MD, MPH

EXECUTIVE SUMMARY

PREFACE

The Center for Medical Technology Policy (CMTP) supports the development of Effectiveness Guidance Documents (EGDs) to provide specific recommendations on the design of prospective studies that will inform decisions by patients, clinicians and payers. The recommended methods aim to achieve a balance between internal validity, relevance and feasibility. EGDs are intended to be analogous to Food and Drug Administration (FDA) guidance documents, but are focused on the design of clinical studies to support clinical and health policy decision making. EGD recommendations are developed through an extensive consultative process involving a broad range of experts and stakeholders. Full details about the EGD development process are available at http://www.cmtpnet.org/cmtp-research/guidance-documents/EGDProcess.pdf. *A list of the experts and stakeholders who provide input during the development of this document are included in appendix A.*

The goal of this EGD is to provide specific recommendations to guide the implementation of more pragmatic study designs for phase 3 pharmaceutical trials. Pragmatic Clinical Trials (PCTs) for pharmaceuticals are prospective studies designed with the specific aim to assist patients, clinicians, and payers in making informed decisions about alternative drug therapies. The ten recommendations outlined in this guidance document cover broad topic areas including: enhancing stakeholder engagement in study design, aspects of trial design, and other operational, analytical and ethical aspects of using pragmatic designs for regulatory approval trials. This guidance is not intended to replace guidance issued by the FDA or other agencies. The objective is to provide recommendations for incorporating pragmatism into Phase 3 clinical trials, while simultaneously meeting regulatory requirements of the FDA. A major conclusion of this work is that any incremental steps towards improving the pragmatic nature of trial design in three major domains of trial design; improving the generalizability of the patient population, selecting active comparators and selecting consistently measured, clinically-relevant outcomes, can markedly improve the utility of information obtained from clinical studies designed for regulatory approval. Other features of pragmatic trial design, such as loosening tight controls on patient adherence with therapy may be better reserved for study in the post-regulatory environment.

INTRODUCTION

Overview

The Center for Medical Technology Policy (CMTP) supports the development of Effectiveness Guidance Documents (EGDs) to provide specific recommendations on the design of prospective studies that will inform decisions by patients, clinicians and payers. The recommended methods aim to achieve a balance between internal validity, relevance and feasibility, with the goal of providing decision makers with studies that allow a reasonable level of confidence that the intervention improves net health outcomes. EGD recommendations address patient inclusion/exclusion criteria, choice of comparators, selection of outcomes, duration of follow-up and other key elements of trial design that are relevant to the specific topic of each guidance.

The target audiences for these study design recommendations include clinical researchers, product developers and research funding organizations. It is also anticipated that organizations developing evidence-based policies will consider these recommendations as they assess the quality and relevance of studies in their evidence reviews. EGDs are intended to be analogous to Food and Drug Administration (FDA) guidance documents, which are also targeted to product developers and clinical researchers, and provide guidance on the design of clinical studies that are intended to support regulatory decision-making. EGDs are intended to serve a comparable function, but are focused on the design of clinical studies to support clinical and health policy decision making. In this respect, they provide methodological guidance for the design of prospective comparative effectiveness research (CER) targeted to specific health technologies or services.

EGD recommendations are developed through an extensive consultative process involving a broad range of experts and stakeholders, including mechanisms for broad public review and comment. The process is guided by a technical working group (TWG) selected by CMTP staff, which consists of about 8-12 individuals with expertise in patient care and research methods specific to the clinical domain that is the focus of the EGD. Initial draft recommendations are developed by CMTP, with input from the TWG as well as interviews with numerous other experts and stakeholders. These recommendations are then discussed in detail during a day-long expert-stakeholder methods symposium involving approximately 45-60 participants including patients, consumers, clinicians, payers, regulators, product developers, clinical researchers and methodologists. Revised draft EGD recommendations are then developed and broadly circulated for public comment. The recommendations are further revised based on public feedback and the resulting EGD is posted on the CMTP website, where it is available for review or download. There is no charge for access to or use of the EGDs. EGDs are updated as new scientific evidence, methodological advances and technologic improvements emerge.

As is the case with FDA guidance documents, EGD recommendations are advisory in nature, and it is anticipated that the recommendations will not be applicable to every conceivable study protocol. To the extent that the recommendations accurately reflect the information needs of patients, clinicians and payers, EGDs are likely to be helpful to product developers and researchers who intend for their research findings to impact clinical and/or health policy decisions.

Full details about the EGD development process are available at http://www.cmtpnet.org/cmtp-research/guidance-documents/EGDProcess.pdf. Information on the members of the TWG, participants in the expert-stakeholder methods symposium, and the names of individuals that provide input during the development of this EGD can be found in appendix A. The recommendations and other content of this EGD do not necessarily reflect the views of the TWG, meeting participants or others whose names are listed in the appendix.

Purpose and Process

Patients, clinicians, payers and policymakers increasingly are interested in understanding the comparative and "real world" effectiveness of pharmaceutical products, and often note that traditional clinical trials performed for regulatory approval may not address important questions needed to guide treatment and coverage decisions. Randomized Clinical Trials (RCTs) that are designed to provide the type of evidence desired by patients, clinicians and payers when making decisions surrounding alternative therapies have been referred to as "pragmatic or practical clinical trials (PCTs)[5]. In May 2009, CMTP convened meeting of experts and stakeholders to characterize the ways in which regulatory trials fail to provide information of potential value to post-regulatory decision makers (patients, clinicians, payers, etc.), explore the reasons for those shortcomings, and to generate ideas for designing more informative phase 3 pharmaceutical trials. A summary of this meeting is available at: http://cmtpnet.org/cmtp-research/guidance-documents/PCT%20Meeting%20Summary%20100609%20no%20watermark.pdf. While pragmatic trial designs are gaining traction in the post-regulatory environment, the vast majority of investment is in pre-market clinical research. The premise of this discussion was that it might be possible and economical to simultaneously satisfy regulatory requirements *and* produce the information needed by the critical post-regulatory decision makers. By strategically investing in more pragmatic trial designs, one could potentially avoid delays in coverage decisions and ultimately speed access to beneficial new therapies. The meeting participants included representatives from pharmaceutical companies and regulatory bodies, clinicians, private and public payers, representatives of organizations within the federal government, clinical researchers, academics, patients/consumers, and technology assessment organizations. This discussion provided the initial insights that were subsequently developed into the methodological recommendations presented in this document.

The goal of this document is to provide recommendations to guide the design and implementation of more informative ("pragmatic") phase 3/3b pharmaceutical trials. The principles outlined below include guidance in addressing methodological issues, and considerations of feasibility and ethical aspects of studies that will make them more informative to "post-regulatory" decision makers. While it will also be useful to provide guidance on the design of PCTs for specific therapeutic areas, a specific drug, or specific drug/indication combination, this document highlights general principles that are applicable across therapeutic areas. This guidance is intended to compliment, not replace guidance issued by the FDA or other agencies.

This document is intended to aid trial designers in the implementation of pragmatic design features without significantly affecting the interpretability of trial results and with limited impact on the efficiency of trial conduct. While trials designed for regulatory approval are typically highly controlled and restrictive in patients studied, there are many examples, particularly in cardiovascular disease, where large, simple trials are both pragmatic and have served as the regulatory basis for a change in drug labeling (Global Utilization of Streptokinase and Tissue Plasminogen Activator to Treat Occluded Arteries, GUSTO-I[8-9]; Heart Outcomes Prevention Evaluation, HOPE[10]; Heart Protection Study, HPS[11]). Thus implementing a pragmatic trial and providing a foundation for regulatory approval are not necessarily mutually exclusive.

However, PCTs raise unique analytical, operational and cost considerations. Some of these issues have been considered by others [4, 12-14]. Thorpe and colleagues recently delineated aspects of trial design that should be considered when designing a pragmatic trial in the Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) wheel [5]. The PRECIS framework includes ten domains determining the pragmatism of a trial. The domains included within the PRECIS framework apply to participant selection, intervention and comparator selection, outcome selection and follow-up design, participant and practitioner adherence, and analysis. The ten domains can be placed on spokes and each rated from more explanatory (at the hub) to more pragmatic (at the rim), resulting in an overview of the degree of pragmatism in that trial. We emphasize to those using this

document that the optimal approach to any given trial may require "complete pragmatism" in only selected domains or moderate pragmatism across all domains.

In fact, the domains that are most necessary for making post regulatory decisions that most often are missing from traditional regulatory trials are centered around three key points; the generalizability of the patient population, active comparators and consistently measured, relevant outcomes.

RECOMMENDATIONS

STAKEHOLDER ENGAGEMENT

I. Engage post-regulatory decision makers (patients, consumers, clinicians, payers) throughout the process of Phase 3 trial design and implementation.

Rationale. Sponsors and investigators should strive to understand the post regulatory perspectives of post-regulatory decision makers, including payers, patient advocates and physicians, in addition to regulators when designing Phase 3 trials in order to improve understanding of the design and outcome parameters that will be most informative to decision makers. It is especially important to engage clinicians, payers and patients early, as these stakeholders are central to decisions concerning whether or not a certain treatment will be offered in the healthcare market and used to manage a particular condition. There must, however, be a common understanding of whether these outcomes will be accepted by the FDA and other international equivalents. Decision makers should be encouraged to explicitly define the clinical/policy question(s) in which they are interested and request that trials be implemented to yield the best answer to that question(s) as early as possible in the drug development cycle. This may involve discussions both with patients and between manufacturers and payers.

Current Standard and Implementation. There are some existing processes. For patients, the FDA established the Cancer Drug Development Program in 2001 to incorporate the perspective of patient advocates into the drug development process by including them in the FDA drug review regulatory process. The British National Institute for Health and Clinical Excellence (NICE) has instituted a consultation service for pharmaceutical companies to inform the design of Phase 3 studies. In Australia, the Department of Health has started discussing with industry sponsors the desired evidence for decision making before new health technologies reach Phase 3 research. Pharmaceutical companies in the United States frequently engage clinician research leaders and convene advisory boards to inform the design of clinical programs and include individuals who work for large, non-governmental payers. There also have been efforts by payer organizations to be more open and transparent. For example, at the Drug Effectiveness Review Project at the Oregon Health and Science University all meetings previewing new technologies are public and industry representatives are welcome to attend. In Europe, the European Medicines Agency (EMA), the body responsible for the evaluation and supervision of medicines for human use in Europe, and the European Federation of Pharmaceutical Industries and Associations (EFPIA), the voice of the research-based pharmaceutical industry in Europe, hold a joint annual Information Day. The German Institute for Quality and Efficiency in Health Care (IQWIG) invites stakeholders who have commented on their protocols for comparative effectiveness reviews for public hearings and discussions. Additionally, in February 2010, EMA announced a collaborative agreement with the European Network for Healthcare Assessment (EUnetHTA). The collaboration is intended to improve availability and optimize data use relevant to healthcare decision makers. While these processes are a start, they may be improved by broadening the range of stakeholder groups and individuals involved in discussions of new technology evaluation. It should

also be noted that attending discussion forums does not necessarily result in involvement in the trial design and evaluation process, so it is important to ensure methods used to involve stakeholders are sufficiently interactive.

Trial designers must remember that different regulatory agencies may have different evidence standards for approval and different payers may have different standards for reimbursement. As a general rule, early engagement of post-regulatory decision makers should be done in conjunction with representatives of regulatory agencies.

STUDY POPULATIONS

II. Inclusion/exclusion criteria should be clearly articulated and justified in the trial protocol. Recruit trials participants that represent the spectrum of patients across disease severity that reflects the targeted labeled indication, including those with co-morbid conditions.

Rationale. Approved drugs are commonly used in patients that differ significantly from those included in the phase 3 trials, and in many cases, important subgroups of patients are routinely treated in the absence of reliable information on the drug's effect on these patients. It would therefore be useful for trials designers to make a serious effort to recruit trial participants that are representative of the patient populations that are likely to receive the therapy in clinical practice.

For example, in key pre-approval trials to generate evidence on the clinical efficacy and safety of Chantix, an aid to smoking cessation treatment, smokers who reported receiving treatment for depression at trial initiation or within the previous 12 months and smokers who reported a past or present history of serious mental illness were excluded from the trial. Exclusion of such patient populations is common in Phase 3 clinical trials and was approved by the FDA. Nonetheless, it is known that there is a higher prevalence of smoking in people with serious mental illness, and the pre-approval trials did not establish the safety and efficacy of Chantix in these groups. In post-marketing experience, serious neuropsychiatric symptoms, including suicidal ideation and suicide, have been reported in some patients taking Chantix, although, as noted in the Chantix label, and the trial evidence lacks information about safety (and efficacy) in this population.

The elderly, who are particularly vulnerable to increased drug exposure and drug related adverse events, tend to be under-represented in, if not excluded from, clinical trials [15]. Women comprise one of the most frequently considered "subgroups" of patients for many medical conditions [16] and the most recent NIH report indicates that women comprise approximately 55% of clinical trial participants [17]. As long as the proportion of treatments for which men and women respond very differently remains uncertain, women's inclusion in trials remains critical. Furthermore, PCTs should include a substantial proportion of patients from minorities including racial/ethnic subpopulations, as well as "vulnerable populations" and the uninsured whenever possible and clinically appropriate.

It must be noted, however, that inclusion of broad patient groups creates heterogeneity, diluting the treatment effect. The only solution is to increase sample size (and thus costs) to estimate effect within these subgroups. It is necessary that PCTs specify subgroup analysis a priori and are adequately powered to allow some subgroup analyses in order to avoid rendering an average result that does not apply to any particular group (see Recommendation 8).

Finally, after making efforts to broaden the population representation in trials, extrapolation of results from pragmatic trials to populations that were not represented in the trial should be done with great care. Although

qualitative interactions of treatment effects in subgroups are uncommon (treatment effects are not qualitatively heterogeneous), given heterogeneity of treatment effects, better known as effect modification, the observed average effects within the pragmatic trial population cannot be assumed to reflect the average treatment effect for patients not included in the trial. Effect modification should be addressed by presenting results according to the subgroups analyzed. If a drug is more effective in one sub-population, these results should be reported rather than reporting an overall study effectiveness estimate.

Current Standard and Implementation. While the FDA has no specific guidance in this area, the NIH requires the inclusion of women and minorities in clinical trials when possible and appropriate[18]. The NIH states that, in addition to the continuing inclusion of women and members of minority groups in NIH-supported biomedical and behavioral research involving human subjects, the NIH must ensure that women and members of minorities and their subpopulations are included in all human subjects research. More specifically, for Phase 3 clinical trials, researchers must ensure that women and minorities and their subpopulations are included such that valid analyses of differences in intervention effect for these groups can be accomplished. The NIH guidelines also state that cost is not an acceptable reason for excluding these groups and that researchers should initiate programs and support for outreach efforts to recruit these groups into clinical studies.

The International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) brings together the regulatory authorities of Europe, Japan and the United States and experts from the pharmaceutical industry in the three regions with the goal of standardizing regulatory requirements across these countries. At the ICH meeting, scientific and technical aspects of product registration are discussed and in 2005 EMA published guidelines about gender inclusion in clinical trials based on consensus from the ICH[19]. This report included a statement that for dose response studies, differences based on demographic characteristics should be explored and reported in subpopulation analyses.

One path to expanding representation in trials is interaction with patient advocacy groups to expand the reach of pragmatic trials into broad populations. It may also be beneficial to include clinical trial designers who have been successful at conducting and implementing such trials. The Heart Protection Study [20] and the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study are such examples of large scale, practical, clinical trials [21].

COMPARATORS

III. Select active comparators that are relevant to healthcare decisions being made by patients, providers, and payers.

Rationale. Trials incorporating only a new technology arm and a placebo arm should be avoided where there exists an unbiased endpoint (e.g., all cause mortality) and a strong, evidence based, understanding about the effectiveness of existing treatments so that an active comparator is ethically necessary, or where the placebo treatment would be ethically unjustifiable. In such cases, the following framework can be used for deliberation: a placebo-control has advantages and can be used in a double-dummy design if there is an active comparator. A placebo only comparison, when there exist proven effective treatments that will be withheld, is only justifiable when the methodological reasons for their use are compelling, a strict ethical evaluation has made it clear that patients who receive placebo will not be subject to serious harm, and provisions have been made to minimize the risks associated with the receipt of placebo[22]. In cases where a substantial portion of the population is intolerant of a class of drugs, use of a placebo may be ethically justified, such as the comparison of an

angiotension receptor blocker versus placebo in heart failure patients intolerant to ACE inhibitors [23]. Even when an active comparator is selected, there are cases where inclusion of a placebo may improve trial interpretability. Examples include the comparison of new antithrombotics to Warfarin in trials of atrial fibrillation [24]. It is particularly important to include a placebo arm, in addition to active comparators of interest, in PCTs that are intended to demonstrate non-inferiority. Recent FDA draft guidance on this topic gives advice on when they can be interpretable, on how to choose the non-inferiority margin and how to analyze the results[25].

Appropriate comparators for PCTs may also include non-drug alternatives, such as devices, surgical interventions and lifestyle and behavioral interventions. For example, the PCT conducted to evaluate the effectiveness of treatment alternatives for recent-onset low back pain compared the non-pharmaceutical alternatives physical therapy, chiropractic care and self care (outlined in an educational booklet) [26].

Current Standard and Implementation. Both the FDA and EMA increasingly are asking for active comparators, as are payers, and industry is beginning to respond with a steady increase in active-comparator studies. Direct comparison is the current standard for PCTs. Indirect comparison compares the efficacy of two different healthcare interventions, tested in separate placebo-controlled trials, in place of a head to head trial. Indirect comparisons are often utilized for clinical indications where numerous treatment options are available and head to head trials with each existing therapy are not possible. Indirect comparisons require that the separate trials being compared are sufficiently similar to be quantitatively combined, the clinical and methodological quality of the trials compared are similar and that the patient cohorts are similar. Currently, due to concerns that indirect comparisons may be subject to greater bias than direct comparisons and may overestimate the efficacy of interventions [27], indirect evidence is viewed as less credible than results from direct comparisons [28]. Indirect comparisons are used, however, as part of “non-inferiority” designs for regulatory approval. For example indirect comparisons were used to compare alternatives for Warfarin in atrial fibrillation [24]. It is also worth noting that there are groups including the Canadian Agency for Drugs and Technologies in Health (CADTH)[29] and NICE [30] that do allow indirect comparisons to be used as part of systematic reviews of interventions when no direct comparison trials are available.

When selecting active comparators, it is also important to consider that standards of care (SOC), and therefore comparators, may vary across countries and even between states in the United States. Examples include the significant differences in evidence based guidance and physician practice around use of ACE inhibitors in Europe [21] [31] and INR control with Warfarin around the world. These differences arise, in part, due to divergent trial results from the two continents, but in some areas due to insufficient evidence for an international consensus to guide daily clinical practice [32-33]. To address this variation in SOC world-wide, trial designers should consider embedding these differences within the standard of care arm (i.e. randomized new drug versus SOC, however defined); this simplifies the analysis but does increase variance within the SOC arm necessitating a larger sample size. Trial designers should compare different SOC regimens to study a drug in a pre-specified analysis using regression techniques.

When guideline directed standards of care, rather than study protocol defined care is considered the appropriate comparator, this arm should be more formally defined, especially for Phase 3 trials, than might be necessary for post-approval studies where heterogeneity of care is more acceptable.

Both the FDA and EMA provide some guidance on comparators from the regulatory standpoint. The FDA guidance addresses the selection of controls including placebo, active concurrent controls and historical controls [34]. This guidance points out both advantages and disadvantages of active comparator trials. The advantages include reducing ethical concerns surrounding non-use of drugs with documented benefits. The use of active comparators also may be more acceptable to patients and their providers, yielding larger samples with fewer

withdrawals due to lack of effectiveness. Where a trial demonstrates superiority to an active comparator, these results are more easily interpreted. Active control trials also can, if properly designed, provide information about relative efficacy. The stated disadvantages include the problem of assay sensitivity (i.e., the ability to distinguish between effective and ineffective drugs) and the ability of these trials to support efficacy statements in equivalence trials. They generally also require larger sample sizes as the expected outcome differences are always smaller for active comparator trials than for placebo controlled trials. The EMA states that, "there is no general rule applicable to all circumstances and the decision to require placebo and/or active control studies has been and will be taken on a case-by case basis although some fundamental principles apply." These are based on the availability and known efficacy of existing therapeutic alternatives to the study drug [35].

TRIAL OUTCOMES

IV. Select outcomes relevant to the healthcare decisions being made by patients, providers, and payers.

Rationale. The issue of outcomes and endpoint selection occurs in all clinical trials. For PCTs, it is critical that the outcomes are clinically important to stakeholders, the most important stakeholder being the patient. While easily measured markers may be selected, they are not always appropriate alternates to clinical endpoints. For example, Fleming and DeMets report on the poor prognostic utility of CD4 count when used as a surrogate for long term outcomes in clinical trials of HIV therapy [36]. Ideally, PCTs are intended to provide both benefit and risk information. Therefore, they would include clinical endpoints that demonstrate benefit to patients in addition to surrogates when they are used.

Current Standard and Implementation. Historically, acceptance of surrogate endpoints requires that the surrogate measure has been repeatedly validated as a highly reliable indicator of clinically meaningful outcomes [37-38]. Fleming[39] summarizes the issues surrounding validation of surrogate endpoints in clinical trials, that is ensuring that these surrogates actually predict patient benefit. Others, including Janet Woodcock, the director of the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration (FDA), have suggested that the definitions and criteria for validation are out of date and that there should be exploration of alternative conceptual frameworks in this area. One recent review covers examples where the use of surrogates was more and less appropriate for drug development trials [40]. Some of the successes include tumor imaging results as a surrogate for progression free survival in cancer and change in LDL as a surrogate for cardiac events, although even that has been challenged by a trial showing increased mortality with an agent highly effective at lowering LDL, likely related to off-target effect [41]. The history of surrogates includes premature ventricular contractions as a surrogate endpoint for sudden cardiac death, which led to use of drugs that caused mortality, as shown in the CAST trial. Lathia et al. point out that, while there is no evidence of causality, there is a strong relationship between the availability of surrogate endpoints and successful drug development. Given this, Lathia et al. are in favor of expanding the use of biomarkers within an evidentiary framework, such as that developed by Altar and colleagues [20].

In April 2009, the FDA reiterated its position on accelerated approval, dating from 1992[42]. It stated that it may grant marketing approval for a new drug product on the basis of adequate and well-controlled clinical trials which establish that the drug product either has an effect on a surrogate endpoint that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit, or on the basis of an effect on a clinical endpoint other than survival or irreversible morbidity [43]. Where appropriate these outcomes can be incorporated alongside outcomes measures that are of greatest interest to decision makers.

V. Include patient-reported outcome measures that are meaningful to patients and important for physicians.

Rationale. Because patient reported outcomes (PROs) such as pain or health related quality of life (HRQOL) often are key factors in treatment selection by patients or physicians, inclusion of PROs in PCTs enhances the value of PCT results for these stakeholders. Roland and Torgerson [44] point out the potential desirability of multiple outcomes as well as the potential for conflicting analytic conclusions resulting from trials with multiple endpoints. They state that in PCTs, a single outcome may not provide adequate information to decision makers, strengthening the desirability of PROs alongside clinical efficacy endpoints in PCTs. In the case of cancer, a great deal of work was done by the National Cancer Institute Cancer Outcomes Medical Working Group (NCI COMWG), which concluded, "The circumstances under which HRQOL measures bring significant information value to outcomes assessment over and above that provided by traditional biomedical endpoints need to be identified, particularly when the study's primary endpoint is survival or disease-free survival [45]. For purposes of COMWG deliberations, HRQOL measures were defined as providing added value when these measures were instrumental in interpreting a study's findings and would be expected to influence clinical recommendations."

Current Standard and Implementation. There needs to be a common understanding of which patient-reported outcomes measures are appropriate in Phase 3 clinical trials and whether these outcomes will be accepted by the FDA, payers and the clinical community. The FDA has issued guidance and the EMA published a reflection paper on the use of patient-reported outcomes in clinical trials. The December, 2009 guidance from the FDA [46], "describes how the Food and Drug Administration (FDA) reviews and evaluates existing, modified, or newly created patient-reported outcome (PRO) instruments used to support claims in approved medical product labeling." The central point of the guidance is that instruments used to measure PROs must have documented evidence of both patient participation in instrument development and of the instrument's reliability in measuring the endpoint or concept of interest for the population and clinical condition of interest for the current clinical trial. The EMA reflection paper makes similar statements, but also contains additional remarks about duration, condition and trial design when assessing HRQOL in clinical trials[47].

TRIAL PROTOCOL

VI. Allow deviation from a strict protocol only insofar as it does not harm the ability to detect the relative efficacy of the experimental drug.

Rationale. While the goal of PCTs is often assessing effectiveness in actual clinical settings, once use is not directed by protocol, allowing unlimited variation in care may lead to difficulties in reaching significance of outcomes (due to high variance) and in understanding what drives clinical effectiveness. Minimizing switching treatment arms and maximizing treatment adherence are generally considered necessary controls to generate valid and reliable data in Phase 3 trials. While understanding how the characteristics of a drug affect real world use is of interest to decision makers, some of these questions are better addressed after efficacy has been clearly established. Phase 3b or Phase 4 trials are more appropriate vehicles for examining variation in prescriber or patient adherence. Eli Lilly, among other companies, has been an early adopter of PCTs, having supported three that have been completed since 1996 (for Olanzapine, Risperidone, and typical anti-psychotics [1], Paroxetine, Fluoxetine and Sertraline [48] and Desipramine, Fluoxetine and Limipramine [49]). They attempted to mirror "real world" treatment including switching, discontinuation, restarting and "add on" therapy. Their experience was that allowing these care patterns made it difficult to determine the effectiveness of a given therapy.

To be most informative to future clinical decision-making, trials need to strike a balance between encouraging high-quality standard care, while allowing the variation in clinical practice that could occur across multiple providers all of whom are treating patients within evidence-based guidelines. If it were possible to predict the background treatments that would be used in the populations being treated with a proposed therapy, replicating that would be ideal to understand the risks and benefits in that context.

Current Standard and Implementation. "Clinically-indicated care" is care that would require investigators to adhere to protocol-prescribed care [50] where the protocol is based on a formal guideline appropriate to the clinical setting in which care is delivered. This may allow incorporation of practice variability without moving away from accepted care practices into "inappropriate care". There are also situations in which tests or further investigations may be aimed specifically at protecting patient safety such as the international study into the efficacy and safety of rivastigmine in patients with Alzheimer's disease[51], the international study of the safety of meloxicam [52] and the study of the efficacy and safety of tifacogin [53]. Particularly for a first-in-class compound, there may be adverse events that are outside the usual clinical experience. The relationship between the clinical trial setting and standard care should be considered when making decisions about both the safety and efficacy of new therapies.

VII. Justify the components and intensity of trial monitoring, aiming for parsimony.

Rationale. Reducing protocol driven care does not imply abandoning trial monitoring. Rather, appropriate alternatives to intensive universal monitoring should be used. When designing a PCT, examine the potential for adaptive or knowledge-based monitoring rather than universal monitoring. Such approaches do not involve monitoring at pre-specified time intervals; rather they begin monitoring according to the occurrence of pre-specified events. For example, monitoring might occur once a certain number of patients have been enrolled, in the case of an adverse event, or at the time of a protocol deviation. However, trial designers should be aware that if monitoring is done in a manner that is not protocol driven and universal, it is possible that only patients with "events" get monitored leading to a database with asymmetric data collection. This would be difficult to analyze and interpret. The protocol driven approach ensures uniform data collection and minimizes bias. Trial designers must make a careful tradeoff analysis before selecting not to use protocol driven monitoring. Eisenstein and colleagues described the significant impact of modest changes in monitoring in Phase 3 cardiology trials [54].

Current Standard and Implementation. When designing external trial monitoring, to approximate standard care environments, as well as to redirect resources into maximizing patient enrollment and narrowing the estimate of treatment effect, appropriate alternatives to intensive universal trials monitoring should be considered. When designing a PCT, examine the potential for adaptive or knowledge-based monitoring rather than universal monitoring. These approaches use an event-determined approach to monitoring rather than monitoring at pre-specified time intervals. For example, monitoring might occur once a certain number of patients have been enrolled, in the case of an adverse event, or at the time of a protocol deviation. This is one area being approached by the Clinical Trials Transformation Initiative, with the FDA's recent call for grants including "Assess existing approaches to monitoring (e.g. onsite monitoring) and, to the extent practicable, propose standards for approaches to monitoring that take into account the specifics of study design, study population and clinical setting."

Going forward, consider the use of electronic medical records (EMRs) as they mature. The information captured in electronic records may need to be expanded, at least in the context of a specific trial, to allow tailoring of

information to the particular question being studied [55]. Contract Research Organizations (CROs) broadly use electronic data capture, but this may not include patient and practice characteristics outside the trial protocol.

PATIENT and PRACTICE HETEROGENEITY

VIII. Pre-specify subpopulations of interest

Rationale. Pragmatic trial designs introduce multiple sources of heterogeneity, including heterogeneity of the population, intervention dose, clinicians and providers. The ideal way to address heterogeneity is though pre-specification or randomization of by subgroup during the design of a trial. Byar et al. recommend relaxing eligibility criteria and analyzing results in sub-groups of patients stratified by more restrictive criteria [56]. They state that the heterogeneity of the larger group allows for better overall clinical generalization. Important subpopulations can be considered even when the study is not powered for hypothesis testing. This will determine if the point estimates for efficacy and safety issues appear to differ for these populations and may be beneficial for informing future study design.

The FDA has observed a movement toward displaying information on a large number of subgroups in forest plots. In some instances, such as the case of the trial of losartan versus atenolol, which showed a reduction in stroke risk for whites, but the reverse for blacks, the retrospective subgroup analysis could not be ignored and was included in the label [57]. Pre-specification may help to limit "data dredging" as it involves hypothesis testing in sub-populations. However, pre-specified subpopulations should be limited to a "reasonable" number based on hypotheses about subgroups and should be selected based on scientific rationale. Pre-specification of subgroups can include clinical, demographic, or other characteristics that would be available and observable to clinicians at the point of care delivery (e.g. disease severity, response to diagnostic genetic test, gender/race/ethnicity or age). Operational definitions of specific subgroups should be based on characteristics unrelated to group assignment or trial outcomes.

Current Standard and Implementation. The EMA provides some guidance in this area in their Statistical Principles for Clinical Trials ICH document which states, "In some instances an adjustment for the influence of covariates or for subgroup effects is an integral part of the planned analysis and hence should be set out in the protocol. Pre-trial deliberations should identify those covariates and factors expected to have an important influence on the primary variable(s), and should consider how to account for these in the analysis in order to improve precision and to compensate for any lack of balance between treatment groups. If one or more factors are used to stratify the design, it is appropriate to account for those factors in the analysis"[58]. FDA regulations require sponsors of New Drug Applications (NDA) to present a summary of safety and effectiveness data by demographic subgroups (age, gender, race), as well as an analysis of whether modifications of dose or dosage intervals are needed for specific subgroups (21 CFR 314.50 (d)(5)(v) and (vi)(a))[43].

Post hoc analyses of sub-populations should be limited to those for which there is a scientific rationale for why the sub-population might reasonably differ from other subpopulations. In addition, justification for why such post hoc analyses were not pre-specified should be provided (e.g. new scientific findings).

Clinical and demographic characteristics, including disease severity and prior treatment, should be captured and reported as part of the heterogeneity of the patient population. Outcomes reporting for all sizable sub-populations in the analysis should include both the rationale for their inclusion and the average treatment effect for that subpopulation.

When existing evidence suggests that there is a scientific reason why the treatment effect may differ in a sizeable population, then that subpopulation should be further explored [59]. Statistical and regulatory aspects of subgroup analyses have been reviewed by Grouin and colleagues [2, 60].

While the desire to observe effectiveness of treatments in "real world" populations necessitates expanding research into a variety of health care settings, studying the heterogeneity of prescriber practice patterns and patient adherence may be better reserved for later phases of drug development.

OPERATIONAL ISSUES and INFRASTRUCTURE

Major barriers to pragmatic clinical trials include cost, regulations, and operational inefficiencies [61]. For a discipline based on principles of evidence generation, a great irony in the conduct of clinical trials is the lack of evidence that the enormously costly and complex approach to assuring regulatory compliance, data quality with monitoring, safety with averse event reporting, and privacy protection, achieves its goal [62]. Central to the objective of PCTs is to inform the medical community about the effects of various treatments on important clinical outcomes. To be successful, trials must be adequately sized and powered, and of long enough duration to accomplish that goal. There is a major need to improve the capacity to conduct large, simple trials in standard practice. A major opportunity, therefore, is to simplify and streamline the conduct of large trials.

IX. Use existing infrastructure, such as practice-based research networks, to broaden the included populations while maintaining data integrity and capture.

Rationale. Practice-based research networks, such as the one used in the pragmatic Dutch study of Initial Management of Newly diagnosed Dyspepsia (DIAMOND) [63] should be considered as one approach to provide an infrastructure to support the rapid design and implementation of pragmatic studies, although supplementary funding may be required to facilitate competitive trials.

Current Standard and Implementation. In addition, major efforts are underway to reduce regulatory, contracting, and financial barriers to conducting pragmatic trials. One of these is the Clinical Trials Transformation Initiative (CTTI) at Duke University. The goal of this new public-private partnership is to improve the quality and efficiency of clinical trials. As founding partners, the FDA and Duke enlisted diverse stakeholders in the clinical trials enterprise to participate in the initiative.

X. Incorporate novel approaches to study design and data analysis, such as adaptive clinical designs, to enhance efficiencies and retain unbiased assessments of comparative effectiveness in PCTs

Rationale. Trials with adaptive features which allow for alterations in design or analyses based on examination of data at an interim point in the trial may make clinical trials more efficient. Adaptive clinical trials run the gamut from "staged" protocols or "group sequential" trials to Bayesian designs that allow both continuous monitoring of trial data and the ability to change pre-specified trial features during trial progress, including the number of patients, eligibility criteria, group assignment, and dosing protocols[64]. An excellent example of this type of approach was Pfizer's ASTIN trial (Acute Stroke Therapy by Inhibition of Neutrophils), a Phase 2 trial which included adaptive allocation of patients to different dose groups and the possibility of a seamless transition to Phase 3 in the event of a strong efficacy outcome.

Current Standard Implementation. There are a number of practice based research networks across specialty areas, including the Federation of Practice based Networks in primary care[65], large cooperative trials networks in oncology such as the Southwest Oncology Group (SWOG), and the Health Maintenance Organization (HMO) Research Network. Payers have expressed interest in models including Practice Based Research Networks (PBRNs) that will allow the integration of research and patient care to make research in general clinical settings financially feasible [66-67]. Many believe that PBRNs can help provide research that can be more rapidly and directly applied to decision making in actual practice[68]. The need for this type of infrastructure has been stated in various clinical areas, including pediatrics by the EMA[69], psychiatry[70], and device related infections [71].

There are several appropriate trial designs that might be considered. In February of 2010, the FDA released a draft guidance for pharmaceutical and biologics developers on adaptive design clinical trials[72]. It describes an adaptive trial as, "a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study." The 50-page guidance covers the clinical, statistical and regulatory aspects of adaptive design clinical studies, and focuses on adaptive designs that may not be familiar to clinical trial designers. Many of these concepts have been covered in publications describing the role of Bayesian analysis and adaptive clinical designs in reducing the sample size, time, and cost required to obtain decision-relevant information from clinical trials [66]. EMA does not provide guidance in this area, but has produced a reflection paper[73].

While the main goal of Phase 2 trials is proof of principle and dose determination, Phase 2 trials should be used to inform the design and improve the implementation of pragmatic Phase 3 trials. Broadening the patient population and addressing specifically questions and outcomes of interest to decision makers should increase the efficiency of Phase 3 trials and potentially reduce the number of post-marketing trials that need to be conducted.

Ultimately trial designers must be strategic when deciding whether and how to implement and utilize PCTs. Not all features of pragmatism are of equal importance to payers and decision makers. By strategically focusing on those domains that are most important to end users, improving the pragmatic design of phase 3 trials may not only be desirable, but may also be feasible.

THE ETHICAL DILEMMA OF PATIENT SAFETY IN PRAGMATIC CLINICAL TRIALS

Pragmatic Clinical Trials give rise to unique ethical and safety issues that must be considered by trial designers.

Following the principles of pragmatism, including relaxing protocol design, decreasing monitoring and broadening the included population introduces ethical challenges. When you include older, sicker patients in a trial, you may increase the number of adverse events. However, it is preferable to discover these risks in the context of a clinical trial where fewer total patients are at risk and there is better monitoring and control of these patients than in general clinical practice.

The Chantix example (Principle 2) illustrates the importance of appropriate selection of patient populations. PCTs typically have broad inclusion criteria and minimal exclusion criteria to enable the study of a more diverse patient population. The goal of PCTs is to enroll patients in the trial with characteristics that reflect the range and distribution of patients observed in clinical practice for a particular problem [1]. This approach prevents the exclusion of higher-risk patients who may receive the greatest benefit from the treatment under investigation [2]. However, trial designers must be aware that they may potentially record a higher number of co-morbidities due to the inclusion of "sicker" patients with other medical issues that may preclude treatment effect. It also satisfies the requirement of decision makers to have more information about applicability of results[3] [4]. Trial designers should note that more extensive use of clinical and administrative records to identify suitable potential study participants may raise a unique set of informed consent and confidentiality issues [1].

Relaxation of the protocol design and reduction of how closely patients are monitored during the trial are important goals in the design of PCTs. Such measures more closely replicate "real world" treatment practice and also maximize the inclusion of treatment centers in the study that are not primarily focused on research. In order for the risk – benefit profile of the preclinical studies / exploratory analyses to be accurately evaluated, trial designers should be aware of the accuracy of signals of excess risk or specific off-target effects. Such results should be addressed in confirmatory trials [6] before excluding a treatment from being tested with a Pragmatic Clinical Trial.

Trial designers should also be aware that, in some cases, a more heterogeneous population may lead to more "false signals", especially in the safety realm. The implications of patient heterogeneity for safety results must be understood and appropriate data analysis techniques and reporting must be employed to prevent "false signals" resulting from spurious heterogeneity of treatment effects. Such errors could lead to inaccurate assessments of the risk – benefit profile of the investigational intervention or termination of the trial. However, if there is indeed a subgroup of patients with high co-morbidity who are at increased risk from the therapy being tested then this will only be established by testing the therapy in that population.

PCTs typically test the treatment under investigation in a head to head trial against another treatment in order to give the best information to decision makers making treatment choices. This allows PCTs to go beyond proving efficacy and giving decision makers information about effectiveness. Trial designers should be aware of the ethical issues surrounding economic motivations for such trials and the question of whether some patients enrolled in these trials are being denied optimal care[1]. Trial designers must determine whether true clinical equipoise exists[7].

REFERENCES

1. Tunis, S.L., et al., *Cost-effectiveness of olanzapine as first-line treatment for schizophrenia: results from a randomized, open-label, 1-year trial.* Value Health, 2006. **9**(2): p. 77-89.
2. Califf, R.M. and D.L. DeMets, *Principles from clinical trials relevant to clinical practice: Part I.* Circulation, 2002. **106**(8): p. 1015-21.
3. Guyatt, G., *Academic medicine and the pharmaceutical industry: a cautionary tale.* CMAJ, 1994. **150**(6): p. 951-3.
4. Dans, A.L., et al., *Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group.* JAMA, 1998. **279**(7): p. 545-9.
5. Thorpe, K.E., et al., *A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers.* J Clin Epidemiol, 2009. **62**(5): p. 464-75.
6. Flemming, D.C., *Value in the placebo response.* BMJ, 2008. **337**: p. a2685.
7. Liliford, R., *Ethics of clinical trials from a Bayesian and decision analytic perspective: whose equipoise is it anyway?* BMJ, 2003. **326**: p. 980-981.
8. Ridker, P.M., et al., *Large-scale trials of thrombolytic therapy for acute myocardial infarction: GISSI-2, ISIS-3, and GUSTO-1.* Ann Intern Med, 1993. **119**(6): p. 530-2.
9. *GUSTO trial results. American Federation of Clinical Research.* Clin. Res., 1993. **47**: p. 207-8.
10. Document, F.B., *The HOPE (Heart Outcomes Prevention Evaluation) Study: A large, randomized study of the ACE inhibitor, ramipril and vitamin E in patients at high risk of cardiovascular disease.* . 1999.
11. Collins, R., et al., *MRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial.* Lancet, 2003. **361**(9374): p. 2005-16.
12. Macpherson, H., *Pragmatic clinical trials.* Complement Ther Med, 2004. **12**(2-3): p. 136-40.
13. Tunis, S.R., D.B. Stryer, and C.M. Clancy, *Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy.* JAMA, 2003. **290**(12): p. 1624-32.
14. Brass, E.P., *The gap between clinical trials and clinical practice: the use of pragmatic clinical trials to inform regulatory decision making.* Clin Pharmacol Ther, 2010. **87**(3): p. 351-5.
15. Lee, P.Y., et al., *Representation of elderly persons and women in published randomized trials of acute coronary syndromes.* JAMA, 2001. **286**(6): p. 708-13.
16. Berlin, J.A. and S.S. Ellenberg, *Inclusion of women in clinical trials.* BMC Med, 2009. **7**: p. 56.
17. NIH, *Comprehensive Report: Tracking of Human Subjects Research as Reported in Fiscal Year 2006 and Fiscal Year 2007.*
18. *NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research – Amended, October, 2001*. 5/1/10]; Available from: http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm.
19. *EMEA ICH Step 5. GENDER CONSIDERATIONS IN THE CONDUCT OF CLINICAL TRIALS*. Available from: http://www.ema.europa.eu/pdfs/human/ich/391605en.pdf.
20. Collins, R., et al., *Effects of cholesterol-lowering with simvastatin on stroke and other major vascular events in 20536 people with cerebrovascular disease or other high-risk conditions.* Lancet, 2004. **363**(9411): p. 757-67.
21. Gaynes, B.N., et al., *What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression.* Psychiatr Serv, 2009. **60**(11): p. 1439-45.
22. Emanuel, E.J. and F.G. Miller, *The ethics of placebo-controlled trials--a middle ground.* N Engl J Med, 2001. **345**(12): p. 915-9.
23. Granger, C.B., et al., *Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function intolerant to angiotensin-converting-enzyme inhibitors: the CHARM-Alternative trial.* Lancet, 2003. **362**(9386): p. 772-6.

24. Jackson, K., et al., *Antithrombotic drug development for atrial fibrillation: proceedings, Washington, DC, July 25-27, 2005.* Am Heart J, 2008. **155**(5): p. 829-40.
25. FDA, *Non-Inferiority Clinical Trials.* Draft Guidance for Industry, March 2010.
26. Cherkin, D.C., et al., *A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain.* N Engl J Med, 1998. **339**(15): p. 1021-9.
27. Bucher, H.C., et al., *The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials.* J Clin Epidemiol, 1997. **50**(6): p. 683-91.
28. Glenny, A.M., et al., *Indirect comparisons of competing interventions.* Health Technol Assess, 2005. **9**(26): p. 1-134, iii-iv.
29. Wells GA, S.S., Chen L, Khan M, Coyle D., *Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis.* Ottawa: Canadian Agency for Drugs and Technologies in Health, 2009.
30. Sutton A, A.T., Abrams K and Cooper N., *Briefing paper for methods review workshop on evidence synthesis (indirect and mixed treatment comparisons).* NICE Decision Support Unit, May 2007.
31. Remme, W.J., et al., *Awareness and perception of heart failure among European cardiologists, internists, geriatricians, and primary care physicians.* Eur Heart J, 2008. **29**(14): p. 1739-52.
32. Connolly, S.J., et al., *Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range.* Circulation, 2008. **118**(20): p. 2029-37.
33. Masuhr, F., M. Busch, and K.M. Einhaupl, *Differences in medical and surgical therapy for stroke prevention between leading experts in North America and Western Europe.* Stroke, 1998. **29**(2): p. 339-45.
34. *FDA Guidance for Industry. E 10 Choice of Control Group and Related Issues in Clinical Trials.* 3/16/10]; Available from: http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm125912.pdf.
35. *EU Standard of Medicinal Product Registration: Clinical Evaluation of Risk/Benefit -The role of Comparator Studies*. 3/16/10]; Available from: http://www.ema.europa.eu/pdfs/human/press/pos/11931904en.pdf.
36. Fleming, T.R. and D.L. DeMets, *Surrogate end points in clinical trials: are we being misled?* Ann Intern Med, 1996. **125**(7): p. 605-13.
37. Echt, D.S., et al., *Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial.* N Engl J Med, 1991. **324**(12): p. 781-8.
38. Deyo, R.A., *Cascade effects of medical technology.* Annu Rev Public Health, 2002. **23**: p. 23-44.
39. Fleming, T.R., *Surrogate endpoints and FDA's accelerated approval process.* Health Aff (Millwood), 2005. **24**(1): p. 67-78.
40. Lathia, C.D., et al., *The value, qualification, and regulatory use of surrogate end points in drug development.* Clin Pharmacol Ther, 2009. **86**(1): p. 32-43.
41. Barter, P.J., et al., *Effects of torcetrapib in patients at high risk for coronary events.* N Engl J Med, 2007. **357**(21): p. 2109-22.
42. FDA, *Code of Federal Regulation, Title 21, Secs. 314.500–314.560.*
43. Kellerman, S.E. and J. Herold, *Physician response to surveys. A review of the literature.* Am J Prev Med, 2001. **20**(1): p. 61-7.
44. Roland, M. and D. Torgerson, *Understanding controlled trials: what outcomes should be measured?* BMJ, 1998. **317**(7165): p. 1075.
45. Lipscomb, J., C.C. Gotay, and C. Snyder, *Outcomes assessment in cancer*. 2005, Cambridge: Cambridge University Press. xiv, 662 p.
46. *FDA Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 3/16/10]; Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.

47. EMEA, C.F.M.P.F.H.U.C. *REFLECTION PAPER ON THE REGULATORY GUIDANCE FOR THE USE OF HEALTHRELATED QUALITY OF LIFE (HRQL) MEASURES IN THE EVALUATION OF MEDICINAL PRODUCTS*. 3/16/10]; Available from: www.emea.europa.eu/pdfs/human/ewp/13939104en.pdf.
48. Kroenke, K., et al., *Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial.* JAMA, 2001. **286**(23): p. 2947-55.
49. Simon, G.E., et al., *Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs tricyclic antidepressants.* JAMA, 1996. **275**(24): p. 1897-902.
50. Meade, M.O. and E.W. Ely, *Protocols to improve the care of critically ill pediatric and adult patients.* JAMA, 2002. **288**(20): p. 2601-3.
51. Rosler, M., et al., *Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial.* BMJ, 1999. **318**(7184): p. 633-8.
52. Distel, M., et al., *Safety of meloxicam: a global analysis of clinical trials.* Br J Rheumatol, 1996. **35 Suppl 1**: p. 68-77.
53. Abraham, E., et al., *Efficacy and safety of tifacogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial.* JAMA, 2003. **290**(2): p. 238-47.
54. Eisenstein, E.L., et al., *Reducing the costs of phase III cardiovascular clinical trials.* Am Heart J, 2005. **149**(3): p. 482-8.
55. Hirst G., D.C., Hovy E., Parsons K., *Authoring and Generating Health-Education Documents That Are Tailored to the Needs of the Individual Patient*, in *User Modeling: Proceedings of the Sixth International Conference, UM97*, P.C. Jameson A., Tasso C., Editor. 1997.
56. Byar, D.P., et al., *Design considerations for AIDS trials.* N Engl J Med, 1990. **323**(19): p. 1343-8.
57. Okin, P.M., et al., *Regression of electrocardiographic left ventricular hypertrophy by losartan versus atenolol: The Losartan Intervention for Endpoint reduction in Hypertension (LIFE) Study.* Circulation, 2003. **108**(6): p. 684-90.
58. (EMA), E.M.A., *ICH Topic E 9: Statistical Principles for Clinical Trials*. 1998.
59. Peto, R. and C. Baigent, *Trials: the next 50 years. Large scale randomised evidence of moderate benefits.* BMJ, 1998. **317**(7167): p. 1170-1.
60. Grouin, J.M., M. Coste, and J. Lewis, *Subgroup analyses in randomized clinical trials: statistical and regulatory issues.* J Biopharm Stat, 2005. **15**(5): p. 869-82.
61. Duley, L., et al., *Specific barriers to the conduct of randomized trials.* Clin Trials, 2008. **5**(1): p. 40-8.
62. Yusuf, S., et al., *Sensible guidelines for the conduct of large randomized trials.* Clin Trials, 2008. **5**(1): p. 38-9.
63. Fransen, G.A., et al., *Pragmatic trials in primary care. Methodological challenges and solutions demonstrated by the DIAMOND-study.* BMC Med Res Methodol, 2007. **7**: p. 16.
64. Chow, S.C. and M. Chang, *Adaptive design methods in clinical trials - a review.* Orphanet J Rare Dis, 2008. **3**: p. 11.
65. Blackburn, G.G., et al., *Cardiac rehabilitation participation patterns in a large, tertiary care center: evidence for selection bias.* J Cardiopulm Rehabil, 2000. **20**(3): p. 189-95.
66. Luce, B.R., et al., *Rethinking randomized clinical trials for comparative effectiveness research: the need for transformational change.* Ann Intern Med, 2009. **151**(3): p. 206-9.
67. Rawlins, M.D., *Cutting the cost of drug development?* Nat Rev Drug Discov, 2004. **3**(4): p. 360-4.
68. Mold, J.W. and K.A. Peterson, *Primary care practice-based research networks: working at the interface between research and quality improvement.* Ann Fam Med, 2005. **3 Suppl 1**: p. S12-20.
69. *The Network of Paediatric Networks at the EMEA Implementing Strategy.*
70. March, J.S., et al., *The case for practical clinical trials in psychiatry.* Am J Psychiatry, 2005. **162**(5): p. 836-46.
71. Califf, R.M., et al., *Novel approaches to clinical trials: device-related infections.* Am Heart J, 2004. **147**(4): p. 599-604.

72. *FDA Guidance for Industry. Adaptive Design Clinical Trials for Drugs and Biologics* 3/16/10]; Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf.
73. EMEA. *REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS WITH FLEXIBLE DESIGN AND ANALYSIS PLAN*. Available from: http://www.ema.europa.eu/pdfs/human/ewp/245902en.pdf.

Appendix A: Stakeholder Committee Members

Academia

Gerald Gartlehner
University of North Carolina

Steve Goodman
Johns Hopkins University

Chris Granger
Duke University

Daniel Mullins
University of Maryland/Center for Medical Technology Policy

Richard Platt
Harvard University, AHRQ's Centers for Education and Research on Therapeutics

Jodi Segal
Johns Hopkins University

Merrick Zwarenstein
University of Toronto

Government Clinical Research

David Gordon
National Institutes of Health

Michael Lauer
National Health, Lung and Blood Institute

Government Clinical Research (Cont.)

Steve Phurrough
Agency for Healthcare Research and Quality

Jean Slutsky
Agency for Healthcare Research and Quality

Patient Representative

John Santa
Consumer Reports Health Ratings Center

Private Payer

Kim Caldwell
Humana

Mark Grant
Blue Cross Blue Shield Association

Peter Juhn
Medco Health Solutions

Ed Lennard
Federal Employee Program, Blue Cross Blue Shield Association

Ed Pezalla
Aetna Pharmacy Management

Private Researcher
Leslye Fitterman
Center for Medical Technology Policy

Bryan Luce
Pragmatic Approaches to Comparative Effectiveness Initiative, United Biosource

Penny Mohr
Center for Medical Technology Policy

Russ Montgomery
Center for Medical Technology Policy

Seema Sonad
Center for Medical Technology Policy

Sean Tunis
Center for Medical Technology Policy

Danielle Whicher
Center for Medical Technology Policy

Public Payer
Chester Good
Department of Veterans Affairs

Louis Jacques
Centers for Medicare and Medicaid Services

David Pass
Health Resources Commission, State of Oregon

Adriana Platona
Department of Health, Australia

Public/Private Partnership
Robert Califf
Duke University

Judy Kramer
Clinical Trials Transformation Initiative, FDA

Regulatory Body
Robert Temple
FDA

Janet Woodcock
FDA

Sponsoring Organization
Carlos Alatorre
Lilly

Roy Baynes
Amgen

Marc Berger
Lilly

Jesse Berlin
Johnson & Johnson

Stephen Bocuzzi
Sanofi-Aventis

Joanne Chang
Novartis

Tom Haverty
Schering-Plough

John O'Donnell
Bristol Meyers-Squibb

Josh Ofman
Amgen

John Orloff
Novartis

Eleanor Perfetto
Pfizer

Sponsoring Organization (Cont.)

Mary Helen Tran
Schering-Plough

Richard Willke
Pfizer

Richard Wolgemuth
Bristol Meyers-Squibb

Technology Assessment Organization

Mark Gibson
Centers for Evidence-Based Policy, OHSU

Mark Helfand
Oregon Health and Science University

Don Husereau
Canadian Agency for Drugs and Technologies in Health

Seren Phillips
National Institute for Health and Clinical Excellence