

**The Prospect of Early Pharmaceutical Pragmatic Clinical Trials:  
A Background Paper**

**Jodi B. Segal, MD, MPH  
Johns Hopkins University School of Medicine**

## **I. Background and Objectives**

A goal of comparative effectiveness research is to provide health care decision makers, particularly patients, clinicians and payers; with the best evidence to make informed health care decisions. Because the pharmaceutical industry invests substantial time and financial resources into *pre-market* clinical research, it might be advisable and efficient for industry to implement pre-market pharmaceutical studies that meet the needs of multiple decision makers, i.e. to simultaneously satisfy the regulatory requirements and produce the information desired by the critical post-regulatory decision makers: patients, clinicians and payers.

Pharmaceutical Pragmatic Clinical Trials (PCTs) are prospective studies designed specifically with the objective of informing patients, clinicians and payers, when making decisions about drug therapies. These studies aim to generate evidence that is applicable to a broad range of patients in usual care settings and conditions. Interest in PCTs has increased with greater attention from payers and policy makers to the need for more reliable evidence of comparative effectiveness in typical practice settings among diverse patient populations.

Pharmaceutical and biotech companies have recognized that achieving optimal market access is not assured by focusing only on meeting regulatory requirements. It is increasingly necessary to design pre-market trials that provide the evidence desired by post-regulatory decision makers. While the demand for such evidence is recognized, there are few examples of pragmatic trials being done for drug licensing. There is also not yet a well-defined framework for the design and implementation of pre-market pharmaceutical pragmatic clinical trials.

In preparation for the development of a guidance document that would lay out a shared set of principles for incorporating more pragmatic features into trials designed for regulatory approval, we have prepared this background paper. In this document, we aim to:

1. Provide historical context for a discussion about PCTs and their place in drug approval;
2. Clarify definitions used in describing PCTs;
3. Identify challenges in designing and implementing these studies for the evaluation of drugs in the licensing phase; and
4. Set a framework for further discussion and methodology development.

It is beyond the scope of this paper to review, in detail, trial designs that may be consistent with the goals of PCTs, as there are numerous options for designs. Similarly, we do not include detailed discussion of the statistical methods for analyzing PCTs. This paper focuses on the “why” of these trials.

## **II. Historical Background and Context**

### *A. Drug Approval in the United States*

We highlight here several key aspects of the drug approval process in the United States that are relevant to our discussion of pragmatic trials.

## Landmark Events in Drug Law in the United States

To provide context, we review just a few key landmark events in the rich history of drug regulation in the United States (U.S.).(1) The modern era of drug safety began with the passage by Congress of the Federal Food, Drug, and Cosmetic Act of 1938. This act required that new drugs be shown to be safe before marketing, and ushered in a new system of drug regulation. This Act was first put to the test with the Food and Drug Administration's (FDA) evaluation of the safety of insulin in 1941, followed soon by safety evaluations of penicillin. The next major milestone came 20 years later with the passage of the Kefauver-Harris Drug Amendments in 1962. These amendments were passed to require demonstration of a drug's *efficacy* and to further strengthen drug safety. Prior to this time, only the demonstration of safety was required before marketing. The passage of these amendments followed closely upon the thalidomide tragedy in Europe.

Legislation regarding human subject's protection was passed in 1981, and the Orphan Drug Act was passed in 1983, enabling the FDA to promote research and marketing of drugs for treating rare diseases. The Food and Drug Administration Act of 1988 officially established the FDA as an agency of the Department of Health and Human Services. The Prescription Drug User Fee Act (PDUFA) was first passed in 1992. This required drug and biologics manufacturers to pay fees for product applications and supplements, and other services. This allowed the hiring of more reviewers to expedite the drug review process.

The Food and Drug Administration Modernization Act of 1997 reauthorized the PDUFA of 1992 and mandated the most wide-ranging reforms in agency practices since 1938. Provisions included measures to accelerate review of devices, to regulate advertising of unapproved uses of approved drugs and devices, and to regulate health claims for foods. PDUFA was reapproved in 2002, and again in 2007 as part of Food and Drug Administration Amendment Acts (FDAAA) of 2007.

The FDAAA, passed in 2007, was legislation with more than 200 provisions designed to inform the public about drug safety and provided new tools for the FDA to reduce risks and unsafe drug use.(2) Among these provisions, the FDAAA established a detailed procedure for the FDA to request certain revisions to approved labeling and for manufacturers to respond to these requests. The FDAAA also granted the FDA the authority to *require* that post-approval studies be done by the manufacturer. If marketed medications are found to be associated with new potential risks, FDA can now require labeling changes or additional research to address these risks.

## Title 21 of the Code of Federal Regulations

The regulatory activities of the FDA are legislated and described in Title 21 of the Code of Federal Regulations. Part 314, *Applications for FDA Approval to Market a New Drug*, details the steps necessary for bringing a drug to market.(3)

Relevant to a discussion of pragmatic trials is the FDA's definition of "adequate and well-controlled studies". The FDA first defined this in 1985 and it was most recently amended in

March 2002. The Federal Register defines the adequate and well-controlled studies that must be presented by industry for determination by FDA of whether there is substantial evidence to support the claims of *effectiveness* [italics added] for new drugs. The regulations describe an adequate study as having the following elements: a clear statement of the objectives of the investigation and a summary of the proposed or actual methods of analysis; a design that permits a valid comparison with a control to provide a quantitative assessment of drug effect; a method of selection of subjects that provides adequate assurance that they have the disease or condition being studied; a method of assigning patients to treatment and control groups that minimizes bias; measures to minimize bias on the part of the subjects, observers, and analysts of the data; methods of assessment of subjects' response that are well-defined and reliable; and an analysis of the results to assess the effects of the drug.(4)

We highlight this description here in order to prompt consideration of whether PCTs would or would not meet the FDA's definition of adequate and well-controlled studies.

### Guidance Documents from the FDA

In addition to regulations, the FDA generates guidance documents.(5) These are documents prepared for FDA staff, applicants/sponsors, and the public that describe the agency's interpretation of a regulatory issue. These documents are intended to assist those who must comply with the regulations. Although many of these guidance documents are relevant to our discussion, one which is particularly relevant is the Guideline for the Format and Content of the Clinical and Statistical Sections of an Application, and two guidelines which are part of the International Committee on Harmonization recommendations: E3 Structure and Content of Clinical Study Reports and M4 Common Technical Document for the Registration of Pharmaceutical for Human Use. Another relevant guideline is the Guideline for the Study of Drugs Likely to be Used in the Elderly, published in 1989, and Patient-Reported Outcome Measures, published in 2006. These can be reviewed from links on the referenced website.

One additional guidance document of high relevance is E 10 Choice of Control Group and Related Issues in Clinical Trials.(6) This document, from May 2001, is intended to assist applicants in choosing a control group for clinical trials intended to demonstrate the efficacy of a treatment. The guidance document states that each type of control group is appropriate in some circumstances, but none is usable or adequate in every situation. Placebo-controlled trials seek to show a difference between treatments when they are studying effectiveness, but may also seek to show lack of difference (of specified size) in evaluating a safety measurement. Active control trials can have two distinct objectives: (1) to show efficacy of the test treatment by showing it is *as good as* a known effective treatment or (2) to show efficacy by showing *superiority* of the test treatment to the active control. The documents states, "In some cases, the focus of the trial is on the comparison of one treatment with another treatment, not the efficacy of the test drug per se... It is not necessary to demonstrate superiority to the active comparator, and, depending on the situation, it may not be necessary to show non-inferiority ... a less effective treatment could have safety advantages and thus be considered useful." The document reminds the readers that in placebo-controlled trials efforts are made to improve compliance and increase the likelihood that the patient population will be responsive to drug effects to ensure that an effective treatment will be distinguished from placebo. In contrast, in trials

intended to show that there is *not* a difference of a particular size (non-inferiority) between two treatments (such as may be the case in trials using active comparators), there may be a much weaker stimulus to engage in efforts to ensure study quality that will help ensure that differences will be detected.

### *B. Needs of Decision-Makers Before, During, and After Drug Approval*

The preceding section provided a brief history of drug approval in the U.S. and a basic review of select FDA regulations and recommendations. Important to a discussion of PCTs, is a discussion of the *needs* of different decision-makers before and after a drug is approved. We suggest that these disparate needs can guide the discussion about the value of PCTs, their design, and potential barriers and challenges.

Regulators We focus here largely on the FDA but this discussion is relevant to regulators world-wide, and particularly in Europe and Japan where the agencies have joined with U.S. regulators in the International Committee on Harmonization. What does the FDA need from phase III or IIIb trials? The FDA needs information that demonstrates that the drug is likely to be safe in much of the population. If the drug is approved, it can be used for any indication and in any patient, with some exceptions for specially regulated drugs. For this reason, the FDA needs pre-approval subgroup information in order to comment on safety in populations that are not necessarily the target population but who may receive the drug. Therefore, FDA needs fairly detailed characterization of the treated population. Regulators need information about the dose-response relationship between the clinical outcome and the drug dose. If safety is inversely related to the dose (as is often the case), demonstration of a dose-response relationship is necessary to minimize drug exposure. Additionally, the FDA clearly needs information about a drug's efficacy. Since the Kefauver-Harris Drug Amendments of 1962, this has been a primary requirement for drug approval.

After approval, regulators need information on safety from a larger population in order to detect infrequent events. They also need information on long-term safety, if it is a drug that is intended for long-term use, or if it is a drug that may be misused over a long-period of time (e.g. chronic use of ibuprofen poses different risks than short term use). These data often comes from safety surveillance systems or from data requested from industry as part of risk management strategies. If a manufacturer is seeking licensing changes or approval for new indications, the FDA requires additional data about efficacy and safety to support these actions. Presently, the FDA has little application for effectiveness data gathered after approval. It is unlikely to inform decisions after approval, unless it is weighed against safety should safety concerns arise.

The question posed here is whether the FDA would benefit from (or might come to require) evidence of *effectiveness* before drug approval, even effectiveness against active comparators. Many will argue that this will substantially slow the drug approval process; however, this may be the logical extension of the *intent* of the 1962 amendments.

Pharmaceutical or biotechnology companies These decision makers are bound by the regulations of the FDA. Simplistically, manufacturers make decisions to increase revenue from their products while not harming patients. Therefore, they are interested in demonstrating

efficacy of their product relative to existing products – prior to approval to get approval, and after approval to increase sales. In the pre-approval period, as well, pharmaceutical companies need pharmacoeconomics information for pricing before launch.

Presently, some postmarketing evaluation is for safety evaluation. Some of these activities are initiated by the manufacturers, but many are part of risk management plans developed along with the FDA. Since passage of the FDAAA, these activities can now be mandated. There is no mandate for industry to produce effectiveness data in the post-approval period. That is, there is no requirement that a product be tested outside of a carefully controlled setting to demonstrate benefits from its use. In the post-approval period, manufacturers proceed with efficacy trials to expand the approved indications for the drug, and acquire both efficacy and effectiveness data to demonstrate superiority over competitors for marketing purposes. Manufacturers aim to safely expand the population of users of their products and therefore may be motivated to conduct trials in diverse patient populations, if they expect benefits to be broadly demonstrated.

Clinicians Clinicians make decisions for patients on an individual basis and are therefore most interested in knowing whether a medication should be used in a given to a specific patient. In the pre-approval period, clinicians may have the opportunity to guide a patient to a clinical trial. In this setting, clinicians are interested in knowing the safety demonstrated in early human testing, and the potential benefit that the patient might experience if he receives the experimental drug.

After approval, a clinician may still be interested in results of drug efficacy trials (over effectiveness data). If results are reported in narrow strata, the trial results can be applied precisely to his/her patient if the patient is like the enrolled patients. Similarly, data on safety in patients much like the patient under consideration will be encouraging to a clinician. However, given that every clinician has a heterogeneous panel of patients, varying perhaps by age, sex, ethnicity, disease severity; a clinician may also be interested in knowing how a drug works in a broad population that *includes* patients like his or her own. The decision-making process for a clinician is simpler if a drug has proven effectiveness across a diverse patient population.

Patients A patient, like a clinician, is primarily interested in safety and in how well the drug will work given their specific situation. In the pre-approval period, patients need data about safety in early human testing, and the potential benefit that he might experience if he receives the experimental drug.

After approval, a patient wants to know how effective the drug is in a population of patients among which he can include himself. Knowledge of safety and effectiveness at a population level is only of interest if he/she is like those in whom it was tested. Patients, however, may value different outcomes than are valued by a clinician, and therefore, data from effectiveness studies may have different relevance to a patient. A patient may value quality of life information or ease of adherence or other patient-relevant outcomes more than the clinician making decisions about use of a drug.

Payers In the pre-approval process, payers have few decisions to make about a drug under investigation except for whether to pay for the drug in the setting of a clinical trial, which rarely

reflects usual care settings. They may also be involved with decisions as to whether to pay for treatments associated with receipt of the drug or injuries incurred from participation in trials of experimental drugs. They may use information from early human trials in this decision making process.

In the post-approval period, payers are tasked with making coverage and reimbursement decisions. These decisions are based on results of efficacy trials and on the pharmacoeconomic information collected (and/or modeled) during these early trials. Payers have a powerful role in limiting access to drugs through coverage decisions. Payers clearly need information on effectiveness in usual care settings, as well as safety and costs in these settings, to make more informed decisions.

Funding Agencies Funding agencies, such as National Institutes of Health (NIH), pay for the vast majority of biomedical research in this country, including early-phase research and some product development. In the pre-approval period, these agencies have relatively limited roles as decision-makers, except for decisions to fund drug development studies. In the pre-approval period, the funding agencies are presumably motivated by the scientific questions that can be answered in trials and the anticipated benefits to the population that warrant support of these activities.

In the post-approval process, these agencies fund evaluations of approved drugs. This is expected to increase with the rapidly growing attention to the evaluation of the comparative effectiveness of interventions. The information that funding agencies need in order to make funding decisions include information about gaps in knowledge that effectiveness studies could answer that pre-approval efficacy studies did not, information as to why industry is not funding such studies, and a strong rationale for how studies will improve the health of the population.

### *B. Pragmatic Clinical Trials Definitions and Background*

In this section, we discuss key definitions and the historical developments relevant to an enlightened discussion of PCTs. We begin with the seminal paper by Schwartz and Lellouch, *Explanatory and Pragmatic Attitudes in Therapeutic Trials*, which provided the early definitions in this field and a framework for later advances.(7)

Dr. Daniel Schwartz (b.1917) began his career as an engineer working for the national tobacco monopoly in France.(8) In this capacity, he helped investigate whether smoking and lung cancer were associated in the French, as had been seen in Britons and Americans. As a result of this work, Schwartz was hired by the Institut Gustave Roussy in 1956 to begin a statistical research unit. Through the 1960's, Schwartz and colleagues developed statistical applications for epidemiology, clinical trials, and laboratory experiments, favoring case-control studies as their investigative method of choice.

The influential early paper about PCTs was published in 1967, before Schwartz had ever done a clinical trial. In this paper, Schwartz and Lellouch asserted that most therapeutic trials are inadequately formulated *at their inception*. They suggested that the design of a trial needs to be directed by the goal of the investigation. Is the goal to acquire information about the true effects

of a treatment (i.e. to verify a biological hypothesis) or is the goal to gather information needed to make a decision about a treatment? They described the former goal as requiring an “explanatory” trial and the latter as requiring a “pragmatic trial.” They maintained that, in general, an explanatory trial will give an answer to a scientific problem but will only sometimes answer questions about the “practicability” of the treatment approach, while the converse is true of pragmatic trials.

Schwartz and Lellouch further explained that the context in which a trial is done determines if the trial is more explanatory or more pragmatic. By context, they meant all aspects of trial design including the setting, the source of participants, and the chosen interventions. The terms they used to describe the setting of the trials are “laboratory”, which is associated with explanatory trials, and “normal”, which is associated with pragmatic trials. Schwartz and Lellouch described in detail the appropriate selection of outcomes to be evaluated in trials, contrasting explanatory and pragmatic trials. They used as an example the outcome of “returning to work”. They acknowledge that this is an important outcome to patients and appropriate for evaluation in a pragmatic trial, but as it conveys little biological information, it may not be a relevant outcome in an explanatory trial. They cautioned against inclusion of multiple endpoints if they are not appropriate to the goals of the trial.

Additionally, they made recommendations about the handling, in the analysis, of subjects who withdraw from the study after enrollment. They suggested that in a pragmatic trial, if it is stated at the outset that in one of the treatment arms subjects are permitted to change therapies or to discontinue therapy, those individuals stopping treatment should not be considered to be withdrawals, as this was established as the time of study design. They contrasted this with the handling of subjects who withdraw from a trial designed to be explanatory (e.g., they withdraw because the drug “tastes nasty”). The other extreme of analyses is to remove these individuals from the trial entirely, as they were not able to receive the treatment intervention. Again, the goal of the trial directs the analysis.

Schwartz and Lellouch described how closely linked the analytic decisions are to the selection of individuals for inclusion in the trial. With an explanatory approach, a strict patient selection criterion may be used in order to render the population homogenous and to reduce the withdrawal rate. However, in a pragmatic trial, a heterogeneous population with more withdrawals is acceptable. Patients should not be turned away from a trial for reasons that would not preclude use of the intervention in usual practice. “The trial must represent as far as possible the population to which the results are to be extrapolated.”

They concluded their manuscript with a discussion of statistical methods for the analyses and a sensitive discussion of the ethics of these trials. We note that regulators presently require that explanatory trials be completed before pragmatic trials. Schwartz and Lellouch question “Should one prefer the goal of immediate applicability with a sacrifice of true understanding, or the more distant goal which may lead to greater enlightenment and which may prove more fertile for the future?”

These concepts were slow to spread in the scientific community. Dr. Peter Armitage at University of Oxford, and Dr. David Sackett at McMaster University were earlier users of this



work, largely in the context of discussion of how to analyze patients who withdraw early from clinical trials.(9;10) Over the next two decades, the 1980s and 1990s, the literature had relatively little about pragmatic trials, although there was interest in defining efficacy trials in contrast to effectiveness trials. These were not phrases used by Schwartz and Lellouch.

Dr. Alvan Feinstein at Yale University appreciated early the limitations of clinical trials. In a Perspective piece in the *Annals of Internal Medicine* in 1983, Feinstein wrote on challenges in trial design that stem from conflicting goals of trials. (11) He said that individuals who want answers to pragmatic questions in clinical management want trials that incorporate heterogeneity and ambiguity, and other “messy” aspects of clinical practice. The opposing viewpoint is that these trials yield “messy” answers; proponents of the latter favor trials using homogeneous groups that reduce or eliminate ambiguity. Feinstein used the terms “pragmatic” and “fastidious” to differentiate among these trial types. He described in this paper issues that had been raised by Schwartz and Lellouch about the design of pragmatic and fastidious trials, including patient selection, choice of comparator treatments, dosage and titration of interventions, and the choice of outcomes (in particular, whether they are patient-relevant or not). He also described conflicts that arise at the time of analysis of study results particularly regarding intent-to-treat or on-treatment analyses, inclusion of patients later found to be ineligible, handling of patients who receive “unauthorized” treatments, and others. He acknowledged that there is no right or wrong approach, necessarily, “but a trial designed or analyzed with one view point will often be unable to satisfy people who hold the opposite view point, and vice versa.” He proposed in this paper some suggestions as to modifications of study designs to better satisfy both view points. He offered many suggestions including patient populations that include a separate heterogeneous group as well as a “pure” group, a pragmatic treatment arm could be added if the chosen comparator treatments seem clinically unsuitable, a “double-observer” procedure could be used to allow flexible dosing; and efforts could be made to “harden” the softer patient-relevant outcomes (such as functional status and quality of life). At the time of writing this paper, Dr. Feinstein was not optimistic that the challenges of analyzing data to satisfy both viewpoints could be overcome.

By the mid-1990’s, the term “mega-trial” was being used to describe large, simple randomized trials that are analyzed on an intention to treat basis.(12) There is overlap between pragmatic trials, as described above, and these large trial designs although they are not necessarily the same. One of the first uses of the word “mega-trial” was to describe the GUSTO trial, which enrolled 41,021 patients and randomized them to one of four thrombolytic regimens for myocardial infarction.(13) Early proponents of these “large simple” trials were Peto and colleagues at the University of Oxford. In a review paper, they provided examples of the use of this trial design to answer important questions, and asserted that there are some underlying assumptions when using this trial design.(14) The first assumption is that the real differences between two treatments in some important outcome will *probably not be large*, but even a moderate difference in an important outcome may be worthwhile [to detect]. The second is that if there is, for some readily identifiable category of patients, a moderate difference between two treatments in their effects on some specific outcome, then this difference might be larger or smaller in other readily identifiable categories of patient, but it is *unlikely to be reversed*. Two examples they describe are ISIS-2, and GISSI-1.(15;16) These trials might be considered to be examples of pragmatic trials as they share many commonalities, including the heterogeneity of

included patients. Detractors, however, argue that the between-subject variation, within each treatment group in large simple trials, makes the results of these trials difficult to apply to an individual patient. Charlton describes these trials as being based on a “methodological mistake”.(17) He states that the mistake is the assumption that a measurement can be made simultaneously more precise *and* more valid by reducing the rigor of its protocol in order to allow increased requirement of patients.

In the next decade, additional terms were used to describe more pragmatic trial designs including “naturalistic trials” and “effectiveness trials”. Efficacy studies are closest to what Schwartz and Lellouch described as explanatory trials. These are studies that aim to investigate whether an intervention works under *optimal* circumstances, or in other words “can it work?” Effectiveness studies are closer in their goals to those of pragmatic studies. They aim to evaluate whether an intervention works under usual circumstances, or in other words “does it work?” Others have since carefully articulated the differences between efficacy studies and effectiveness studies, such as is shown in this table from Bombardier and Maetzel.(18) These authors nicely highlight the difference in *motivation* for performing these studies; efficacy studies being appropriate for regulatory approval and effectiveness studies being appropriate for formulary approval. While necessarily simplistic, this reminds the reader that the design of the study has to be driven by its intended use.

**Table 1 Efficacy versus effectiveness studies**

	<i>Efficacy studies</i>	<i>Effectiveness studies</i>
Objective	Does it work under optimal circumstances?	Does it work under usual circumstances?
Motivation	Regulatory approval – FDA	Formulary approval
Intervention	Fixed regimen / forced titration	Flexible regimen
Comparator	Placebo Arbitrarily chosen comparator	Usual care Least expensive / most efficacious
Design	RCT - strict control	RCT or open label - minimum control
Subjects	Selected or "eligible" subjects High compliance	Any subjects Low compliance
Outcomes	Condition-specific Strong link to mechanism of action	Comprehensive (for example, QoL, utilities) Weak link to mechanism of action
Analysis	Short-term horizon Protocol adherers	Short and long term horizon Intent to treat

QOL-quality of life, RCT-randomized controlled trial

Garlehner and colleagues later operationalized the distinction between efficacy and effectiveness studies.(19) They created a tool to assist in distinguishing between these two designs. They reviewed published literature for definitions of effectiveness and efficacy and proposed criteria that might differentiate between the two. They then asked experts in evidence-based medicine to classify articles as efficacy or effectiveness studies, based on their own knowledge, and then applied their instrument against this reference standard. The proposed criteria included the following: the population is in primary care, there were less stringent

eligibility criteria, the study evaluated health outcomes, the study had a long duration and clinically relevant treatment modalities, it assessed adverse events, the samples size was adequate to assess a minimally important difference from a patient perspective, and it was analyzed using intent-to-treat methods. Studies fulfilling 6 or more of the criteria were found to be very likely to be effectiveness studies.

A recent advance in differentiating pragmatic and explanatory trials was the exercise by Thorpe and colleagues who devised a graphical method by which an investigator, or reader, can evaluate where on the explanatory – pragmatic continuum a study lies.(20) This grew out of discussion among investigators involved in the PRACTiHC project, a Canadian and European Union initiative to promote pragmatic trials in low and middle-income countries. They called this the Pragmatic-Explanatory Continuum Indicator Summary (PRECIS).The key domains which distinguish explanatory and pragmatic trials are shown in Table 2.

**Table 2. Domains for the PRECIS Graphic**

1. The eligibility criteria for trial participants.
2. The flexibility with which the experimental intervention is applied.
3. The degree of practitioner expertise in applying and monitoring the experimental intervention.
4. The flexibility with which the comparison intervention is applied.
5. The degree of practitioner expertise in applying and monitoring the comparison intervention.
6. The intensity of follow-up of trial participants.
7. The nature of the trial's primary outcome.
8. The intensity of measuring participants' compliance with the prescribed interventions, and whether compliance-improving strategies are employed.
9. The intensity of measuring practitioners' adherence to the study protocol, and whether adherence-improving strategies are employed.
10. The specification and scope of the analysis of the primary outcome.

Thorpe and colleagues, in this paper, urged investigators to clarify which question the trial is designed to answer (i.e. is it an explanatory or pragmatic trial). They acknowledged that for some interventions, this distinction is of little practical difference – such as for trials of simple interventions in an acute care setting, like aspirin for myocardial infarction. They recommended that the investigator specify the settings or conditions for which the trial is intended to be applicable, specify the design options *at each extreme* of the explanatory-pragmatic domains, and then see how pragmatic or explanatory the trial under consideration is. They suggested that this exercise is useful to make the investigator focus on each decision that is made when designing a trial so that the trial can be made to be more or less pragmatic, as needed to answer the question.

These careful definitions that differentiate pragmatic from explanatory trials led to an extension of the CONSORT statement in 2008.(21) This, too, originated from discussions in the meetings that yielded the PRECIS tool. The CONSORT guidelines are intended to help investigators report the results of trials in the medical literature. The original CONSORT statement, published in 1996, was developed to improve the reporting of parallel group randomized trials. Over the years, the original 22-item checklist has been updated and modified to improve the

reporting of more specific trial designs. The extension of CONSORT for PCTs retains the original 22-items but provides additional text for 8-items thought to be unique to the reporting of PCTs. (See Appendix). The authors were cautious in not wanting to promote the belief that there is a dichotomy between explanatory trials and PCTs. They encouraged thinking of these trials on a continuum so that elements important to report in explanatory trials remain important when reporting on PCTs.

There has been some dissension about definitions, recently. Commentaries published in the Journal of Clinical Epidemiology in 2009 have highlighted differences in opinion.(22-25) Karanicolas et al challenged the prevailing explanatory-pragmatic framework.(22) They maintained that this framework “confounds purpose with structure” and “ignores the varying perspectives of those using RCT results” to make decisions. They thought that most authors have severed the link between the goals of the trial – answering questions relevant to decision-makers – and the design of the trial. They feared that the design of pragmatic trials, as discussed above, exclusively answers questions from a public health perspective, and provides little information that is relevant to clinicians caring for individual patients. They felt that the latter are the “real-world” health care decisions that trials should be designed to address. They used the term “practical” rather than pragmatic to describe trials that yield comprehensive information to guide health care decisions, and “mechanistic” to address trials that explain biological relationships. In this paper, the authors used the concept of minimally important difference (MID) to represent the smallest difference in treatment effects that would lead to a change in management. Acknowledging that the MID is often hard to know, they stressed that practical trials that demonstrate differences between groups that are bigger than the MID are trials that have definitively demonstrated a benefit to an intervention. A practical trial can be, legitimately, directed at highly compliant patients managed by skilled specialists if this is the setting in which the intervention is intended for use. In other words, this may be the usual care setting for some interventions (perhaps cancer therapies)

The “pragmatists”, lead by Oxman and colleagues, countered that Karanicolas’ use of the term “practical” to describe trials which may be performed in carefully selected patients in optimal clinical settings distorts the idea of a pragmatic trial.(25) These trials are more likely to demonstrate that an intervention is effective and may not provide valuable information to most decision makers, including most clinicians who do not exclusively engage with highly compliant patients in supportive clinical settings. They feared that these trials would look much like explanatory trials, except perhaps for the choice of outcomes.

The response by Karanicolas and team was that there are many points of agreement between the two “sides”; however, differences remain.(23) They contended that clinicians find more value in trials that are designed to answer a question in a narrower population if it maximizes applicability or directness to his/her patient. They encouraged researchers to carefully define the settings in which their trial results will be practical, and report this explicitly in the protocol and in the report of the results. They also discussed the difference between *context* and *perspective*, suggesting that pragmatists are focused most on context – i.e. the question being asked by the study. They prefer to focus on perspective – i.e. who needs to learn information from the study.

The last word was had by the pragmatists who remained unconvinced by Karanicolas, et al. Their definition of context was “the circumstances where and when an intervention is implemented that modify (or might modify) the effectiveness of the intervention.”(24) This does not seem to be how Karanicolas and team defined context. Furthermore, there remained disagreement about the interpretation of the word *perspective* as well. The pragmatists were unconvinced they had confounded purpose with structure. They conceded, however, that there are sometimes reasons for not having broad inclusion criteria, but cautioned that trial results are *always* average results and there is never information specific to an individual patient outside of an n-of-1 trial. This piece concluded with a clear summation:

“Although explanatory trials may help to understand mechanisms of treatment effects, they are primarily designed to test whether interventions have hypothesized effects under optimal circumstances, not necessarily to investigate ‘possible mechanisms of effect.’ The strength of explanatory trials is that a ‘negative’ result can directly inform practice, because an intervention that does not work under optimal circumstances is unlikely to work under usual circumstances. The weakness of explanatory trials is that ‘positive’ results do not directly inform practice, although they may directly inform practice under a narrow set of optimal circumstances and they can inform decisions about future research. Pragmatic trials, on the other hand have the opposite strengths and weaknesses. The weakness of pragmatic trials is that with ‘negative’ results it is unclear whether the intervention is ‘worthless’ or whether it might, in fact, be worthwhile under some (more optimal) circumstances or for a subgroup of patients. The strength of a pragmatic trial is that ‘positive’ results can directly inform decisions under the ‘usual’ conditions for which the trial was intended to be applicable.”

### *C. Why Conduct Pragmatic Clinical Trials?*

With some clarity now about definitions, we suggest possible uses of PCTs, describing how the evidence generated in PCTs differs from and complements information from explanatory trials, and how this evidence might be used by decision makers. Maclure, in a recent commentary, described how he would explain pragmatic trials to policy makers.(26) Although he used the word “policy makers”, the broader phrase “decision-makers” seems equally appropriate. He highlights in this brief piece the distinctions between explanatory and pragmatic trials that arise from the real-world variation among providers of interventions (often clinicians, but not always) and recipients of interventions (often patients, but not always). The distinctions that he makes are drawn from the work of Thorpe, et al, and provide a useful framework for highlighting the information that comes from PCTs that can inform decision makers.(20)

*PCTs provide information that differs from information from explanatory trials of the same intervention*

To illustrate, we provide an example of what explanatory and pragmatic trials of a diabetes drug, exenatide, might look like. The drug underwent the requisite pre-approval clinical trials,(27-29) but no pragmatic trials before approval. We acknowledge that not all of the

domains below may be relevant in the pre-approval period. We describe here what could be learned in pragmatic trials because of differences in the following factors:

i. Patient heterogeneity

In an explanatory trial of an injectable diabetes medication, exenatide, patients are selected to have similar severity of illness (similar duration of diabetes, similar hemoglobin A1c measures, similar extent of diabetes-related complications). In a pragmatic trial of exenatide, the enrolled patients will be at different stages in their illness, be on any number of combinations of other diabetes therapies, and have varying degrees of glycemic control at baseline. What do we learn in the pragmatic trial? We learn how exenatide works within patient subgroups, if the trial is large enough, and we get an estimate of the effectiveness and safety of exenatide *across* these subgroups (an estimate of absolute risks and risk reductions).

ii. Patient adherence

In an explanatory trial of exenatide, patients may have been involved in a run-in process where they demonstrated their ability to adhere to the use of an injectable medication. They would have had scheduled follow-up visits with a study nurse to assess and encourage adherence to the study drug (or comparator). In a pragmatic trial, adherence is expected to be variable, as adherence is in practice, and possibly more so given that this is an injectable drug which may pose more adherence challenges than an oral medication. What do we learn from the pragmatic trial? We learn how exenatide works in patients who have varying degrees of adherence (which might be a stratifying variable in an analysis), and, importantly, we learn whether adherence to this medication is *so* challenging that it is a useless medication in a usual clinical setting.

iii. Comparator flexibility

In an explanatory trial, the comparator may be a placebo or it may be an established second line therapy for treating diabetes, such as a thiazolidinedione. In a pragmatic trial, the comparator may be chosen by the treating clinician – he/she may be advised by the investigator to add therapy as clinically indicated, or it may be required to be an active comparator. What do we learn from the pragmatic trial? We learn about the risks and benefits of exenatide compared to the usual “next” therapies that are chosen in practice.

iv. Intervention flexibility

In an explanatory trial, there is little flexibility in use of the intervention under investigation – the dosing is carefully specified and the allowable concurrent therapies are specified. In a pragmatic trial, there may be more flexibility allowing some titration-to-response (5mg or 10mg of exenatide) by the clinicians and freedom regarding adjunctive therapies such as second or third agents for glycemic control. What do we learn from the pragmatic trial? We learn how clinicians use the drug in a usual care setting, including dose, titration rate, and choices when additional medications are added, and we learn the risks and benefits of these practices.

v. Clinician adherence

In an explanatory trial, adherence by the clinician to the study protocol is strongly encouraged by study staff. Clinicians are visited by staff to encourage compliance and incentives are provided to maximize attentive participation. In a pragmatic trial, the clinician will not experience intensive involvement of study staff. The clinician is free to stop exenatide or prescribe another medication in place. What do we learn from the pragmatic trial? We learn whether it is a challenging drug for clinicians to prescribe and monitor. We learn whether clinicians are faced with requests by patients to discontinue exenatide and how they respond.

vi. Clinician heterogeneity

In an explanatory trial, clinicians are selected to be fairly homogenous so that the patient populations are homogenous. Endocrinologists at academic centers may be selected as the prescribing physicians, or primary care doctors in community practices. In a pragmatic trial, clinicians from many different subspecialties and practice settings would be invited to participate. What do we learn from the pragmatic trial? We learn whether there is willingness of physicians across practice settings and specialties to use exenatide; we learn whether these diverse clinicians can prescribe the medication (have sufficient office staff to do the requisite patient education for an injectable medication); we learn if the use of the medication (rapidity of titration, selection of concomitant therapies) differs by practice setting or specialty and whether these differences translate into different risks and benefits for the patient.

vii. Outcomes under investigation

In an explanatory trial, the outcomes are carefully specified and may include intermediate outcomes (such as change in HbA<sub>1c</sub>) and clinical outcomes (admission for hypoglycemia). In a pragmatic trial, the primary outcome will be a patient relevant outcome such as hypoglycemia, a diabetes complication like amputation, death, or a measure of quality of life. What do we learn from the pragmatic trial? The outcomes are different and provide different information to decision-makers. Exenatide is expensive, but if it reduces downstream costs associated with amputation or retinopathy, it may be a worthwhile drug to payers. If patients feel well on the drug, this is a relevant outcome to patients choosing to use exenatide or not.

*PCTs provide information needed by decision-makers*

Tunis and colleagues prepared a special communication for JAMA in 2003 describing the role that PCTs may have in generating evidence for decision makers.<sup>(30)</sup> At the time of their writing, the NIH did not have an organized, systematic mechanism for identifying the highest priority questions of decisions makers, i.e. questions that would be amenable to answering with a PCT. They recommended that there be an institution, perhaps the Institute of Medicine, with primary responsibility for identifying and prioritizing clinical research questions. A strategy to improve clinical research might include encouraging decision-makers to *require* high quality evidence (including from PCTs), the development of an infrastructure within the primary care setting for the conduct of PCTs, and the training of investigators in the conduct of these challenging trials.

Appropriately, Tunis and colleagues discuss the challenges of funding options for PCTs, undeniably expensive trials. They noted that industry funding of trials is largely for phase I, II, and III studies, with only 10% of funding going to phase IV studies at that time, which may be understandable as the regulatory framework of the FDA is “not structured to ensure that research is conducted that will inform optimal clinical use of . . . technologies.” Thus, industry has little motivation to conduct these trials. They anticipated that NIH funding for these trials would increase with its commitment to improving public health through research. In addition, Tunis, et al, suggested that payers and purchasers may have the most motivation to fund PCTs as they are most likely to benefit, financially, from a strong evidence-base. Payers such as the Veterans Affairs and the National Health Services of Canada and England have indeed funded PCTs, although comparatively few are tests of medications. (31-35)

Freemantle, with colleagues in France and Canada, made recommendations that are complementary to Tunis et al. about the use of “real-world” trials to answer questions relevant for health policy and reimbursement.(36) They remind the reader that trials designed for regulatory purposes (“confirmatory” trials) many often not answer important clinical and economic question about how a new treatment should best be applied. They reason that trials for licensing require demonstration of an effect that is unlikely to be due to chance. The demonstration of the value of a product, to patients and payers, is not a requirement for licensing, but is undeniably necessary information.

Maclure suggested that although we can enumerate many differences in the design and conduct of PCTs as compared to explanatory trials, the principle difference is how they will be used by decision-makers. Explanatory trials are used, like briefing notes “*For Information*”, while PCTs are used “*For Decisions*.”(26)

#### **IV. Examples of Pragmatic Clinical Trials: Design and Implementation**

We are highlighting the designs of a select group of pragmatic trials to illustrate key aspects of their design and implementation. None of the trials was of drugs in a pre-approval setting. We note that it is challenging to identify pragmatic trials in the literature. The term “pragmatic” is still infrequently used by study authors. A systematic review of this topic, of low quality, identified 283 citations from 1995-2002 that used the keyword “pragmatic”. Of these, 95 were clinical trials. However, only 4 of these followed methodology resembling that described by Schwartz and Lellouch.(37)

##### *a. Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT)*

We begin with this familiar trial designed to test the effectiveness of chlorthalidone (a diuretic) in comparison to three different medications from other therapeutic classes. (38) This trial was funded by the National Heart Lung and Blood Institute (NHLBI) of the NIH along with the Department of Veteran’s Affairs. A subset of the participants in this antihypertensive trial was enrolled in a lipid-lowering trial simultaneously. The primary hypothesis of the antihypertensive trial component was “that the combined incidence of fatal coronary heart disease and nonfatal myocardial infarction (first or recurrent) will be lower in hypertensive patients randomized to 1) calcium antagonist (amlodipine), 2) an ACE inhibitor (lisinopril), or



3) an  $\alpha$ -adrenergic blocker (doxazosin) as first-line therapy than in those randomized to a thiazide-like diuretic (chlorthalidone) as first-line therapy.” The trial had 8 secondary outcomes, as well. The inclusion criteria were blood pressure measurements within a specified range, and 6 additional criteria used to define a group of participants at elevated risk for the primary outcome. The exclusion criteria were largely for patient safety and included recent myocardial infarction or stroke, requirement for a specific antihypertensive medication, contraindications to the study drug, and illnesses likely to lead to death from a non-cardiovascular cause in the near future. There was, additionally, an exclusion of participants deemed unlikely to comply with the protocol.

Participants meeting inclusion criteria were randomized to one of four arms after having been safely “stepped-down” from their existing antihypertensive medication. The participants and investigators were masked as to the treatment arm. The treating study investigator was free to add a second agent (and third agent) as needed to achieve blood pressure control and was provided with a list of acceptable agents that did not include the drug classes under consideration. Patients were seen at one-month intervals for dose-titration until successful control was achieved. Endpoints were noted at each visit by the study investigator and confirmed by chart review. The primary analysis was time to development of fatal coronary heart disease or a non-fatal myocardial infarction and groups were compared with a log-rank test.

By their definition, this large, simple design was appropriate because “1) a very large sample size is needed, 2) a streamlined protocol is possible, 3) the targeted conditions are commonly encountered in clinical practice, and 4) there is widespread interest in the study question among clinicians.” They anticipated that 400-500 physician-investigators would need to enroll 40,000 patients to meet the recruitment goals. These treating physicians were considered to be co-investigators and were required to participate in a training session as well as periodic “refresher courses”.

#### *Challenges Highlighted by this PCT*

The great number of participants required for this study is a challenge. However, for trials in conditions that are highly prevalent such as hypertension, this may not be a major obstacle to completing the trial, although it is expensive. The investigators chose to use an active comparator. This is appropriate as their primary hypothesis was whether the other agents were superior to chlorthalidone. They might have used a placebo comparator if they had assurance that the treating doctors would responsibly add additional agents as needed. The use of placebos in PCTs can be problematic as there is the risk of under-treatment of the participants and harm. These trials are typically longer than explanatory trials so the argument that the participant is only briefly exposed to the harm of under-treatment cannot be made. ALLHAT used a combined endpoint which is common, particularly in cardiology trials. Combined endpoints may be even more commonly used in PCTs than in traditional efficacy trials if it is thought that the patient-relevant outcome is a combined outcome. Combined outcomes, however, present analytic challenges if competing risks of events are not appropriately managed.

This trial had the participants presenting monthly for dose titration. This highlights the challenge of treating patients safely in a trial setting (with frequent dose titration) while attempting to mirror usual practice. Could this trial have been done in a pre-approval setting? It is possible. Because the clinicians were permitted to titrate additional medications for blood pressure control if the trial medication was ineffective, the safety of the participants was assured. Would it have been ethical to expose thousands of patients to medications that did not necessarily have proven efficacy let alone effectiveness? This is less clear.

#### *b. Effectiveness of Antipsychotic Drugs in patients with Chronic Schizophrenia*

This trial is from the Clinical Antipsychotic Trials of Intervention Effectiveness Investigators (CATIE).(39) This trial was sponsored by the National Institute of Mental Health and was designed to compare the effectiveness of atypical and conventional antipsychotic medications. The primary aim was to determine the comparative effectiveness of a representative antipsychotic (perphenazine) and several atypical antipsychotic medications for a representative sample of patients seeking treatment for chronic schizophrenia, as measured by all-cause treatment discontinuation rates and associated measures of effectiveness and safety. Patients without adequate symptom relief during the first phase of the trial were offered enrollment in a second phase in which atypical antipsychotics were compared to clozapine, or enrollment in a trial where atypical antipsychotics were compared to ziprasidone. There was also a phase 3.

The investigators intentionally specified few exclusion criteria, aside from those required for safe use of these drugs. They also excluded treatment refractory individuals who were extremely unlikely to respond in any of the treatment arms. They wanted to mirror treatment decisions that a clinician in practice would make when deciding whether a patient should use a conventional or atypical antipsychotic. The patients and clinicians were masked to the treatment drug except for clozapine which requires white blood cell count monitoring. The primary outcome was time to all-cause treatment failure, marked by the discontinuation of the study medication. The investigators chose this measure as it is a distinct measure that reflects both efficacy and side effects, and is clinically meaningful to patients and clinicians.

The investigators aimed to enrolled 1,500 individuals from 50 clinical sites and follow them for 18 months. The treating physicians were allowed to titrate the medications to effectiveness. All patient participants were offered psychosocial interventions as well as an educational plan.

#### *Challenges Highlighted by this PCT*

This is a challenging patient population to study made even more challenging by the broad inclusion criteria. The patients did not need to prove themselves to be compliant patients or to have strong family support for enrollment. This trial could not have been done with a placebo comparator – the primary outcome was treatment discontinuation and it would have been unsafe to have a patient on no medication. Could this trial have been done in a pre-approval setting? This would have been more challenging than ALLHAT (above) because there was no provision for adding additional medication for disease control. If any the drugs under consideration was truly ineffective, the patient would be receiving no treatment and would be in a less closely monitored setting than in a traditional efficacy trial.

### *c. Similar Effectiveness of Paroxetine, Fluoxetine, and Sertraline in Primary Care*

This trial was designed to compare the effectiveness of 3 antidepressants in depressed primary care patients. (40) It was sponsored by Eli Lilly and Co. but the investigators did not need the company's approval to publish the results. Six-hundred and one patients were enrolled from clinical practices that were part of two primary care research networks. Patients were included in the study if their *own* primary care physicians determined that they had depression that warranted treatment with medication. Patients were excluded if they were cognitively impaired or had bipolar illness, were using cocaine or opiates, were pregnant, were presently on another antidepressant medication, were terminally ill, in a nursing home, did not speak English or could not tolerate the starting dose.

Patients were randomly assigned to one of three treatments. The trial was intentionally not blinded to allow clinicians to manage patients as they would in usual practice. All decisions regarding switching medication including adding additional medications were left up to the clinician and the patient. Outcomes were assessed using computer-assisted telephone interviews. The primary outcome was the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) Mental Component Summary. Several other depression scoring scales were also used, as well as assessment of 5 other psychological outcome measures. Social function and work function were also assessed from patient self-report. Antidepressant use, reasons for changes, and adverse effects were other outcomes of interest. Results were analyzed using an intent-to-treat analysis.

#### *Challenges Highlighted by this PCT*

This trial did not need a particularly large number of participants. The outcome measure was sufficiently sensitive to detect changes in this relatively small sample of patients. This was not a blinded study so clinicians could switch or add medications at will. Clearly this raises major challenges in the analysis as many of the patients were not on the drug initially assigned. Attributing changes on the SF-36 and other symptom surveys to medication regimens for patients on a myriad of different treatment regimens is challenging. Non-adherence to a regimen would increase the likelihood of demonstrating no difference between groups (and the possible conclusion that they are equally effective). The secondary outcomes, antidepressant use and reasons for change, are straightforward to analyze by this design. Could this have been done in a pre-approval setting? It could have been done without major safety concerns since physicians had freedom to change and add medications. This trial may not have answered the questions of the regulators regarding efficacy of the drugs in compliant patients who are encouraged to remain on their medications.

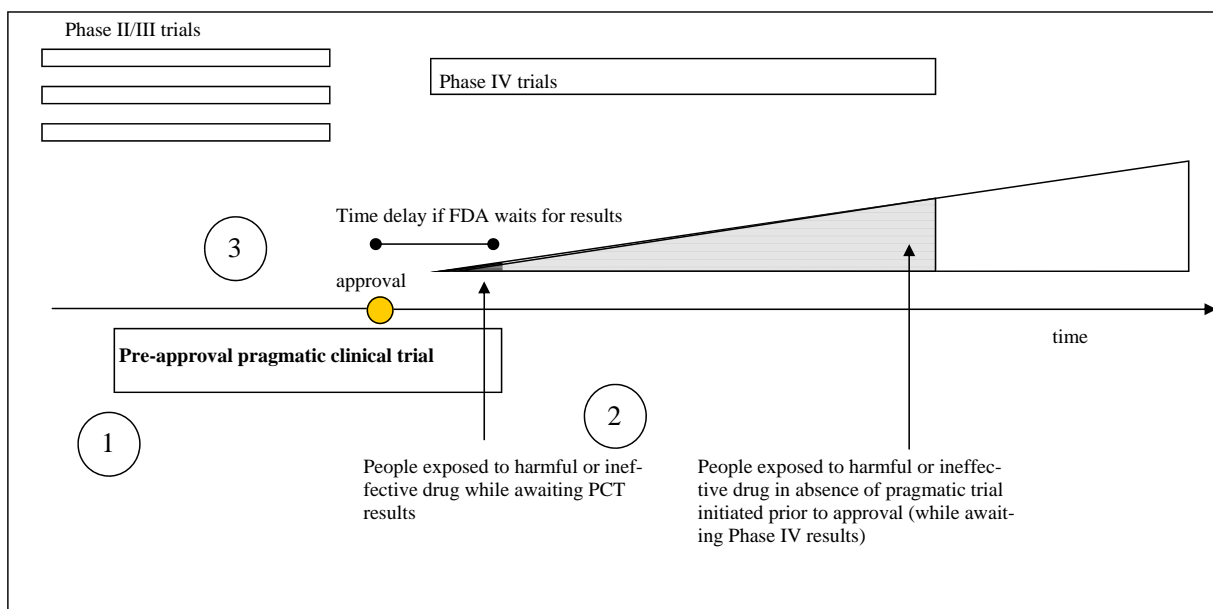
### **V. Challenges to be Anticipated**

As stated, the trials described above were all conducted after the drugs had FDA approval. These were large, expensive trials that required recruitment from many different settings. Trials with pragmatic features may be even more challenging to accomplish in a pre-approval setting. We do not mean to imply that pre-approval pragmatic trials are intended to supplant traditional

phase III trials valued by regulators. Nor are we implying that pragmatic trials are only an adjunct to phase III trials. This discussion is meant to explore the possibility of incorporation of elements of “pragmatism” earlier in the drug evaluation process so that pre-approval trials yield results valuable to a broader range of decision-makers.

In Figure 1, we illustrate some key variables that may influence the acceptability or feasibility of a PCT in a *pre-approval* setting. The challenge identified by ① is the decision about the timing of initiation of the PCT and its duration. One scenario is that the PCT is initiated early, at the traditional start time of phase III trials. FDA could make a decision to approve the drug based on results of, perhaps concurrent, conventional efficacy trials or could wait for the results of the PCT. If the FDA waits for PCT results, drug approval will be delayed, denying patients access to these medications during this time, and denying the manufacturer revenue from the drug. If the FDA proceeds with approval prior to the results, and the results are not favorable about effectiveness or safety, there will be people who have been exposed to an ineffective or harmful drug that would not have been had the FDA waited for the trial results. Of course, this would not be different from the present situation, in which people are exposed to drugs with uncertainty about effectiveness and safety because there is no requirement for PCTs.

**Figure 1.**



The necessary duration of the PCT will be determined by the success at participant recruitment, by the number of participants required for an adequately powered study, and by the clinical outcomes (including risks) and whether they require a long observation period. The shorter the PCT, the easier it would be for the FDA to delay approval until the results are available. Incorporation of pragmatic features into traditional efficacy trials seems likely to lengthen the approval process.

The challenge identified by ② are challenges specific to the patient population and the drug under consideration. If the FDA aims to minimize harm, it will ideally individualize decisions regarding the need for information from PCTs depending on the drug and the condition for which it is indicated. Drugs which are “me-too” drugs that may contribute relatively little to the population’s health, might be best approved after completion of trials with pragmatic features. Decision-makers other than the FDA (such as clinicians) will gain little new information from traditional efficacy trials of these drugs but may gain important information from pragmatic studies. Drugs for treating serious conditions with few alternatives might be approved solely on the basis of traditional efficacy studies. In this situation, the lack of effectiveness data (given efficacy data) may be tolerable. One could argue that drugs that are first-in-class, where there are no comparable therapies, should be approved without awaiting longer, pragmatic-type trials. However, these drugs may also have unanticipated risks which could be observed in a PCT which would suggest that awaiting the results is prudent.

Similarly, the setting of use may determine the need to await PCT results or not for approval. Drugs which are administered in a setting in which effectiveness is likely to *closely mirror* efficacy might be approved without PCTs. An example would be a drug that is used in a closely monitored setting by equivalently trained clinicians, such as an intravenous cancer chemotherapeutic agent. In contrast, a drug that is challenging to use by patients in an outpatient setting, such as an injectable like exenatide or a drug with substantial side effects like some antiepileptic drugs, might generate different results in a PCT than in phase III trials. Incorporation of pragmatic features into these pre-approval trials may be very valuable. Likewise, if the treatment effect is expected to be heterogeneous across the population of users or if there is expected to be a lot of off-label use of the drug, the FDA might opt to require pragmatic trials before approval, enrolling a broad swath of the population of potential users.

The challenge identified by ③ are ethical issues. PCTs enroll a large number of participants. If these trials are initiated prior to completion of traditional phase III studies, a lot of people are exposed to drugs with potentially limited efficacy, let alone effectiveness. Presumably, the early phase studies of risks associated with the drug will be complete before a PCT is initiated. While there is uncertainty about the harms of the drug when people are exposed during a pre-approval PCT, this is probably comparable to the uncertainty when a drug is approved without such information. (Table 3)

**Table 3. Factors which Could Determine Need for Pragmatism in a Pre-approval Trials**

	<i>Favors a Trial with Pragmatic Features</i>	<i>Favors a Traditional Efficacy Trial</i>
<i>Me-too drug</i>	<b>x</b>	
<i>Drug expected to have extensive off-label use</i>	<b>x</b>	
<i>First-in-class</i>	<b>x</b>	<b>x</b>
<i>Serious condition with few treatments</i>		<b>x</b>
<i>Drug administered in a controlled setting</i>		<b>x</b>

***Who May be Affected by a Requirement of Inclusion of Pragmatic Features in Trials of Drugs Before Approval***

Pharmaceutical or biotechnology companies The costs of bringing a drug to market were estimated at \$800 million dollars in 2001.(41) This estimate included the estimate for the costs of development of the many drugs that do not make it through the development process. Longer drug development times shorten the period during which a company can earn the returns they need to make investment financially viable.(42) Any regulation, therefore, that delays approval of a drug is costly to the manufacturer. However, early termination of the development of drugs that are not going to ultimately succeed saves money.(43)

Drug promotion, by direct-to-consumer advertisements, journal advertisements, and the use of sales representatives makes up a large proportion of companies expenses. Manufacturers might *benefit* if these costs are averted when the drug is withdrawn early, based on results of a PCT initiated prior to approval.

Manufacturers could be harmed by PCTs if there is identified a large subset of the population for whom the drug is ineffective or harmful. Manufacturers seek approval for marketing based on phase III studies conducted in carefully chosen populations. They are not limited in their marketing to any one population, although they are limited to the indication for which approval was granted. The company’s goal is that the drug will be used widely for the approved-indication, across patient populations. If a PCT demonstrates lack of effectiveness in a subset (perhaps African-Americans or women over 75 years old), this information could encroach upon sales in those subsets. The same things would happen if harms are demonstrated disproportionately in a subset of the population.

The research staff of drug companies may be challenged by these study designs. Certainly, companies are familiar with designing large, even international trials, particularly for phase IV studies. Some of the assumptions underlying the design of PCTs; however, are different –even from phase IV trials – which may require education of the research staff.

Patients and Clinicians Patients and clinicians should almost universally benefit from a requirement that PCTs be done prior to approval, except if it substantially delays access to a drug that is effective with low risk of harms. Emergency access to and compassionate use of unapproved drugs would presumably still be available. However, this is only for life-threatening or serious conditions. First-in-class drugs may be drugs that patients and clinicians are most interested in prompt access to, but these are also the drugs with the most uncertainty regarding harms.

Will clinicians be challenged by how to interpret results of PCTs? Yes, possibly. As described above, clinicians want results that are applicable to an individual patient. Results from a heterogeneous population with variable adherence and assorted co-treatments are challenging to apply to an individual. However, if the drug is demonstrated to be broadly effective and broadly without substantial risks, the clinician can have confidence in using the drug regardless of the unique clinical situation.

Funders Presumably if these trials are done prior to approval, the funders of the trials will largely be the manufacturers. It is conceivable however, that these trials may be of sufficient interest to payers that they may have a role in funding these studies, particularly if it is very likely that the drug will be approved. This may create unique situation in which studies of unapproved drugs are designed and implemented by manufacturers in coordination with payers.

Is it likely that federal agencies would fund these trials? It is conceivable that if it is a novel drug expected to make a substantial contribution to health, that there could be federal support for the PCTs of these unapproved drugs. There will be the potential for this to be challenging if there is perceived favoritism for drugs from one company over another. Presumably if there is a transparent process which an unapproved drug needs to meet to be eligible for federal support (serious conditions, few alternatives).

## VI. Suggested Discussion Topics

Certainly, PCTs and preapproval trials for licensing that incorporate pragmatic features raise interesting questions. We outline here some issues needing discussion. If the goal is to accommodate the needs of different decision makers, the voices of different decision makers should be heard on these topics.

The following set of topic deal with the key issues in study design. These topics will be discussed in the morning breakout group session.

### *Selecting Appropriate Comparators*

- How does one select the appropriate comparator(s) in head-to-head trials and how might selection vary when one takes into consideration different end users (e.g., FDA versus payers)? What are some of the issues related to this and how might they be resolved?
- Are there ways to better classify or specify ‘usual care’ prospectively in studies, without completely jeopardizing this aspect of pragmatic design? What are some options for doing this?
- When is it important to include a placebo comparator? Under what circumstances might this not be important? Under what circumstances might this be acceptable to the FDA?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs? Are there topics related to the design of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

#### *Enhancing Generalizability of Patient Population*

- What are some of the major barriers to reducing inclusion/exclusion criteria for RCTs intended for drug approval?
- Are there certain criteria one might be able to use in order to determine whether a particular exclusion criterion was essential for study validity? What statistical tests might one use to be able to identify nonessential exclusion/inclusion criteria (e.g., elderly, children, specific racial groups)?
- Which groups of patients have been underrepresented in traditional RCTs?
- What are some specific strategies to improve enrollment those groups?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to the design of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

#### *Dealing with Heterogeneity Overall*

- When designing a pragmatic trial, are there statistical or analytic methods in study design or analysis (apart from increasing sample size) that can be used to improve statistical precision?
- What issues do these approaches raise regarding their acceptance in trial design in the licensing phase for a new drug?
- What analytic approaches or tools are available to define more homogeneous subgroups a priori for analysis? What issues are raised with incorporating these analyses into trials used for drug licensing?



- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to the design of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the breakout groups?
- What are these and what are the important issues to explore within each of these topic areas?

#### *Relaxing Intensity of Monitoring and Protocol-driven care*

- In what ways might the intensity of protocol-driven monitoring or care be relaxed that may be more tolerable in Phase III/IIIb pharmaceutical trials (e.g., monitoring patient compliance, monitoring physician compliance to dosage, requiring follow up visits at specific intervals, allowing switching among treatment regimens)?
- Are there mechanisms to loosen protocol driven controls that would not jeopardize being able to assess the safety and efficacy of the drug?
- Are there specific environmental situations where it may be more acceptable to relax protocol-driven monitoring or care (e.g., specific features of the drug, or better knowledge of the disease pathway)? If so what are these?
- Are there design innovations that would allow one to relax protocol-driven care after immediate requirements are met for safety and efficacy, but would still utilize the same patient population and be completed prior to drug launch? What are the benefits and drawbacks of these approaches with respect to meeting the needs of decision makers? What specific questions would these study designs answer of interest to decision makers?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to the design of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the breakout groups?
- What are these and what are the important issues to explore within each of these topic areas?

The following set of topics deal with implementation barriers and potential solutions. These topics will be discussed in the afternoon breakout group session.

#### *Regulatory Barriers*

- What are the major statutory or regulatory barriers to greater use of pragmatic clinical trials in the design phase?
- Do any of these affect specific features of pragmatic trials more than others? If so, which ones and in what ways?

- Are there recent changes in the regulatory environment that might be more conducive to designing pragmatic trials? What are they and how do they improve the acceptance of pragmatic design features?
- What options are there to reduce these regulatory barriers? Are there legislative vehicles in the near term to do this? Are there FDA initiatives that could be used as a vehicle to further lay the groundwork for acceptance of some pragmatic design features?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to administrative or operational issues of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

#### *Expanding Research to the Community Setting*

- What are some of the difficulties one encounters when trying to expand trial settings outside of the traditional academic centers (e.g., quality of data collection, training costs, overall trial costs, recruitment)?
- What are some of the key issues related to these difficulties that are most pertinent to the Phase III/IIIb trials?
- What are some affordable options for ensuring high-quality data are received by community physicians?
- What are ways to help minimize expense per enrolled patient when one reaches out to physicians who may have a smaller patient base than the typical academic setting?
- Are there salient issues specific to conducting research in a managed care population that are important to address in pragmatic design? What are they?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to administrative or operational issues of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

#### *Reducing the Cost and Improving the Efficiency of Pragmatic Trials*

- What are some approaches that could reduce the cost and improve the efficiency of pragmatic trials (discuss separately for trial design innovations, modes of data collection, approaches to patient recruitment, or any other categories you feel are important)?
- Are any of these more amenable to use in the licensing phase of a drug? Which ones and why?

- What are the major barriers to using these approaches? How might those barriers be circumvented or mitigated?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to operational or administrative issues of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

### *Dealing with Cultural Acceptance of Pragmatic Trials*

- What are some approaches that could mitigate internal company opposition to incorporating more pragmatic design features into trials designed for drug approval?
- What are some approaches that could mitigate regulatory agency trial reviewers' opposition to incorporating more pragmatic design features into trials designed for drug approval?
- Are there some features of pragmatic design that trialists are more likely to accept? Which ones and why? Which ones will they likely not be willing to adopt, and why? Are the features they may be more willing to adopt the ones that health payors feel are most important?
- How does one protect against 'design regression' (where one starts out with the idea of designing a pragmatic trial, but it evolves into a more explanatory trial design)?
- Under this specific topic, what subtopics should be included in an Effectiveness Guidance Document that lays out the principles for designing more pragmatic trials in the licensing phase for drugs?
- Are there topics related to operational or administrative issues of pragmatic trials that you feel are important to include in the Effectiveness Guidance Document, but were not a focus of one of the break out groups?
- What are these and what are the important issues to explore within each of these topic areas?

In summary, PCT's or trials which incorporate some of the pragmatic features of PCTs may be useful additions to the traditional Phase III trials conducted to evaluate new drugs in the pre-approval setting. Given that the goal is to generate data that is useful to both regulators and decision makers (clinicians, patients, payors, manufacturers, funders) who need this information after approval, innovative methodologies need to be developed to maintain the internal validity of these trials, assure generalizability, and make sure that the questions being asked and the outcomes being measured are the right ones.

## Reference List

1. Center for Drug Evaluation and Research. Time Line: Chronology of Drug Regulation in the United States. <http://www.fda.gov/cder/about/history/time1.htm> . 1-1-1998. 3-31-2009.
2. Food and Drug Administration Amendments Act of 2007'. [http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110\\_cong\\_public\\_laws&docid=f:publ085.110](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ085.110) H.R. 3580. 9-27-2007.
3. CFR Title 21 - Food and Drugs: Parts 1 to 1499. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm> . 4-1-2008. 4-12-2009.
4. Code of Federal Regulations. PART 314 -- APPLICATIONS FOR FDA APPROVAL TO MARKET A NEW DRUG. 21CFR314.126. 3-4-2002.
5. Office of Training and Communications. Center for Drug Evaluation and Research. List of Guidance Documents. <http://www.fda.gov/cder/guidance/index.htm> . 2-4-2008. 4-9-2009.
6. Food and Drug Administration. Guidance for Industry E 10 Choice of Control Group and Related Issues in Clinical Trials. <http://www.fda.gov/CDER/GUIDANCE/4155fnl.htm> . 2001.
7. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis 1967;20:637-48.
8. Berlivet L. Between expertise and biomedicine: public health research in France after the Second World War. Med Hist 2008;52:471-92.
9. Armitage P. Controversy in the interpretation of clinical trials. Ann Neurol 1979;5:601-2.
10. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. N Engl J Med 1979;301:1410-2.
11. Feinstein AR. An additional basic science for clinical medicine: II. The limitations of randomized trials. Ann Intern Med 1983;99:544-50.

12. Charlton BG. Mega-trials: methodological issues and clinical implications. *J R Coll Physicians Lond* 1995;29:96-100.
13. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. The GUSTO investigators. *N Engl J Med* 1993;329:673-82.
14. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *Ann N Y Acad Sci* 1993;703:314-40.:314-40.
15. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. *Lancet* 1988;2:349-60.
16. GISSI-2: a factorial randomised trial of alteplase versus streptokinase and heparin versus no heparin among 12,490 patients with acute myocardial infarction. Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico. *Lancet* 1990;336:65-71.
17. Charlton BG. Megatrials are based on a methodological mistake. *Br J Gen Pract* 1996;46:429-31.
18. Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheum Dis* 1999;58 Suppl 1:I82-5.:I82-I85.
19. Gartlehner G, Hansen RA, Nissman D, Lohr KN, Carey TS. A simple and valid tool distinguished efficacy from effectiveness studies. *J Clin Epidemiol* 2006;59:1040-8.
20. Thorpe KE, Zwarenstein M, Oxman AD et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009;62:464-75.
21. Zwarenstein M, Treweek S, Gagnier JJ et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390. doi: 10.1136/bmj.a2390.:a2390.
22. Karanickolas PJ, Montori VM, Devereaux PJ, Schunemann H, Guyatt GH. A new 'mechanistic-practical' framework for designing and interpreting randomized trials. *J Clin Epidemiol* 2009;62:479-84.

23. Karanickolas PJ, Montori VM, Devereaux PJ, Schunemann H, Guyatt GH. The practicalists' response. *J Clin Epidemiol* 2009;62:489-94.
24. Oxman AD, Lombard C, Treweek S, Gagnier JJ, Maclure M, Zwarenstein M. A pragmatic resolution. *J Clin Epidemiol* 2009;62:495-8.
25. Oxman AD, Lombard C, Treweek S, Gagnier JJ, Maclure M, Zwarenstein M. Why we will remain pragmatists: four problems with the impractical mechanistic framework and a better solution. *J Clin Epidemiol* 2009;62:485-8.
26. Maclure M. Explaining pragmatic trials to pragmatic policymakers. *J Clin Epidemiol* 2009;62:476-8.
27. Buse JB, Klonoff DC, Nielsen LL et al. Metabolic effects of two years of exenatide treatment on diabetes, obesity, and hepatic biomarkers in patients with type 2 diabetes: an interim analysis of data from the open-label, uncontrolled extension of three double-blind, placebo-controlled trials. *Clin Ther* 2007;29:139-53.
28. Buse JB, Henry RR, Han J, Kim DD, Fineman MS, Baron AD. Effects of exenatide (exendin-4) on glycemic control over 30 weeks in sulfonylurea-treated patients with type 2 diabetes. *Diabetes Care* 2004;27:2628-35.
29. Kolterman OG, Buse JB, Fineman MS et al. Synthetic exendin-4 (exenatide) significantly reduces postprandial and fasting plasma glucose in subjects with type 2 diabetes. *J Clin Endocrinol Metab* 2003;88:3082-9.
30. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003;290:1624-32.
31. Massie BM, Collins JF, Ammon SE et al. Randomized trial of warfarin, aspirin, and clopidogrel in patients with chronic heart failure: the Warfarin and Antiplatelet Therapy in Chronic Heart Failure (WATCH) trial. *Circulation* 2009;119:1616-24.
32. Fried LF, Duckworth W, Zhang JH et al. Design of combination angiotensin receptor blocker and angiotensin-converting enzyme inhibitor for treatment of diabetic nephropathy (VA NEPHRON-D). *Clin J Am Soc Nephrol* 2009;4:361-8.
33. Bond J, Wilson J, Eccles M et al. Protocol for north of England and Scotland study of tonsillectomy and adeno-tonsillectomy in children (NESSTAC). A pragmatic randomised

controlled trial comparing surgical intervention with conventional medical treatment in children with recurrent sore throats. *BMC Ear Nose Throat Disord* 2006;6:13.:13.

34. Hay EM, Foster NE, Thomas E et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomised trial. *BMJ* 2006;333:995.
35. Kennedy A, Reeves D, Bower P et al. The effectiveness and cost effectiveness of a national lay-led self care support programme for patients with long-term conditions: a pragmatic randomised controlled trial. *J Epidemiol Community Health* 2007;61:254-61.
36. Freemantle N, Blonde L, Bolinder B et al. Real-world trials to answer real-world questions. *Pharmacoeconomics* 2005;23:747-54.
37. Vallve C. [A critical review of the pragmatic clinical trial]. *Med Clin (Barc)* 2003;121:384-8.
38. Davis BR, Cutler JA, Gordon DJ et al. Rationale and design for the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). ALLHAT Research Group. *Am J Hypertens* 1996;9:342-60.
39. Stroup TS, McEvoy JP, Swartz MS et al. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr Bull* 2003;29:15-31.
40. Kroenke K, West SL, Swindle R et al. Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. *JAMA* 2001;286:2947-55.
41. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. *J Health Econ* 2003;22:151-85.
42. DiMasi JA. What's time got to do with it? *Health Aff (Millwood)* 2006;25:1188.
43. DiMasi JA. The value of improving the productivity of the drug development process: faster times and better decisions. *Pharmacoeconomics* 2002;20 Suppl 3:1-10.