Published in final edited form as:

Contemp Clin Trials. 2016 November ; 51: 34–43. doi:10.1016/j.cct.2016.09.006.

Ethical and Clinical Safety Considerations in the Design of an Effectiveness Trial: A Comparison of Buprenorphine versus Naltrexone Treatment for Opioid Dependence

Edward V. Nunes, MD¹, Joshua D Lee, MD, MSc³, Dominic Sisti⁴, Andrea Segal⁵, Arthur Caplan⁶, Marc Fishman⁷, Genie Bailey⁸, Gregory Brigham⁹, Patricia Novo², Sarah Farkas², and John Rotrosen, MD²

¹Columbia University Medical Center; New York State Psychiatric Institute

²NYU School of Medicine, Department of Psychiatry

³NYU School of Medicine, Department of Population Health

⁴Perelman School of Medicine, University of Pennsylvania

⁵Perelman School of Medicine, University of Pennsylvania

⁶NYU School of Medicine, Division of Medical Ethics

⁷Johns Hopkins University School of Medicine, Mountain Manor Treatment Program

⁸Stanley Street Treatment and Resources

⁹ADAPT, Roseburg, OR

Abstract

We examine ethical challenges encountered in the design of an effectiveness trial (CTN-0051; X:BOT), comparing sublingual buprenorphine-naloxone (BUP-NX), an established treatment for opioid dependence, to the newer extended-release injectable naltrexone (XR-NTX). Ethical issues surrounded: 1) Known poor effectiveness of one possible, commonly used Treatment as Usual control condition—detoxification followed by counseling without medication; 2) The role of patients' preferences for treatments, given that treatments were clinically approved and available to the population; 3) Differences between the optimal “usual treatment” clinical settings for different treatments making it challenging to design a fair comparison; 4) Vested interest groups favoring different treatments exerting potential influence on the design process; 5) Potentially vulnerable populations of substance users and prisoners; 6) Potential therapeutic misconception in the implementation of safety procedures; and 7) High cost of a large trial limiting questions that could be addressed. We examine how the design features underlying these ethical issues are characteristic of effectiveness trials, which are often large trials that compare treatments with

Corresponding Author: Edward V. Nunes M.D., Address: 1051 Riverside Drive (unit 51), New York, NY 10032, nunesed@nyspi.columbia.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

varying degrees of existing effectiveness data and familiarity to patients and clinicians, in community-based treatment settings, with minimal exclusion criteria that could involve vulnerable populations. Hence, investigators designing effectiveness trials may wish to remain alert to the possibility of similar ethical issues.

Keywords

ethics; effectiveness; clinical trial; opioid dependence; buprenorphine; naltrexone

Introduction

Effectiveness trials constitute an important step in the translational spectrum of treatment development. Efficacy trials stress internal validity and address whether a new treatment can be shown to work under ideal conditions, often within research centers. Once efficacy is established, effectiveness trials take the next step, emphasizing external validity and addressing how a new treatment will perform under real-world clinical conditions and in comparison to current standards of care in the community.

The National Drug Abuse Treatment Clinical Trials Network (CTN), funded by the National Institute on Drug Abuse (NIDA), was founded in 1999, charged with the mission to conduct randomized, controlled effectiveness trials, in order to help move new efficacious treatments for substance use disorders out of research centers and into widespread clinical use. An Institute of Medicine report [1] had concluded that while a number of new treatments for substance use disorders had been developed through efficacy trials in research settings, these treatments were not being adopted into routine clinical practice, and that community-based effectiveness trials were needed to bridge this gap. The CTN provided a collaborative structure within which researchers could partner with community-based treatment programs and NIDA to conduct such effectiveness trials on new interventions and services for treating substance use disorders.

The first decade of experience in the CTN yielded a number of lessons learned about the design and analysis of effectiveness trials in the addictions field [2] [3] [4] [5] [6]. However, to date there has been relatively little inquiry around bioethical issues involved in the design and implementation of CTN effectiveness trials. In this paper, we review the ethical issues encountered in the design of one CTN study, CTN-0051, “Extended-Release Naltrexone vs. Buprenorphine for Opioid Treatment (X:BOT)”, led by several of us (JR, JL, EN). While the ethical framework for traditional efficacy trials is relatively well worked out, in the design of CTN-0051 we encountered issues that seemed related to the distinct aims and circumstances of a community-based effectiveness trial and comparative effectiveness research. We examine how these ethical issues relate to specific design features of CTN-0051 and the extent to which such design features may reflect general characteristics of community-based effectiveness trials. It is hoped that this may provide a useful framework for the design of future studies.

Methods

The CTN-0051 protocol development team was charged by NIDA to design a trial to evaluate the effectiveness of extended release injectable naltrexone (XR-NTX, Vivitrol®), a sustained-release injectable formulation of the opioid receptor antagonist naltrexone, for treatment of opioid dependence in real world, community-based treatment settings in the U.S. Oral naltrexone had been indicated and available for several decades for treatment of opioid dependence, but its effectiveness in practice had been limited by poor adherence [7] [8] [9] [10]. Long-acting injected or implanted naltrexone has the potential to circumvent problems with adherence to daily pill taking. XR-NTX, which has a month-long duration of action, had received FDA approval for treatment of alcohol dependence in 2006, and for treatment of opioid dependence in 2010, and had seen limited clinical use at the time the present trial launched in 2014. Approval was based on a pivotal trial conducted in Russia among hospitalized, detoxified opioid dependent patients, which demonstrated that 51% of patients on active XR-NTX plus outpatient counseling had what could be considered a good clinical response--remaining in treatment for 6 months (6 monthly injections) with minimal evidence of opioid use--compared to 31% of patients treated with placebo plus outpatient counseling [11] [12]. Other forms of long-acting, implanted naltrexone, had also shown efficacy compared to oral naltrexone, or to placebo [13] [14][15].

The final design of CTN-0051 is a randomized, comparative effectiveness trial, in which individuals with opioid dependence, seeking treatment at community-based, short-term inpatient/residential addiction treatment programs, were offered the opportunity to consent to being randomly assigned to 6-months of treatment with either monthly injections of XR-NTX, or daily sublingual buprenorphine-naloxone (BUP-NX). Buprenorphine is a high-affinity, mu-opioid receptor partial agonist and kappa receptor agonist with well-established efficacy and effectiveness for the long-term, maintenance treatment of opioid dependence [16]. BUP-NX had been FDA approved for the long-term treatment of opioid dependence for 10 years at the time CTN-0051 was designed. The final design and its rationale is described in detail elsewhere [17].

Early in project development of CTN-0051, it became clear that several design issues were associated with bioethical concerns, some of them unfamiliar. What followed was a yearlong process of design reconsiderations with input from investigators, the treatment community, the Sponsor (NIDA), the manufacturer of XR-NTX (Alkermes, Inc.), and an independent Protocol Review Board. In the process, the lead team engaged two bioethicists (DS and AC) to join the development team. In what follows we describe the design features with associated ethical issues encountered during the design of CTN-0051 and how these may relate to the general features of effectiveness trials.

Ethical Issues Encountered in the Design Process

1. Treatment as Usual (TAU) Control Condition

Effectiveness trials often seek to compare a new treatment such as XR-NTX to existing treatments in use in the community, referred to as 'treatment as usual' (TAU) [2]. Therein lay the first and most difficult ethical rub in the design of CTN-0051, namely the choice of

TAU control condition(s) that were scientifically appropriate, and ethically acceptable. Usual treatments, having already been in widespread clinical use, may already have substantial evidence characterizing their effectiveness and safety, or lack thereof. For CTN-0051, one of the potential control conditions in common use in community-based practice for opioid dependence, namely detoxification followed by counseling without maintenance medication, had evidence of being inferior and potentially dangerous.

Usual treatments for opioid dependence in the U.S. at the time included opioid agonist therapies, methadone maintenance or buprenorphine maintenance [18], which had substantial evidence from clinical trials for superiority over placebo with counseling alone [19] [16]. Such trials typically showed rates of good clinical response retained throughout a 3 to 6 month course of treatment, predominant abstinence from opioids) of at least 40% to 50%. Observational studies also suggest agonist maintenance therapy protects against overdose death [20]. These could be considered gold standard treatments, and a trial comparing the effectiveness of the new treatment, XR-NTX, to methadone or buprenorphine maintenance, was straightforward from an ethical perspective. To date, no such comparisons of a long acting injected or implanted naltrexone formulation, compared to opioid agonist therapy, have been performed. Such a trial would ask how the new treatment, XR-NTX, compares to the gold standard.

However, another usual mode of treatment in the U.S., and many other parts of the world, consists of hospitalization for detoxification from opioids, followed by psychosocial, counseling-based treatment without medication, either on an outpatient basis or as part of long term residential treatment. There were significant concerns about high risk of relapse [21] [22], and associated risk of overdose and overdose death after detoxification and discharge to outpatient status. Large observational studies have shown a spike in relapse deaths after discharge from controlled settings like prison without a maintenance medication [23]. The mechanism of this is well understood, namely that detoxification reverses tolerance, and patients are then more susceptible to the powerful respiratory depressant effects of opioids, at doses that would previously have been tolerated. Yet for numerous reasons this approach remains common in the U.S. [24] [25]. Strong traditions or beliefs among both clinicians and patients favor a “drug-free” treatment, and certainly, some patients succeed in achieving long term recovery from opioid dependence with this route, although little is known about how to identify such good prognosis patients ahead of time. In particular, there is a tradition of considering inpatient rehabilitation treatments definitive, perhaps almost curative, by a process of psychological and spiritual epiphany. Significant stigma and lack of acceptability surround agonist treatments [26] [27]. Agonist treatment is actually illegal in Russia, where the pivotal trial of XR-NTX was conducted [11]. Many parts of the U.S. simply lack sufficient methadone maintenance programs or physicians certified to prescribe buprenorphine for treatment of opioid dependence, leaving detoxification followed by counseling based treatment without medication as the only option.

Thus, the CTN-0051 design team faced the difficult decision whether or not to include, as a third arm, a TAU control condition consisting of inpatient detoxification, followed by counseling-based treatment without medication. To be considered ethically acceptable, a

trial needs to address a question that is scientifically and clinically important and about which there is equipoise (uncertainty about the answer), risks need to be reasonable, and adequate protections in place [28]. For inpatient detoxification followed by counseling with no medication, there was, arguably, both absence of equipoise and serious risk. Despite this, a compelling argument in support of this TAU lay in the very fact that detoxification, followed by counseling without medication, remained a predominant standard of care in many if not most communities in the United States. Thus, there was a desire for a definitive trial to demonstrate the inferiority of this approach, especially within the very delivery setting context of inpatient rehabilitation treatment where it is commonly attempted, in hopes that this would help stimulate the widespread implementation of effective medications. Voices in favor of this argument included members of the Protocol Development Team who were running community-based treatment programs. In their communities medication maintenance was not available to many or most patients due to absence of third party payment and/or lack of practitioners. They argued that a definitive demonstration of the inferiority of counseling alone without medication maintenance might help influence policy and practice.

Another argument was that some patients do achieve long-term abstinence and recovery with detoxification followed by counseling alone, even if it is not the majority. For example, in the Russian pivotal trial 30% of patients on placebo (who received inpatient detoxification alone followed by counseling without medication treatment) achieved stable abstinence over the 6 month trial, albeit inferior to the 50% randomized to XR-NTX with, stable abstinence [11] [12]. Other studies, such as CTN-0030 [21] showed more than 90% relapse after discontinuation of buprenorphine despite the presence of counseling, suggesting a much lower rate of successful outcome without medication. A long term follow-up on this sample showed slightly higher rates of abstinence without medication, compared to the end of the acute trial, but engagement in agonist medication treatment was still associated with abstinence [29]. Studies following patients discharged for inpatient detoxification or residential treatment programs have similarly shown high relapse rates, but that a proportion of patients reduce their drug use or even sustain abstinence [30] [31] [32]. It was of interest to try to understand predictors or clinical characteristics of which patients might be expected to succeed without medication. However, such predictive or moderator analyses require large sample sizes, and are usually considered exploratory, with concerns about both replicability (Type I error) and power (Type II error), hence not a strong rationale for a major clinical trial.

Ultimately, for CTN-0051 it was decided to implement the 2-Arm comparison of active medications, XR-NTX versus buprenorphine over a 6-month trial—i.e. random assignment to 6 monthly injections of XR-NTX vs 6 months of daily sublingual buprenorphine-naloxone (BUP-NX). This design was felt to address the most important effectiveness question, comparing the new treatment (XR-NTX) to a current evidence-based standard (buprenorphine) in community-based treatment settings, with the most favorable risk/benefit ratio. A proposed design involving a third arm of detoxification followed by counseling without medication was eliminated because of the concerns about safety and absence of equipoise. A recently completed randomized trial that compared treatment as usual without medication to XR- NTX, supports this decision in that rates of both relapse to regular opioid

use and overdose events were lower among patients assigned to XR-NTX [33]. The CTN-0051 design team was aware of the existence of this trial, but its findings were not available until long after CTN-0051 was designed and launched.

When designing comparative effectiveness trials, the existing evidence on efficacy and safety of TAU control conditions needs to be evaluated. When there is evidence that a candidate TAU is ineffective, or risky, absence of equipoise and risk/benefit ratio need to be considered. The risk/benefit analysis will be complicated to the extent that a given field is divided on what constitutes good or bad TAU, and how to interpret the evidence (see also below: Section 4. Vested Interests). In such instances, the potential benefit of a definitive randomized trial with an arguably inferior TAU control condition needs to be weighed against the potential harms to patient-participants associated with that control condition.

2. Patient Preferences and Choice Designs

As noted above, detoxification followed by counseling-based treatment without medication is considered a preferred treatment approach by many clinicians and patients, despite concerns and evidence on relapse risk and overdose. Others in the field advocate medications, and the clinical trials that support them. Patients and clinicians are likely to have beliefs and preferences about treatments that are already in use, and that are going to be compared in an effectiveness trial. These preferences may or may not be based on evidence from clinical trials, but may be rather based on personal belief systems, or personal experience with treatments. In clinical efficacy trials, treatments are typically assigned randomly. However, in real world clinical care, when treatment options are available, choices are made by patients and their clinicians. This raises the scientific question of whether free choice and preference may contribute to the effectiveness of treatments for opioid dependence. This question could be addressed by a design in which patients are randomly assigned to either 1) choose among available treatments or 2) be randomly assigned to available treatments. Such a 4-cell (2 by 2) design would likely require at least twice the sample size of a straight, 2-cell random assignment trial.

Another design option that was considered during protocol development was to forgo randomization and instead set up an observational trial or registry study in which opioid dependent patients would choose from among available treatment options. An observational study such as this addresses somewhat different questions, compared to a randomized trial, such as characteristics of patients choosing different treatments and predictors of outcome within each treatment. Also, a practical issue was that XR-NTX, though backed by an efficacy trial [11] and FDA approval, was new enough that it was not yet in widespread clinical use, and hence less likely to be chosen naturalistically for that reason.

Allowing patients to choose a medication-free treatment, or a medication-based treatment, after due informed consent, does model what would happen in a community-based treatment setting where all options were available. Thus, from an ethical perspective including choice in the design arguably may produce the benefit of a better, more generalizable scientific yield. Allowing choice also respects patients' autonomy. Allowing choice was viewed as a potential solution to the ethical concerns about efficacy and safety of medication-free treatment that were noted above. Patients would receive a medication-free TAU, thought to

be an inferior treatment, only if that was their choice. Choice would weaken the internal validity of the study, since without randomization, and despite statistical methods like propensity score matching, confounds cannot be ruled out. Another argument leveled against choice was that it may do the field and the patients a disservice to study choices, if those choices are irrational—i.e. not based on data. Certainly, there is an imperative to generate good data on patient-treatment matching, but the question is: does one need random assignment for the best data on matching, or does preference-based choice actually contribute to effectiveness of a matching strategy?

Another concern is that, in a fully randomized trial, patients who are not assigned to their preferred choice of treatment may refuse participation and drop out, and such refusal, particularly if there is more dropout in one condition than another, could bias the results of the trial and threaten its internal validity. Such a phenomenon was observed in a CTN trial of buprenorphine detoxification, where the control arm, a clonidine-based detoxification, had a poor reputation and excess early dropout from that arm was observed [34].

Ultimately, for CTN-0051 it was decided to preserve random assignment, not allowing choice of treatments, but to elicit patients' stated preferences prior to randomization and study this as a covariate, lending itself to scientific study of preference, if not choice per se. In the design of effectiveness trials, potential preferences for treatments being compared need to be considered in terms of how those preferences might impact the scientific validity of the trial and its risk/benefit ratio.

3. Inherent Differences between Treatments: What is a Fair Comparison?

Effectiveness trials may compare distinct multi-component treatment regimens, in contrast to efficacy trials, which typically compare treatment elements (e.g. active medication versus placebo) within an otherwise constant background regimen or so called platform treatment. For CTN-0051, key differences between the treatment regimens or platforms that accompany antagonist (XR-NTX) versus agonist (buprenorphine) treatments included the substantial “induction hurdle” required for XR-NTX [35], compared to the ease of induction onto buprenorphine, and differences in failure modes between the two treatments. This led to two key design issues which were contentious—namely the clinical setting or ‘platform’ for the trial and timing of randomization, and the choice of primary outcome measure. These have ethical implications, related to the validity of the comparison between treatments, and appropriate clinical management and patient safety.

Choice of Primary Outcome—Naltrexone (XR-NTX) and buprenorphine have different mechanisms (partial agonist versus antagonist) and somewhat different patterns of benefit and of failure. With naltrexone (XR-NTX) the outcome is rather binary in nature—while patients adhere they do well, rarely using opioids because the effects of opioids are completely blocked. But, if a patient stops naltrexone, relapses to opioid use, and physical dependence is re-established, then that treatment episode is effectively over, in that another detoxification/induction is needed in order to resume naltrexone because giving naltrexone to an individual currently physically dependent on opioids would precipitate severe opioid withdrawal. Thus, for treatment with XR-NTX, a discrete event of relapse is, arguably, the

most meaningful clinical outcome, and time to relapse the most appropriate outcome analysis. With buprenorphine, on the other hand, a patient may drop off the medication temporarily, resume illicit opioids, and then easily resume medication within the same treatment episode. Or a patient may substantially reduce use of opioids, reflecting a clinically significant improvement, but not achieve abstinence. Thus, for buprenorphine, an outcome of dimensional frequency of opioid use over time, or sustained abstinence at the end of a trial period (e.g. at 3 or 6 months post randomization) might be more appropriate.

For an effectiveness trial, the primary outcome should be the outcome that is most clinically meaningful. The quandary here was that the most meaningful outcome might be different for the treatments being compared. Choosing the outcome that is best for one treatment might bias the trial in favor of that treatment. For example, choosing a discrete event of relapse as the outcome might bias in favor of naltrexone, to the extent that patients on buprenorphine might be called relapsed prematurely and their trial ended, when more time on treatment might ultimately yield a good outcome. On the other hand, choosing sustained distal abstinence at the end of the trial as the outcome might bias in favor of buprenorphine, to the extent that patients on naltrexone would have difficulty getting back onto the medication in the event of intermediate difficulties while patients on buprenorphine could more easily “bounce back” to resume the medication and stabilize over time. Ultimately, an event of relapse was chosen as the primary outcome, but defined to require enough sustained opioid use to be indicative of treatment failure for buprenorphine as well as naltrexone. Another way to view this is that successful treatment with XR-NTX and buprenorphine looks similar, consisting of sustained adherence plus abstinence. Relapse as the main outcome sets a high bar, suggesting that any significant lapse from sustained abstinence should trigger a new treatment plan.

Timing of Randomization in the Context of Inpatient Treatment—After an opioid dependent patient is admitted to an inpatient setting, buprenorphine can be started fairly easily at any point from the first day of admission onward and has the beneficial effect of relieving acute opioid withdrawal. In fact, buprenorphine does not usually even require hospitalization, since it can usually be initiated on an outpatient basis. In contrast, naltrexone requires a completed detoxification and washout over 5 to 7 days, typically accompanied by a certain amount of withdrawal discomfort, before the naltrexone injection (XR-NTX) can be given safely. Inpatient treatment is the most reliable way to get a patient through this “induction hurdle”. Even so, a certain proportion of patients, 30% or more based on prior experience, dropout from inpatient before starting XR-NTX, likely subsequently relapsing [36].

This creates a conundrum around the setting of randomization and its timing. On the one hand, randomization could occur on an outpatient basis, with inpatient only afterwards for those randomized to XR-NTX, since BUP could be initiated on an outpatient basis. Such a trial might favor buprenorphine, since it would be logistically easier for patients to initiate BUP than XR-NTX. Another proposal was to require all prospective participants to fully detoxify and washout (i.e. be ready to take naltrexone) before being randomized to either naltrexone or buprenorphine. However, this would arguably favor naltrexone. It was not consistent with good, or most common clinical practice for buprenorphine initiation and

would put some patients, destined to be treated with buprenorphine, at unnecessary risk for dropout and relapse. The Protocol Review Board in its review of CTN-0051 design options emphasized this concern about “tilting the playing field” in favor of XR-NTX. Ultimately, it was decided to base the trial on a platform of inpatient/residential treatment, but to allow randomization to occur at any time after admission to the inpatient setting. Thus, the platform for the trial is the inpatient setting. It was argued that a large proportion of patients seeking treatment for opioid dependence begin with inpatient treatment, so the inpatient platform would have good generalizability. Within that platform each medication would be initiated as it would be in common clinical practice, hence not “tilting the playing field” toward one treatment or the other within the inpatient platform.

4. Vested Interests

A further factor complicating the design process was the presence of potential vested interests. The manufacturers of the two medications had a potential interest in the design placing their respective medication in the best light. There were also factions among clinicians in the field, as discussed above. Attitudes and beliefs among clinicians and other stakeholders favoring an abstinence-based treatment without medication have been documented and suggested to be slowing the adoption of evidence-based medications for opioid use disorder [37] [38]. Strong advocates of opioid agonist therapy (methadone or buprenorphine maintenance) had attacked previous trials of extended release naltrexone (XR-NTX) based on a strong belief in its inferiority [39]. Other groups sometimes opposed to opioid agonist therapy include criminal justice professionals dismayed by poor outcomes among some patients and diversion of methadone and buprenorphine onto the black market, and proponents of ‘drug-free’ abstinence as the best path to recovery. From these points of view, naltrexone--a drug that blocks effects of opioids while having no drug-like effects of its own, and no ‘street value’ or potential for diversion—may seem acceptable. As mentioned earlier, some members of the design team wanted to see a counseling only/no medication TAU comparison condition, despite the concerns about equipoise and safety, in hopes of influencing attitudes and policies that were hampering the implementation of medications for treatment of opioid dependence in their communities.

In the case of the present trial, none of these groups attempted to exert undue influence. But the design team was aware of their presence, and it increased the pressure to design a trial that would be a fair test. Since effectiveness trials typically compare treatments that are already in use, and such treatments may have proponents with vested interests, be they financial, ideological, or pragmatic, design teams need to be mindful of how such interests might sway the design of a trial.

5. The Generalizable Sample and Participant Vulnerability

Effectiveness trials seek to recruit samples that are maximally representative of the relevant patient population. They do this by recruiting and treating patients within typical community-based treatment settings, and by making eligibility criteria as open as possible, with only those exclusionary criteria that reflect clear contra-indications to the treatments being studied. As a result, effectiveness trials may recruit participants who may be considered to be members of vulnerable populations. In the case of CTN-0051, the

participants had opioid dependence, which is often accompanied by other substance use and/or mental disorders, thus raising concerns about capacity and informed consent. Further, many opioid dependent individuals are under various levels of criminal justice supervision (e.g. parole, probation, recommended treatment, court-mandated treatment, or arrested and incarcerated), which raised ethical concerns around prisoner participation in research.

Capacity and Informed Consent—Adequate informed consent is a critical component of ethical research. Instantiating the principle of respect for persons, the informed consent process requires three basic elements: information, comprehension, and voluntariness. In clinical research, capacitated volunteers should understand the purpose, and potential risks and benefits of the study in order to make an informed choice about whether to enroll. The issue of whether drug dependent patients have diminished capacity has been raised [28]. Even if not acutely intoxicated, drug craving, or changes in motivation and decision making associated with addiction might impair capacity [40] [41]. In CTN-0051 this concern was addressed through a thorough informed consent process. Informed consent occurred after patients had been admitted to an inpatient unit, assuring at least that potential participants were not acutely intoxicated during consent. The process could occur over several days and required prospective participants to pass a detailed quiz demonstrating understanding of the study.

A more subtle issue with informed consent in effectiveness trials may be the preconceived notions of patients and their clinicians (some already alluded to above), given that the treatments compared in effectiveness trials are typically already in clinical use. In the addiction field there is street lore about treatments, some of it based on no evidence, or contradicted by existing evidence (e.g. “methadone rots your bones”), which might influence patients' decisions about trial participation. The informed consent document and process needs to be designed in an effort to provide balanced information, mindful of strongly held beliefs that may abound.

Prisoner Participants—In designing CTN-0051 the investigators wanted to be able to follow participants who may become incarcerated during their study participation, and to include participants who were under various forms of legal mandate to participate in treatment for their substance use problems. With the advent of drug treatment courts and other diversion programs the number of persons with opioid addiction living in the community under court supervision has increased steadily. For example, in 2008, of 116,000 criminal offenders served by a drug court, between 9% and 19% had heroin or prescription drugs as the primary drug of abuse [42] [43]. Hence, in order to achieve a generalizable sample, there was a methodological imperative to be able to include such individuals.

Prisoners present a special problem with respect to voluntariness and informed consent, given the inherently coercive nature of the criminal justice system. CTN-0051 did not allow recruitment of individuals currently incarcerated within prisons or jails. However, authorization was sought from the Office for Human Subjects Protections (OHRP) (hss.gov/ohrp) to interview study participants who might become incarcerated during study treatment (up through week 24) or at time of follow up (week 28 and week 36 post randomization).

Further, CTN-0051 was recruiting exclusively within residential treatment programs, and a certain percentage were expected to be involved with the criminal justice system, in some cases for “court mandated” treatment, which under 45 CFR 46.303(c) might meet the definition of “prisoner” [44]. The CFR definition of “prisoner” includes individuals “...detained in other facilities [other than penal institutions] by virtue of statutes or commitment procedures which provide alternatives to criminal prosecution or incarceration in a penal institution...”. It is open to interpretation whether a patient, although court mandated, is actually “detained”, given that: 1) Courts usually give such individuals a choice between a treatment program and incarceration; and 2) The inpatient and residential treatment programs in CTN-0051, typical of such programs nationwide, are not locked units, and patients are free to leave if they so choose, though they would face consequences from the court for leaving. Nonetheless, as court direction, court mandates and alternative sentencing vary substantially from state-to-state and district-to-district, we sought and received OHRP approval to include (both at intake and at follow-up) individuals who might meet the definition of “prisoner”. This is consistent with a recent Institute of Medicine report, which suggested that the definition of prisoner, for research purposes, should be broadened beyond physical detention, to include individuals under various forms of criminal justice supervision in the community [45]. Approval was based on the article from the regulations (45 CFR 46.306 (a) (2) (iv)) that more-than-minimal-risk research with prisoners may be acceptable if it constitutes: “...research on practices, both innovative and accepted, which have the intent and reasonable probability of improving the health or well-being of the subject.”

There are potentially far-reaching ramifications in the determination that persons on court recommended or mandated inpatient or outpatient treatment are to be considered prisoners for research purposes, given how prevalent court supervision is among substance dependent patients. Since the definition in the regulations of “prisoner” is open to interpretation, an IRB may determine that the definition of “prisoner” is not met, depending on the nature of the subject population, and whether or not the population is considered “detained”. Or, the IRB and study leadership may seek OHRP approval (ultimately approval of the Secretary), as was done for CTN-0051.

6. Risk-Benefit Ratio, Participant Safety, and Therapeutic Misconception

Ensuring participant safety and minimizing risk is a central ethical consideration across all Human Subjects research. Effectiveness trials may focus on disorders and/or treatments with significant risks. In the case of CTN-0051, opioid dependence carries a significant risk of death from drug overdose among other risks. As noted in Section 1, above, one of the possible Treatment as Usual control groups was eliminated due largely to concerns about risk.

Another important dimension of safety in clinical trials is the management of patients with poor outcome or significant adverse events. This involves appropriate criteria for removing patients from the trial, as well as procedures for clinical management or “rescue” after removal from the trial. Effectiveness trials typically take place within community-based treatment settings, where patient-participants will be receiving a combination of usual clinical management, plus treatment conditions randomly assigned per protocol. Further, one

or both of the treatments being compared are typically already in common clinical use. This can lend itself to the “therapeutic misconception” [46] [47], where distinction is blurred between treatment that is delivered as part of usual care and for the benefit of the patient, and treatment delivered for the purposes of the research study. Therapeutic misconception is an important issue in regard to informed consent for research. During a trial it could also lead clinicians or patients to persist with a study-assigned treatment, even if the outcome is not good, and an alternative treatment might be pursued in routine clinical practice. In CTN-0051 these issues were addressed in part by setting a primary endpoint as relapse (see also Section 3, above), such that patients only remain in the trial if they are having a good clinical response (sustained abstinence), and are considered relapsers and exited from the trial if there is sustained opioid use, either daily for 7 consecutive days or at least once-a-week for four consecutive weeks, as determined by urine toxicology and self-reports. Patients who relapse revert to usual clinical care in the community-based treatments system.

Once a patient in an effectiveness trial develops poor clinical outcome or an adverse event that triggers exit criteria, it is important to have rescue procedures for clinical management after study treatment is ended. In effectiveness trials this may be facilitated since community-based treatment settings in which such trials are situated would likely already have standard clinical procedures for handling patients with poor outcome. In CTN-0051, for patients exited due to relapse or an adverse event, study personnel and clinical personnel at the participating clinical treatment programs (CTP) are responsible to maximize safety through clinical examination and appropriate procedures, either implementing further treatment at the CTP, or referrals to other settings, if necessary, for continued treatment.

Rescue procedures can also be studied systematically, through the use of adaptive designs such as the Sequential Multiple Assignment Randomized Trial (SMART) design [48] [49], where patients who trigger a study endpoint such as relapse, are re-randomized to alternative next-step treatments. The SMART design goes beyond a simple head to head comparison of treatments in a standard comparative effectiveness trial, providing data upon which to build adaptive treatment algorithms, addressing questions on what is the best next step if a given treatment succeeds or fails. It is argued that real world clinical treatment consists of sequences of treatment maneuvers, rather than discrete treatment episodes, so that the scientific yield of the SMART design is thus greater. It also provides patient-participants with potential benefits in the form of built in rescue procedures and continued study-based treatment. Again, care would be needed to distinguish study-based treatment from treatment that might be delivered outside of the treatment study purely for the benefit of the patient. A SMART design was considered for CTN-0051, but ultimately abandoned due to concerns about sample size, power, and cost.

Since patients who relapse are offered rescue treatment or referral by the community-based treatment programs, according to clinical judgment, this does afford the opportunity to study the outcome of patients depending on the treatment offered and implemented after a relapse naturalistically, if not systematically as in the SMART design. In CTN-0051 this data is collected from all participants 6 months after randomization (corresponding to the time at which study treatment would end), as well as at 7 months and 9 months after randomization. An effort is made to locate all randomized patients, including those relapsed and/or dropped

out, at these time-points to measure their clinical status and document what treatment they had received and their outcomes under naturalistic clinical treatment after completion of the 6 month trial. A key safety issue is that the risk of relapse, overdose, and overdose death likely increase if medication treatment for opioid dependence (either methadone, buprenorphine, or naltrexone) is discontinued. Clinical teams were encouraged to arrange further medication treatment for patients after relapse or after the end of the study at month 6. However, this would not always be possible, due to limitations in third party coverage or availability of services in the community, or patients choosing not to continue treatment.

Large scale effectiveness trials are typically monitored by an independent Data Safety and Monitoring Board (DSMB) and/or other independent safety monitor, to review significant or serious adverse events and outcomes as the trial progresses. In this way a trial design might be modified to enhance safety procedures, or even halted early, if significant safety or efficacy concerns emerge. CTN-0051 employed a DSMB, as well as quality assurance and safety monitoring by a contract research organization (EMMES, Inc), both reporting to the Sponsor, NIDA.

Another important safety challenge in addiction research, and in clinical practice, is that patients who discontinue medication and/or relapse often do not return for treatment and cannot be reached for follow-up. They drop out of treatment entirely. This scenario is far from ideal and prevents “rescue” interventions. In an effort to address this, CTN-0051 put into place extensive procedures for tracking and locating dropouts, based on procedures that achieved 90% successful follow-up rates in a prior trial [50]. This is important both scientifically, to ascertain outcome as completely as possible with minimal missing data, and for patient safety, to make the best effort to stay in touch with patients and encourage ongoing treatment participation. For a study of the treatment of opioid dependence, such as CTN-0051, the stakes are high since death by drug overdose is a not uncommon outcome of opioid dependence. It would be expected that the risk of overdose might go up among patients who discontinue medication either during or after completion of the trial.

7. Study Cost, Feasibility, and Ideal Evidence

There is always a tension between financial constraints and the capacity to collect the quantity or quality of data researchers might desire in order to answer their hypotheses. But, even in an era of cost-containment, it is not ethical to utilize methods that cannot answer the question being addressed in that such research exposes subjects to risk and inconvenience for no purpose. Nor is it ethical to accept tradeoffs in the collection of data such that significant degrees of uncertainty remain about research findings. In such cases, subject safety may be compromised or interventions may be introduced into practice that turn out to be ineffective, dangerous or both. Research design including sample size and mode of inquiry must be able to satisfy peer expectations about the quality of research, the degree of confidence that can be accorded results and the ability to build on findings in future inquiries.

In the case of CTN-0051 this tension was particularly apparent insofar as budget considerations precluded extending the treatment duration to study rescue approaches, such as the SMART design, reviewed above, that re-randomizes relapsed patients to subsequent

treatment options. The additional study-driven treatment would have increased costs, although it would address the important question of what is the best next treatment to try if a given initial treatment fails. It would also have improved safety, in that it would have offered further medication treatment to patients after a relapse. Under naturalistic clinical treatment after relapse, medication treatment may, again, not be available to some patients, due to limitations in their third party coverage, or the availability of such treatment in their communities. Treatment provided and financed by the study was limited to only that provided up to the primary endpoint (relapse or completion of the 6-month trial).

Budget considerations also limited the length of the follow-up period after completion of the 6 month trial to follow-ups at 1 and 3 months after the month 6 study endpoint. How long medication treatment needs to be continue, before it can be safely terminated with little risk of relapse or overdose, and predictors of which patients would be able to safely discontinue, are important clinical questions. Many patients want to limit their exposure to medications, but available evidence suggests there is a significant risk of relapse, even after 6 months or more of treatment [33] [51].

For the SMART design there was also the concern that the sample sizes for the secondary randomizations would be too low, yielding insufficient power to derive valid conclusions. To adequately power the secondary randomizations would have required substantially increasing the sample size entering the initial randomization to BUP versus XR-NTX.

In summary, budget and cost considerations limited both the scientific questions that could be addressed and the extent of treatment that could be offered by the study. The basic question, “what is good enough evidence?” is an ethical one; it becomes more salient as government agencies and sponsors attempt to carefully leverage limited resources. Emerging “pragmatic” trials, built on electronic health record and registry platforms, may be able to more cost-effectively address these longitudinal adaptive treatment questions critical for chronic disease management.

Discussion

The design considerations for CTN-0051, described above, serve as an example of how ethical, scientific, and economic factors compound to create challenging ethical questions for the design and conduct of effectiveness research. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research [52] promulgated the ethical principles of respect for persons, beneficence, and justice to guide human subjects' research. Emanuel, et al. [53] argue these principles and others, derived from several codes of medical ethics, entail at least seven requirements that make an experiment involving human subjects, and specifically a clinical trial, ethically acceptable (Table 1). Ethical issues encountered in the development of CTN-0051 are summarized in Table 2 and included: 1) Concerns around poor effectiveness (absence of equipoise) and risk associated with one of the possible Treatment as Usual control conditions; 2) The appropriate role of patients' preferences for treatments in the design; 3) Differences between the optimal platforms for different treatments that made it challenging to design a fair comparison; 4) Vested interest groups favoring different treatments exerting potential influence on the design process; 5)

Potentially vulnerable populations of substance users (issue of capacity to consent) and prisoners (issues of voluntariness and coercion); 6) Safety procedures to manage risks and potential therapeutic misconception between usual treatment and research treatment; and 7) overall design cost placing limitations on questions that can be addressed.

Effectiveness trials compare treatments already in widespread clinical use in the community (comparative effectiveness), or some new treatment compared to a usual treatment. Treatments already in widespread use are likely to have substantial evidence already in place regarding their effectiveness and safety. The most difficult ethical problem in the design of CTN-0051 surrounded a potential Treatment as Usual (TAU) control condition—detoxification followed by counseling without medication—wherein substantial evidence pointed to its lack of effectiveness, and risk of death from drug overdose. Concern about the safety and effectiveness of comparison conditions was central to the ethical controversy over the ethics of the Surfactant, Positive Pressure, and Oxygenation Randomized Trial (SUPPORT) [54], where premature infants were randomly assigned to one of two different target levels of blood oxygen saturation, 85% versus 95%. The stakes were high since death, or severe lifetime disability were common outcomes of this syndrome. One line of criticism held that the optimal management of blood oxygen saturation was already known. Another held that standard treatment would involve clinicians titrating the oxygen level within the 85% to 95% range, but the trial did not include such a group. In the design of effectiveness trials, existing evidence on widely used treatments needs to be carefully weighed to insure that there is uncertainty or equipoise surrounding the effectiveness and safety of treatments to be compared, including treatment as usual (TAU) control conditions, that inferior treatments with greater risks are not studied, and that the right treatments are compared to address the most important clinical questions.

Other challenges encountered in the design of CTN-0051 also involved scientific value and validity. Widely available treatments are likely to be well known to patients. That patients may have preferences raises a scientific question of whether being able to choose one's preferred treatment contributes to its effectiveness. Choice also invokes the ethical principle of respect for patients. An observational study, where patients would choose treatments, forgoes the key strength of randomization in protecting against selection bias, although the findings would generalize only to those who agree to be randomized. A simple solution within a randomized trial is to measure patients' preference prior to randomization to treatment, in order to examine preference as a moderator of outcome, comparing outcomes of patients matched to their preferred treatment to those mismatched. This approach was chosen for CTN-0051. A stronger experimental design would first randomize patients to either choose their treatment, or to be randomized to treatments. The complexity, larger sample size, and cost of such a trial ultimately discouraged its use.

Treatments to be studied in effectiveness trials may be best suited to different clinical settings or platforms. Trials comparing surgical versus medical interventions are possible examples [55] [56]. For CTN-0051 naltrexone (XR-NTX) required an approximately week long inpatient detoxification at the outset, whereas buprenorphine (BUP) was easily initiated on an outpatient basis. The two treatments also had different typical failure modes, based on their different mechanisms of action. This made it difficult to choose a common platform for

the trial, and common outcome measure, which would provide a fair comparison, not favoring one treatment or the other. Vested interests, preferring one treatment or the other, could sway the design process, and need to be recognized and their influence neutralized.

Other ethical issues encountered in CTN-0051 related to the research participants being individuals with substance use disorders, who could be considered vulnerable due to the nature of the illness. Substance use disorders are to some extent disorders of motivation and volition. And, substance exposure itself can produce cognitive impairment [28]. Further, many individuals entering treatment for substance use disorders are under mandate or supervision from the criminal justice system, raising concerns about potential coercion and voluntariness of consent and whether such individuals meet the definition of prisoners under the Federal regulations governing human subjects research (Code of Federal Regulations 45 CFR 46, Subpart C). Effectiveness trials seek highly representative samples with minimal necessary exclusionary criteria. This is favorable with respect to the principle of fair distribution of research risks and benefits across the population. However, this may engage vulnerable populations, who require special protections with respect to informed consent and voluntariness of research participation.

That the risk/benefit ratio should be reasonable, and safety procedures in place to minimize risks, is essential to all human subjects research. Since effectiveness trials take place in routine clinical settings, a research protocol needs to carefully define safety procedures that are part of routine care at the participating site, and safety procedures determined by the protocol, in order to avoid therapeutic misconception [46]. CTN-0051 addressed this in part by constructing a primary trial endpoint of relapse, such that patients who begin to show poor outcome were exited from the trial back to individualized clinical treatment.

Effectiveness trials are typically large and expensive endeavors, and limited funds will limit the type and number of questions that can be addressed. For CTN-0051 an observational design or registry was considered and would have been potentially simpler and cheaper, but ultimately a randomized trial was felt to be needed to best address the central question of comparative effectiveness. More elaborate designs considered for CTN-0051 included choice designs (randomization to choose treatments versus be randomly assigned to treatments), and an adaptive SMART design with re-randomization of relapsed patients to alternative treatments. Examples of such adaptive designs include STAR*D, which randomly assigned depressed outpatients who had failed an initial anti-depressant medication trial, to successive waves of alternative treatments [57]. The size and scope of STAR*D, which started with a sample of roughly 4,000 outpatients, was greater than that of CTN-0051 with its final sample size of 570. For CTN-0051, the choice and SMART designs would have answered additional questions, but it was ultimately decided to study the simple comparison between BUP and XR-NTX in a two group randomized trial. Whether a trial addresses the most important clinical and scientific questions, and whether the ability to address additional questions is worth the additional cost, is ultimately in part an ethical one, relating to the key imperative that a trial be of scientific and social value.

An important limitation of this paper is that it represents a case study of the ethical issues encountered in the design of one clinical trial, CTN-0051. In an effort to consider how this

experience might generalize, we discuss how the specific design features driving the ethical issues in CTN-0051 reflect general characteristics of effectiveness trials. However, a systematic review of the extent to which such ethical issues are encountered across effectiveness trials is beyond the scope of this paper. Whether a given trial would involve any of these types of ethical issues would depend on its specific design features such as the treatments compared and the settings and population involved.

The experience described here designing CTN-0051, a comparative effectiveness trial of two medication treatments for opioid dependence, suggests effectiveness trials may pose unique ethical challenges, given that such trials involve treatments already in widespread use with evidence already available on their effectiveness and safety, preferences on the part of patients, vested interests, variations inherent in the treatments and in their implementation in community-based treatment settings, and highly representative samples that may include vulnerable populations. It is hoped that the ethical issues encountered and the corresponding design solutions developed for CTN-0051 may be useful to future design teams developing effectiveness trials comparing treatments for substance use disorders, as well as other mental or medical disorders.

Acknowledgments

X:BOT is supported by the National Institute on Drug Abuse Clinical Trials Network grants U10DA013046 (NYU, New York Node), U10DA013035 (RFMH, Greater New York Node), UG1DA013035 (NYU, Greater New York Node), UG1DA013732 (OVN, Ohio Valley Node), UG1DA015831 (New England Node), UG1DA013034 (Mid-Atlantic Node), and NIDA Contracts HHSN271201500065C (CCC, The Emmes Corp.) and HHSN271201200017C (DSC, The Emmes Corp.), K24 DA022412 (Dr. Nunes). Indivior (formerly, Reckitt Benckiser) provided Suboxone Film in-kind.

We thank the members of the protocol development team for their work, as well as the staff and participants of the Community Treatment Programs and the associated Nodes for their involvement in the project. We thank the staff of the Clinical Coordinating Center and the Data and Statistics Center for their support, and the staff at the Center for the Clinical Trials Network for their work on this project.

References

1. Institute of Medicine. Bridging the Gap Between Practice and Research: Forging Partnerships with Community-Based Drug and Alcohol Treatment. 1998
2. Brigham GS, Feaster DJ, Wakim PG, Dempsey CL. Choosing a control group in effectiveness trials of behavioral drug abuse treatments. *J Subst Abuse Treat*. 2009 Dec; 37(4):388–97. [PubMed: 19553062]
3. Carroll KM, Ball SA, Jackson R, Martino S, Petry NM, Stitzer ML, Wells EA, Weiss RD. Ten take home lessons from the first 10 years of the CTN and 10 recommendations for the future. *Am J Drug Alcohol Abuse*. 2011 Sep; 37(5):275–82. [PubMed: 21854269]
4. Nunes EV, Ball S, Booth R, Brigham G, Calsyn DA, Carroll K, Feaster DJ, Hien D, Hubbard RL, Ling W, Petry NM, Rotrosen J, Selzer J, Stitzer M, Tross S, Wakim P, Winhusen T, Woody G. Multisite effectiveness trials of treatments for substance abuse and co-occurring problems: have we chosen the best designs? *J Subst Abuse Treat*. 2010 Jun; 38(1):S97–112. [PubMed: 20307801]
5. Nunes EV. The design and analysis of multisite effectiveness trials: a decade of progress in the National Drug Abuse Clinical Trials Network. *Am J Drug Alcohol Abuse*. 2011 Sep; 37(5):269–72. [PubMed: 21854267]
6. Wells EA, Saxon AJ, Calsyn DA, Jackson TR, Donovan DM. Study results from the Clinical Trials Network's first 10 years: where do they lead? *J Subst Abuse Treat*. 2010 Jun; 38(1):S14–30. [PubMed: 20307792]

7. Kirchmayer U, Davoli M, Verster AD, Amato L, Ferri A, Perucci CA. A systematic review on the efficacy of naltrexone maintenance treatment in opioid dependence. *Addiction*. 2002; 97(10):1241–9. [PubMed: 12359026]
8. Johansson BA, Berglund M, Lindgren A. Efficacy of maintenance treatment with naltrexone for opioid dependence: a meta-analytical review. *Addiction*. 2006; 101(4):491–503. [PubMed: 16548929]
9. Nunes EV, Rothenberg JL, Sullivan MA, Carpenter KM, Kleber HD. Behavioral therapy to augment oral naltrexone for opioid dependence: a ceiling on effectiveness? *Am J Drug Alcohol Abuse*. 2006; 32(4):503–17. [PubMed: 17127538]
10. Roozen HG, de Waart R, van der Windt DA, van den Brink W, de Jong CA, Kerkhof AJ. A systematic review of the effectiveness of naltrexone in the maintenance treatment of opioid and alcohol dependence. *Eur Neuropsychopharmacol*. 2006; 16(5):311–23. [PubMed: 16361086]
11. Krupitsky E, Nunes EV, Ling W, Illeperuma A, Gastfriend DR, Silverman BL. Injectable extended-release naltrexone for opioid dependence: a double-blind, placebo-controlled, multicentre randomised trial. *The Lancet*. 2011; 377(9776):1506–1513.
12. Nunes EV, Krupitsky E, Ling W, Zummo J, Memisoglu A, Silverman BL, Gastfriend DR. Treating Opioid Dependence With Injectable Extended-Release Naltrexone (XR-NTX): Who Will Respond? *J Addict Med*. 2015; 9(3):238–43. [PubMed: 25901451]
13. Krupitsky EM, Blokhina EA. Long-acting depot formulations of naltrexone for heroin dependence: a review. *Curr Opin Psychiatry*. 2010 May; 23(3):210–4. [PubMed: 20224403]
14. Krupitsky E, Zvartau E, Blokhina E, Verbitskaya E, Wahlgren V, Tsoy-Podosenin M, Bushara N, Burakov A, Masalov D, Romanova T, Tyurina A, Palatkin V, Slavina T, Pecoraro A, Woody GE. Randomized trial of long-acting sustained-release naltrexone implant vs oral naltrexone or placebo for preventing relapse to opioid dependence. *Arch Gen Psychiatry*. 2012; 69(9):973–81. [PubMed: 22945623]
15. Hulse GK, Morris N, Arnold-Reed D, Tait RJ. Improving clinical outcomes in treating heroin dependence: randomized, controlled trial of oral or implant naltrexone. *Arch Gen Psychiatry*. 2009; 66(10):1108–15. [PubMed: 19805701]
16. Mattick RP, Breen C, Kimber J, Davoli M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database Syst Rev*. 2014 Feb 6; 2:CD002207.
17. Lee JD, Nunes EV, Novo P, Bailey GL, Brigham GS, Cohen AJ, Fishman M, Ling W, Lindblad R, Shmueli-Blumberg D, Stablein D, May J, Salazar D, Liu D, Rotrosen J. NIDA Clinical Trials Network CTN-0051, Extended-Release Naltrexone vs. Buprenorphine for Opioid Treatment (X:BOT): Study Design and Rationale. Submitted manuscript.
18. Turner L, Kruszewski SP, Alexander GC. Trends in the use of buprenorphine by office-based physicians in the United States, 2003–2013. *Am J Addict*. 2015; 24(1):24–9. [PubMed: 25823632]
19. Clark RE, Samnaliev M, Baxter JD, Leung GY. The evidence doesn't justify steps by state Medicaid programs to restrict opioid addiction treatment with buprenorphine. *Health Affairs*. 2011; 30(8):1425–33. [PubMed: 21821560]
20. Clausen T, Waal H, Thoresen M, Gossop M. Mortality among opiate users: opioid maintenance therapy, age and causes of death. *Addiction*. 2009; 104(8):1356–62. [PubMed: 19426288]
21. Weiss RD, Potter JS, Fiellin DA, Byrne M, Connery HS, Dickinson W, Gardin J, Griffin ML, Gourevitch MN, Haller DL, Hasson AL, Huang Z, Jacobs P, Kosinski AS, Lindblad R, McCance-Katz EF, Provost SE, Selzer J, Somoza EC, Sonne SC, Ling W. Adjunctive counseling during brief and extended buprenorphine-naloxone treatment for prescription opioid dependence: a 2-phase randomized controlled trial. *Arch Gen Psychiatry*. 2011 Dec; 68(12):1238–46. [PubMed: 22065255]
22. Bailey GL, Herman DS, Stein MD. Perceived relapse risk and desire for medication assisted treatment among persons seeking inpatient opiate detoxification. *J Subst Abuse Treat*. 2013; 45(3):302–5. [PubMed: 23786852]
23. Binswanger IA, Blatchford PJ, Mueller SR, Stern MF. Mortality after prison release: opioid overdose and other causes of death, risk factors, and time trends from 1999 to 2009. *Ann Intern Med*. 2013; 159(9):592–600. [PubMed: 24189594]

24. Substance Abuse and Mental Health Services Administration. National Survey of Substance Abuse Treatment Services (N-SSATS): 2013. Rockville, MD: 2014. HHS Publication No. (SMA) 14-4890
25. Nosyk B, Anglin MD, Brissette S, Kerr T, Marsh DC, Schackman BR, Wood E, Montaner JS. A call for evidence-based medical treatment of opioid dependence in the United States and Canada. *Health Aff.* 2013; 32(8):1462–1469.
26. Olsen Y, Sharfstein JM. Confronting the stigma of opioid use disorder--and its treatment. *JAMA.* 2014; 311(14):1393–4. [PubMed: 24577059]
27. Wakeman SE. Using Science to Battle Stigma in Addressing the Opioid Epidemic: Opioid Agonist Therapy Saves Lives. *Am J Med.* 2016; 129(5):455–6. [PubMed: 26836384]
28. Geppert CMA, Bogenschutz MP. Ethics in substance use disorder treatment. *Psychiatr Clin N Am.* 2009; 32:283–297.
29. Weiss RD, Potter JS, Griffin ML, Provost SE, Fitzmaurice GM, McDermott KA, Srisarajivakul EN, Dodd DR, Dreifuss JA, McHugh RK, Carroll KM. Long-term outcomes from the National Drug Abuse Treatment Clinical Trials Network Prescription Opioid Addiction Treatment Study. *Drug Alcohol Depend.* 2015; 150:112–9. [PubMed: 25818060]
30. Gossop M, Green L, Phillips G, Bradley B. Lapse, relapse and survival among opiate addicts after treatment. A prospective follow-up study. *Br J Psychiatry.* 1989; 154:348–53. [PubMed: 2597836]
31. Chutuaape MA, Jasinski DR, Fingerhood MI, Stitzer ML. One-, three-, and six-month outcomes after brief inpatient opioid detoxification. *Am J Drug Alcohol Abuse.* 2001; 27(1):19–44. [PubMed: 11373035]
32. Broers B, Giner F, Dumont P, Mino A. Inpatient opiate detoxification in Geneva: follow-up at 1 and 6 months. *Drug Alcohol Depend.* 2000; 58(1-2):85–92. [PubMed: 10669058]
33. Lee JD, Friedmann PD, Kinlock TW, Nunes EV, Boney TY, Hoskinson RA Jr, Wilson D, McDonald R, Rotrosen J, Gourevitch MN, Gordon M, Fishman M, Chen DT, Bonnie RJ, Cornish JW, Murphy SM, O'Brien CP. Extended-Release Naltrexone to Prevent Opioid Relapse in Criminal Justice Offenders. *N Engl J Med.* 2016; 374(13):1232–42. [PubMed: 27028913]
34. Ling W, Amass L, Shoptaw S, Annon JJ, Hillhouse M, Babcock D, Brigham G, Harrer J, Reid M, Muir J, Buchan B, Orr D, Woody G, Krejci J, Ziedonis D. Buprenorphine Study Protocol Group. A multi-center randomized trial of buprenorphine-naloxone versus clonidine for opioid detoxification: findings from the National Institute on Drug Abuse Clinical Trials Network. *Addiction.* 2005; 100(8):1090–100. [PubMed: 16042639]
35. Sigmon SC, Bisaga A, Nunes EV, O'Connor PG, Kosten T, Woody G. Opioid detoxification and naltrexone induction strategies: recommendations for clinical practice. *Am J Drug Alcohol Abuse.* 2012 May; 38(3):187–99. [PubMed: 22404717]
36. Mogali S, Khan NA, Drill ES, Pavlicova M, Sullivan MA, Nunes E, Bisaga A. Baseline characteristics of patients predicting suitability for rapid naltrexone induction. *Am J Addict.* 2015; 24(3):258–64. [PubMed: 25907815]
37. Knudsen HK, Abraham AJ, Roman PM. Adoption and implementation of medications in addiction treatment programs. *J Addict Med.* 2011; 5:21–7.
38. Wallack SS, Thomas CP, Martin TC, et al. Substance abuse treatment organizations as mediators of social policy: Slowing the adoption of a congressionally approved medication. *J Behav Health Serv Res.* 2010; 37:64–78. [PubMed: 18668369]
39. Wolfe D, Carrieri MP, Dasgupta N, Wodak A, Newman R, Bruce RD. Concerns about injectable naltrexone for opioid dependence. *Lancet.* 2011; 377(9776):1468–70. [PubMed: 21529930]
40. Miller NS, Goldsmith RJ. Craving for alcohol and drugs in animals and humans: Biology and behavior. *Journal of Addictive Diseases.* 2001; 20(3):87–104. [PubMed: 11681596]
41. Jeste D, Saks E. Decisional Capacity in Mental Illness and Substance Abuse: Empirical Database and Policy Implications. *Behavioral Sciences and the Law.* 2006; 24:607–628. [PubMed: 16883611]
42. Huddleston W, Marlowe DB. Painting the Current Picture: A National Report on Drug Courts and Other Problem-Solving Court Programs in the United States. National Drug Court Institute. 2011

43. Matusow, Harlan, et al. Medication assisted treatment in US drug courts: Results from a nationwide survey of availability, barriers and attitudes. *Journal of substance abuse treatment*. 2013; 44(5):473–480. [PubMed: 23217610]
44. Title 45 Public Welfare. Department of Health and Human Services; Code of Federal Regulations (45 CFR 46). Part 46 Protection of Human Subjects
45. Institute of Medicine. *Ethical Considerations for Research Involving Prisoners*. Washington, DC: The National Academies Press; 2007.
46. Appelbaum PS. Clarifying the ethics of clinical research: a path toward avoiding the therapeutic misconception. *Am J Bioeth*. 2002; 2(2):22–3.
47. Lidz CW, Appelbaum PS. The therapeutic misconception: problems and solutions. *Med Care*. 2002; 40(9 Suppl):V55–63. [PubMed: 12226586]
48. Murphy SA1, Lynch KG, Oslin D, McKay JR, TenHave T. Developing adaptive treatment strategies in substance abuse research. *Drug Alcohol Depend*. 2007; 88(2):S24–30. [PubMed: 17056207]
49. Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. *Statistics in Medicine*. 2012; 31(17):1779–1902. [PubMed: 22576824]
50. Campbell AN, Nunes EV, Matthews AG, Stitzer M, Miele GM, Polsky D, Turrigiano E, Walters S, McClure EA, Kyle TL, Wahle A, Van Veldhuisen P, Goldman B, Babcock D, Stabile PQ, Winhusen T, Ghitza UE. Internet-delivered treatment for substance abuse: a multisite randomized controlled trial. *Am J Psychiatry*. 2014; 171(6):683–90. [PubMed: 24700332]
51. Krupitsky E, Nunes EV, Ling W, Gastfriend DR, Memisoglu A, Silverman BL. Injectable extended-release naltrexone (XR-NTX) for opioid dependence: long-term safety and effectiveness. *Addiction*. 2013; 108(9):1628–37. [PubMed: 23701526]
52. National Commission. *The Belmont report: ethical principles and guidelines for the protection of human subjects of research*. Washington DC: US Government Printing Office; 1979.
53. Emanuel E, Wendler D, Grady C. What makes clinical research ethical? *JAMA*. 2000; 283(20):2701–2711. [PubMed: 10819955]
54. Lantos JD, Feudtner C. SUPPORT and the Ethics of Study Implementation: Lessons for Comparative Effectiveness Research from the Trial of Oxygen Therapy for Premature Babies. *Hastings Cent Rep*. 2015 Jan-Feb;45(1):30–40. [PubMed: 25530316]
55. Chaitman BR, Hardison RM, Adler D, Gebhart S, Grogan M, Ocampo S, Sopko G, Ramires JA, Schneider D, Frye RL. Bypass Angioplasty Revascularization Investigation 2 Diabetes (BARI 2D) Study Group. The Bypass Angioplasty Revascularization Investigation 2 Diabetes randomized trial of different treatment strategies in type 2 diabetes mellitus with stable ischemic heart disease: impact of treatment strategy on cardiac mortality and myocardial infarction. *Circulation*. 2009; 120(25):2529–40. [PubMed: 19920001]
56. Jacobs WC, van Tulder M, Arts M, Rubinstein SM, van Middelkoop M, Ostelo R, Verhagen A, Koes B, Peul WC. Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. *Eur Spine J*. 2011; 20(4):513–22. [PubMed: 20949289]
57. Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA, Thase ME, Nierenberg AA, Quitkin FM, Kashner TM, Kupfer DJ, Rosenbaum JF, Alpert J, Stewart JW, McGrath PJ, Biggs MM, Shores-Wilson K, Lebowitz BD, Ritz L, Niederehe G. STAR*D Investigators Group. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*. 2004; 25(1):119–42. [PubMed: 15061154]

Table 1
Requirements for determining whether a trial is ethical (adapted from Emanuel, et al. 2000)

Requirement	Description
1. Social or scientific value and equipoise	Research will improve health and well-being or increase scientific knowledge. There should be reasonable uncertainty about the outcome.
2. Scientific validity	Use of accepted scientific methods and analyses to produce valid data.
3. Fair subject selection	Vulnerable populations are not targeted for risky research, and powerful individuals are not favored for beneficial research.
4. Favorable risk-benefit ratio	Risks are minimized, benefits are enhanced; risks are proportionate to benefits to individual and society.
5. Independent review	Ethics review conducted by individuals unaffiliated with the research
6. Informed consent	Disclosure of risks, benefits, and alternatives to volunteers who can make a decision to participate or not in research.
7. Respect for potential and enrolled subjects	Permitting withdrawal from research; protecting privacy; informing subjects of newly discovered risks or benefits; informing subjects of results of clinical research; maintaining the welfare of subjects.

Table 2

Ethical issues in the design of effectiveness trials: Lessons learned from CTN-0051, a comparative trial of maintenance medication treatments for opioid dependence—buprenorphine-naloxone (BUP-NX), a partial opioid receptor agonist, versus extended release injection naltrexone (XR-NTX), an opioid receptor antagonist.

Design Issues	Ethical Dimensions	CTN-0051 Examples
Treatment as Usual (TAU) control condition	<u>Equipose, risk</u> : Existing evidence on potential usual treatment control conditions may suggest inferior outcome, or dangers.	One potential TAU control condition—detoxification followed by counseling without medication—was in widespread use, but had evidence of poor outcome, and risk of death from drug overdose. It was ultimately not included in the design.
Patient preferences for treatments	<u>Scientific validity, respect for patients</u> : In real world practice patients choose treatments with their clinicians. Since treatments may already be in widespread use, patients or clinicians may have preferences. Does choice contribute to effectiveness? Should patients be allowed to choose treatments?	Widespread treatments for opioid dependence include agonist maintenance with methadone or buprenorphine, and detoxification followed by counseling (so called “drug free” treatment). Naltrexone, an antagonist, is more consistent with the “drug free” approach. Many patients and clinicians have preferences. Randomization to either choose treatment, or to be randomized to treatment, was considered, but rejected due to complexity and cost. Instead, patients rate their preference prior to randomization, for study as a covariate.
Inherent differences between treatments and their regimens or platforms	<u>Equipose, designing a fair comparison</u> : Different treatments may be best suited for different treatment platforms or regimens. Choosing a common platform or setting for a clinical trial may put one or another treatment at a disadvantage.	Buprenorphine is easily initiated on an outpatient basis over one or two days. Naltrexone requires complete detoxification prior to initiation, typically requiring hospitalization for 5 to 7 days or more.
Vested interests	<u>Equipose, designing a fair comparison</u> : Treatments in widespread use may have vested interests, financial (manufacturers, clinicians whose practices center on a particular treatment) or ideological (clinicians' or patients' beliefs). The Design Team of trial may feel pressure from such groups.	Agonist maintenance treatment and “drug free” treatment have long traditions and strong proponents in communities of patients and clinicians. Entire systems of care are organized around each. BUP-NX and XR-NTX both have interested manufacturers.
Generalizable sample	<u>Vulnerable populations</u> : Effectiveness trials seek broadly representative samples. This is favorable to fair distribution of research risks and benefits. It may also engage vulnerable populations with issues around capacity to consent, and voluntariness.	Substance use disorders may impair capacity to consent, due to direct effects of substances or more subtle impairments in motivation underlying the disorders. Patients with opioid dependence are often under some form of supervision by the criminal justice system, and might be considered prisoners under the research regulations.
Safety: During-trial management, exit criteria, and rescue procedures	<u>Risk/benefit ratio, therapeutic misconception</u> : Research protocol must have adequate safety procedures (e.g. management plans for anticipated risky situations, exit criteria, rescue procedures after protocol exit). Community treatment sites for effectiveness trials will already have such procedures—hence potential confusion between usual treatment and research treatment.	Relapse to opioid use carries significant risks, including death from drug overdose. CTN-0051 included clear procedures for managing risky situations (e.g. where patients have missed doses and are using opioids). The primary study endpoint, Relapse, sets a high bar for good clinical response, and patients with persistent opioid use are called Relapsed and exited from the protocol into individualized clinical care.
Study cost and feasibility	<u>Scientific and social value, ideal evidence</u> : Effectiveness trials are typically large and expensive, placing constraints on design features and how many questions can be addressed.	A simple 2-group trial of BUP-NX vs XR-NTX was ultimately chosen. More complex and costly designs were considered, including choice designs, and a SMART design, where patients who relapse under the first assigned treatment are re-randomized to alternative treatments.