Published in final edited form as:

Drug Alcohol Depend. 2009 December 1; 105S1: S14–S25. doi:10.1016/j.drugalcdep.2009.04.003.

Principles of laboratory assessment of drug abuse liability and implications for clinical development

Lawrence P. Carter¹ and Roland R. Griffiths²

¹ Department of Psychiatry, University of Arkansas for Medical Sciences

² Department of Psychiatry and Behavioral Sciences and Department of Neuroscience, Johns Hopkins School of Medicine

Abstract

Abuse liability testing plays an important role in informing drug development, regulatory processes, and clinical practice. This paper describes the current “gold standard” methodologies that are used for laboratory assessments of abuse liability in non-human and human subjects. Particular emphasis is given to procedures such as non-human drug discrimination, self-administration, and physical dependence testing, and human dose effect abuse liability studies that are commonly used in regulatory submissions to governmental agencies. The potential benefits and risks associated with the inclusion of measures of abuse liability in industry-sponsored clinical trials is discussed. Lastly, it is noted that many factors contribute to patterns of drug abuse and dependence outside of the laboratory setting and positive or negative signals in abuse liability studies do not always translate to high or low levels of actual abuse or dependence. Well-designed patient and physician education, pharmacovigilance, and postmarketing surveillance can reduce the diversion and misuse of drugs with abuse liability and can effectively foster the protection and promotion of public health.

Keywords

abuse liability; clinical trial; drug abuse; drug development; drug discrimination; self-administration

1. Introduction

The relative abuse liability of a novel drug is an important consideration in the development of that drug as a new medication. An accurate estimate of the abuse liability of a new drug informs the decisions of the commercial sponsor, governmental agencies, and ultimately, the physicians and patients who might use the drug product. In the process of drug development, a commercial sponsor must make a number of decisions when designing clinical trials, estimating safety and commercial risks, implementing risk mitigation strategies, and designing postmarketing surveillance programs, each of which are affected by the abuse liability of the drug under development (see Mansbach et al., 2003 for review).

Corresponding Author: Roland R. Griffiths, Ph.D. Johns Hopkins University – BPRU, 5510 Nathan Shock Drive, Baltimore, Maryland 21224, Telephone: (410) 550-0034, Fax: (410) 550-0030, E-mail: rgriff@jhmi.edu. Lawrence P. Carter, Ph.D. University of Arkansas for Medical Sciences, Psychiatric Research Institute – Center for Addiction Research, 4301 W. Markham Street #843, Little Rock, AR 72205.

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 citable 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.

Abuse liability studies are used by the Drug Enforcement Agency for placement of a drug into one of five categories or schedules under the Controlled Substances Act, and are required by the Food and Drug Administration (FDA) as part of a new drug application. The FDA considers both the *promotion* of public health (encouraging the development of safe and effective new pharmacotherapeutic agents) and the *protection* of public health (ensuring that new products are introduced with an adequate degree of scientific knowledge and, when indicated, an appropriate degree of regulatory control related to their abuse liability) in their decisions and recommendations to other agencies (1990 FDA Draft Guidelines for Abuse Liability Assessment, in Balster and Bigelow, 2003). Taken together, these two considerations emphasize the need for balance between ensuring that marketed drugs are safe and effective, and encouraging the development of new pharmacotherapies.

Abuse liability assessment is also important for the physicians who will prescribe and for the patients who will use new therapeutic drugs. Both physicians and patients should be aware of the potential risks and benefits of a therapeutic drug, including its abuse liability. No drug is completely safe and every drug has specific risks associated with its use. Both the overestimation and underestimation of abuse liability can negatively impact public health through the failure to adequately treat legitimate illnesses such as pain and anxiety (due to an unreasonable fear of abuse liability) or through the failure to prescribe drugs with abuse liability in a judicious manner, particularly to individuals who might be at higher risk for abuse (McCabe et al., 2008; Wu et al., 2008).

FDA Guidelines define abuse liability as “the likelihood that a drug with psychoactive or central nervous system (CNS) effects will sustain patterns of non-medical self-administration that result in disruptive or undesirable consequences” and the guidelines differentiate between the likelihood/severity of self-administration and the likelihood/severity of undesirable consequences (1990 FDA Draft Guidelines for Abuse Liability Assessment, in Balster and Bigelow, 2003). As described in this volume by Leiderman, the FDA Amendments Act of 2007 has increased FDA’s focus on the safety of pharmaceutical products and the requirement for pharmacovigilance throughout a product’s life cycle. The FDA Amendments Act of 2007 gives FDA new authority and resources to require and enforce Risk Evaluation and Mitigation Strategies (REMS) for high risk pharmaceutical products (Leiderman, 2009).

The purpose of the present paper is to provide a general overview of abuse liability assessment in non-human and human subjects with particular attention paid to issues regarding study design, dependent measures, and the potential applicability of those measures to industry-sponsored clinical trials. The aim is to describe the principal methodologies for the assessment of the likelihood and consequences of abuse of centrally acting compounds in a context that is applicable to the drug development process. The selected methodologies that are discussed in this review represent the most commonly used procedures for the assessment of abuse liability for regulatory purposes. This paper is not intended to be a review of all methodologies or a critical evaluation of the different procedures used to assess abuse liability. Indeed, detailed and comprehensive reviews covering abuse liability testing in non-human and human subjects have been previously written on these topics (e.g., Ator and Griffiths, 2003; Balster and Bigelow, 2003; Griffiths et al., 2003; Panlilio and Goldberg, 2007; Comer et al., 2008; Schoedel and Sellers, 2008).

2. Behavioral procedures for the assessment of abuse liability in non-human subjects

Some behavioral pharmacology procedures such as drug discrimination and drug self-administration have been developed for use in both non-human and human subjects; however, the use of non-human subjects for these procedures has several distinct advantages over using

similar procedures in human volunteers. Studies in human volunteers can be limited by the drugs and procedures that can be safely and ethically administered to humans, the amount of time and effort that human volunteers are willing to commit to a scientific study, and the oversight and cost that is required for the proper conduct of human research. Studies in non-human subjects often allow for studies of a longer duration, which can result in a more thorough examination of a larger range of doses or the effects of chronic treatment of a drug of interest. Also, there is a greater variety of pharmacological tools (i.e., chemical compounds that are not approved for use in humans) that are available for characterizing the pharmacology of drugs in non-human subjects, and the relative cost of non-human studies can be markedly less than those in humans.

2.1. Assessment of the likelihood of abuse in non-human subjects

In non-human subjects, data from drug discrimination and drug self-administration studies are used to predict the likelihood that a drug will be abused. This section will describe the experimental designs and the dependent variables typically used to predict the likelihood of abuse for regulatory purposes.

Alternative procedures, such as conditioned place preference, have also been used to provide information about the likelihood that a drug will be abused. However, unlike drug self-administration procedures, conditioned place preference procedures do not provide a direct assessment of the reinforcing effects of drugs. Thus, they provide neither a face-valid nor a functional model of the human condition of compulsive drug taking. Also unlike drug self-administration procedures, conditioned place preference procedures tend to be conducted only in rodents and usually require the use of a larger number of animals to study a range of doses due to the between-subjects design of the studies (Bardo and Bevins, 2000). It is possible that these differences have limited the use of the conditioned place preference procedure in studies designed to examine the abuse liability of drugs for regulatory purposes.

2.1.1. Drug discrimination in non-human subjects—The pharmacological mechanism of action of a drug can be indicative of its relative risk of being abused. It is often the case that at least some information regarding the mechanism of action of a drug is available from *in vitro* studies before a drug is studied *in vivo*. Such information, for example, might include a receptor binding profile or data from a functional assay in cell culture or a tissue preparation. Drug discrimination provides information regarding the pharmacological mechanism of action of a drug in an intact behaving animal. Although drug discrimination procedures are not designed to examine the reinforcing effects of drugs per se, drugs that share discriminative stimulus effects are likely to have a common pharmacological mechanism of action and might also have similar reinforcing effects (Schuster and Johanson, 1988).

2.1.1.1. Experimental designs used in drug discrimination studies: A drug vs. saline discrimination procedure generally involves first training animals to make an operant response to receive a positive reinforcer or avoid a negative reinforcer in an operant conditioning chamber. For example, animals might be trained to make a nose poke (mouse), key peck (pigeon), or lever press (rat, monkey) in order to receive food or water. After the operant behavior is established, animals can be trained to discriminate drug from saline with two response options (e.g., two levers). In a drug discrimination procedure with two response options, animals are typically trained such that responding on one of the two options is reinforced following administration of a dose of the training drug and responding on the other option is reinforced following administration of vehicle. Testing is generally conducted under conditions in which responding on either response option is reinforced, although testing can also be performed under conditions in which responding is not reinforced on either response option. One of the greatest values of the drug discrimination procedure is its pharmacological

selectivity: drugs that share a pharmacological mechanism of action, not drugs in general, tend to occasion responding on the training drug-paired response option.

The drug discrimination procedure offers a remarkable amount of flexibility in that it has been used to study a wide variety of drugs using different species, routes of administration, operant responses, and reinforcers. Training animals to discriminate a novel drug from saline or vehicle, particularly when the mechanism of action is unknown or thought to occur through multiple receptors, allows for the study of a range of agonists and antagonists from different drug classes in the same animals. However, if the drug of interest acts at multiple receptors or has several mechanisms of action, drugs from multiple classes might occasion drug-appropriate responding (e.g., Winter et al., 1981; Carter et al., 2003), resulting in data that are challenging to interpret and might require further investigation.

In addition to training animals to discriminate the drug of interest, an alternative strategy can be used in which the drug of interest is studied in animals trained to discriminate another drug with a selective, known mechanism of action (e.g., a known drug of abuse). Using this approach, members of the Drug Evaluation Committee of the College on Problems of Drug Dependence have routinely studied the pharmacology of new potential drugs of abuse in established drug discrimination procedures in which animals have been trained to discriminate different drugs of abuse from saline or vehicle (e.g., Woolverton et al., 1999; Fantegrossi et al., 2005). An advantage of such an approach is that many novel drugs can be studied in the established procedures without training additional groups of animals in new behavioral procedures.

Given the increasingly common development of drugs with novel pharmacological mechanisms of action, multiple drug discrimination studies might be necessary to thoroughly investigate the mechanism of action of a drug *in vivo*. For example, training animals to discriminate the novel drug and studying different drugs of abuse in those animals could be used to identify the most appropriate comparator drug or drugs (i.e., those that occasion some drug-appropriate responding) for use as a training drug in subsequent drug discrimination procedures designed to examine a full range of doses of the novel drug as well as related agonists, antagonists, etc. (cf., Carter et al., 2003 and Carter et al., 2004).

In training animals in a drug discrimination procedure, the dose of the training drug and the behavioral and pharmacological history of the animals can qualitatively and quantitatively influence the results of the study. Quantitative differences in the potency of the drugs to occasion discriminative stimulus effects (i.e., leftward or rightward shifts in the dose effect curves) might be observed as the training dose of a drug is decreased or increased. Qualitative differences might include fewer test compounds occasioning drug-appropriate responding at larger training doses, which usually results in greater pharmacological specificity (cf., Rowlett et al., 1999 and Rowlett et al., 2000; Zhang et al., 2000). Training animals to discriminate multiple drugs in successive or concurrent drug discrimination procedures reduces the pharmacological specificity of the procedure (McMillan et al., 1996; Koek et al., 2005), although this approach is not typically employed for the purpose of abuse liability assessment. However, the administration of different drugs under test conditions does not appear to markedly affect the pharmacological selectivity of the procedure. In fact, one advantage of the drug discrimination procedure is that animals can be tested repeatedly over many years without loss of pharmacological selectivity or apparent tolerance to the discriminative stimulus effects of the training drug (Colpaert, 1995).

2.1.1.2. Dependent measures used in drug discrimination studies: The primary dependent measure in a drug discrimination procedure is the operant behavior that occurs on the response option that has been associated with the training drug (drug-appropriate responding). Drug-

appropriate responding is frequently summarized as the percentage of total operant responses that are made on the response option that has been reinforced after administration of the training drug. It can also be informative to analyze drug discrimination data by calculating the number or percentage of animals that “choose” the response option associated with the training drug, with the “choice” typically defined as either the response option on which 80% or more responding occurs or the response option on which the first reinforcer is received.

Another useful dependent measure in drug discrimination studies is the rate of responding across all of the response options. Increases or decreases in the rate of responding indicate that behaviorally-active doses of a drug have been studied. In the absence of significant effects (typically decreases) in the rate of responding, it is usually unclear whether a drug of interest was studied up to large enough (behaviorally-active) doses.

2.1.2. Drug self-administration in non-human subjects—One of the aims of abuse liability testing is to accurately predict whether a compound will maintain patterns of non-medical self-administration that result in disruptive or undesirable consequences (1990 Draft Guidelines for Abuse Liability Assessment, in Balster and Bigelow, 2003). Operant drug self-administration is considered the “gold standard” of non-human abuse liability testing because of its high level of face validity and predictive validity (see Haney and Spealman, 2008). Data from a drug self-administration study in animals provide information about whether a drug is likely to maintain non-medical self-administration (i.e., be abused) in humans. Moreover, there is a generally good correspondence between drugs that are self-administered by non-human subjects and those that function as reinforcers in humans (Schuster and Thompson, 1969; Griffiths et al., 1980; Haney and Spealman, 2008).

2.1.2.1. Experimental designs used in drug self-administration studies: A reinforcer can be defined as a stimulus (e.g., drug administration) that increases the future probability of behavior on which its presentation is contingent. Drugs of abuse such as cocaine or heroin are considered to be reinforcers because the opportunity to self-administer those drugs typically increases rates of self-administration relative to saline or vehicle. In contrast, other classes of psychotropic drugs (e.g., phenothiazine antipsychotics) are not considered to be reinforcers because they are not reported to increase rates of self-administration behavior by human or non-human subjects.

The earliest drug self-administration procedures were designed to allow rats or rhesus monkeys to make an operant response to receive an intravenous dose of morphine through an indwelling catheter (Weeks, 1962; Thompson and Schuster, 1964). Currently, most drug self-administration procedures in non-human subjects use an intravenous route of administration, regardless of species and drug class (e.g., Thomsen and Caine, 2006; Winger et al., 2006). Given that a faster rate of onset is associated with greater reinforcing effects and abuse liability (e.g., Balster and Schuster, 1973; de Wit et al., 1993; Mumford et al., 1995a), intravenous drug self-administration is likely a more rigorous test of abuse liability than self-administration via oral or intragastric routes. Additionally, intravenous self-administration procedures eliminate the potential problems of drug taste, pharmacokinetic effects related to the presence or absence of food in the stomach, or drug fading procedures that can occur with oral drug self-administration (Carroll et al., 1984; Meisch et al., 1993). Another advantage of the intravenous drug self-administration procedure is that studying different drugs, doses, and saline (or vehicle) in the same animals can be readily accomplished by changing the solution available for infusion without altering other stimuli (e.g., taste) that might otherwise affect drug-taking behavior via other routes of administration.

Drug self-administration procedures often involve first training animals to make an operant response to receive a positive reinforcer such as palatable food (e.g., a banana-flavored food

pellet for a monkey) or a drug that is known to serve as a reinforcer (e.g., cocaine). Cocaine, pentobarbital, and methohexital have been used as baseline or initial reinforcing drugs for many abuse liability evaluations by members of the Drug Evaluation Committee of the College on Problems of Drug Dependence, because these drugs are readily self-administered and have relatively short durations of action, allowing for many opportunities for self-administration within a fixed amount of time (e.g., Woolverton et al., 1999; Fantegrossi et al., 2005). There are several advantages to substituting a dose of a drug of interest for another reinforcer that is already maintaining a high rate of self-administration compared to training naïve animals to self-administer a drug: 1) animals are more likely to come into contact with (i.e., self-administer a dose of) the drug of interest if they have a history of responding to receive a reinforcer; 2) the initial reinforcer that maintained a high rate of responding can serve as a positive control; and 3) if the initial reinforcer is delivered intravenously, after a negative result with a test drug, reinstatement of self-administration with the initial drug reinforcer will ensure that an animal's catheter is not damaged or compromised (Ator and Griffiths, 2003).

The "substitution procedure" of studying a test drug in animals that have been previously trained to self-administer another drug of abuse has been criticized by some for using animals that have previously "learned" to self-administer drugs; however, studies have shown that reinforcers that are self-administered by animals with histories of drug self-administration are also typically self-administered by animals without such histories (cf., Tanda et al., 2000 and Justinova et al., 2003). However, the drug history or the context of drug substitution, can influence the self-administration of a drug (e.g., Ator and Griffiths, 2003; Collins and Woods, 2007). Therefore, it might be advisable to study self-administration behavior over multiple sessions and to specify a criterion for stable behavior *a priori* when a drug of interest is studied in animals trained to self-administer another reinforcing drug (e.g., Weerts et al., 1999). A negative control (e.g., intravenous saline) should always be studied to ensure that the drug, and not other stimuli, is responsible for maintaining the behavior.

In operant conditioning procedures, a schedule is a rule that specifies the contingency between the behavior and the delivery of stimuli. Fixed ratio (FR) schedules, fixed interval (FI) schedules, and combinations and variations thereof are commonly used in self-administration studies for abuse liability assessment. FR schedules, which require that an animal make a fixed number of responses to receive an infusion of drug, allow the animal to control the frequency of drug delivery (by responding) and allow the investigator to specify the amount of behavior (e.g., an FR50 schedule would require 50 responses) for the delivery of a reinforcer. Alternatively, FI schedules, in which the first response after a fixed period of time (i.e., interval) results in delivery of a reinforcer, allows the investigator to specify how often a reinforcer is available for delivery. Under each of these schedules, the unconditioned effects of drugs (e.g., ataxia, stereotypy, which might be independent of the drug's reinforcing effects) can affect the ongoing rate of responding. To reduce or eliminate the potential influence of direct drug effects on responding, some investigators have used second order schedules that include a long FI component (e.g. 30 minutes) to study behavior that occurs prior to drug delivery. In some cases, sessions conclude with a single administration of drug (Panlilio and Goldberg, 2007). Progressive-ratio procedures and choice procedures have also been used to provide measures of drug reinforcement under conditions that reduce the influence of the direct drug effects on responding (Panlilio and Goldberg, 2007).

A drug that serves as a reinforcer under one schedule of reinforcement will typically serve as a reinforcer under other schedules of reinforcement (e.g., Czoty et al., 2007); however, different schedules allow for experimenter control of different variables. The examination of self administration behavior under different schedules of reinforcement or under different conditions within a single schedule can provide valuable information with regard to the range of conditions under which doses of drugs will maintain behavior or the sensitivity of the

reinforcing effects of the drug to increases in price or effort required to obtain or administer the drug. For example, progressive ratio (PR) schedules are modified FR schedules in which the number of responses required for reinforcer delivery is increased after a specified number of reinforcers have been delivered. As such, the maximum amount of behavior an animal expends to receive a dose of drug (i.e., the final FR at which a reinforcer is self-administered) is assessed. Similarly, the examination of self-administration behavior at different unit prices (i.e., by manipulating FR values or drug dose) allows for the behavioral economic analysis of the demand for a drug at different unit prices (Hursh and Winger, 1995).

A timeout in an operant drug self-administration procedure is a period of time after the delivery of a putative reinforcer in which responding has no programmed consequence and in which responding will not result in the delivery of the reinforcer. Under a FR schedule in which the frequency of reinforcer delivery is determined by the response rate of the subject, the use of a timeout allows for greater experimenter control over the frequency of reinforcer delivery. There are several advantages to using a timeout period in which drug is not available for self-administration: 1) the period of drug unavailability can reduce the likelihood of drug accumulation; 2) direct behavioral effects (e.g., sedation) of longer acting drugs can subside prior to the next opportunity for self-administration; 3) reduced availability of drug might allow for a broader range of doses to be safely studied; and 4) the patterns of drug self-administration (e.g., every few hours) and availability (e.g., 24 hours a day) might more closely model the reported or anticipated recreational use of the drug by human drug users (Ator and Griffiths, 2003). Moreover, the duration of the timeout period has been shown to affect the results of drug self-administration studies. Sedative drugs of abuse that were not self-administered under conditions of relatively brief timeouts and limited access (1–2 hours) have been subsequently shown to be self-administered at rates comparable to stimulant drugs of abuse when studied under conditions of relatively lengthy 3 hour timeouts and 24 hour/day drug availability (cf., Goldberg et al., 1971 and Griffiths et al., 1981; Woolverton et al., 1999 and Weerts et al., 2008). Thus, although the use of a lengthy timeout period might not be necessary for all drugs, a self-administration procedure that includes a lengthy timeout period might represent a more rigorous and widely applicable test of the abuse liability of novel drugs for which the pharmacological mechanism of action or pharmacokinetic characteristics are not completely known.

2.1.2.2. Dependent measures used in drug self-administration studies: A reinforcer is defined as a stimulus that increases the future probability or likelihood of behavior on which its presentation is contingent. As such, the amount or rate of drug-taking behavior is typically the primary dependent measure in a drug self-administration study. A drug is considered to function as a reinforcer if one or more doses of the drug maintain responding that is significantly higher than the drug vehicle. Specific measures of drug-taking behavior will depend on the schedule employed and the design of the study. In studies that use FR schedules, many investigators report the number of drug infusions or reinforcers earned, and might also include the rate of responding and total drug intake as a function of dose. Under FI and some second order schedules, the number of injections or drug intake may be determined by the schedule, and the rate of responding is usually the primary dependent variable. In studies utilizing progressive ratio schedules, the “breakpoint” (sometimes defined as the ratio at which the animal fails to self-administer a minimum defined number of reinforcers) is typically considered to be the primary dependent variable. In studies using choice procedures, relative preference or choice of one drug dose over another option is usually the primary dependent variable.

In each case, a full dose effect curve should be determined for a drug that is being evaluated for abuse liability, in addition to data for at least one positive and negative control (e.g., cocaine and vehicle, respectively). If drug reinforcement is demonstrated, reinforcing effects are

typically an inverted U-shaped function of dose. If drug reinforcement is not demonstrated, confirmation that behaviorally active and sufficiently large doses of the test compound were studied might be verified by studying effects of different doses of the drug on other physiological or behavioral endpoints such as locomotor behavior or food-reinforced responding. Ideally, an appropriate range of doses should be identified via effects on other behaviors prior to evaluation in self-administration studies.

For studies in which behavioral economic analyses are to be applied, the “cost” (i.e., fixed-ratio value) and/or magnitude of reinforcer (i.e., dose) are varied to allow the plotting of demand curves in which self-administration data are presented as a function of unit price. Whereas the breakpoint from a PR schedule would be the FR value at which self-administration approaches zero, behavioral economic analyses often report the FR or unit prices (or FR values if dose is not manipulated) at which maximum responding is observed (Hursh and Winger, 1995; Bickel et al., 2000).

2.2. Assessment of the consequences of abuse in non-human subjects

One of the aims of abuse liability testing is to predict the consequences of abuse of a drug of interest (1990 Draft Guidelines for Abuse Liability Assessment, in Balster and Bigelow, 2003). In non-human subjects, data from chronic administration studies examining physical dependence and withdrawal are typically used to predict whether the abrupt discontinuation of a drug will result in behavioral disruption, although any number of behavioral or biological endpoints (e.g., seizure threshold, neuronal damage) may be used to examine the toxicity of a drug. This section will describe the experimental designs and dependent variables typically used in studies of physical dependence and withdrawal.

2.2.1. Physical dependence and withdrawal in non-human subjects—Physical dependence is manifested by time-limited biochemical, physiological, or behavioral changes (i.e., a withdrawal syndrome) that occur upon termination of chronic drug administration. Physical dependence is distinct from the absence of a drug effect (i.e., a return to baseline) and occurs as a result of an organism’s acclimation to chronic drug administration. Historically, physical dependence has long been associated with abuse liability (e.g., of opioids, benzodiazepines) and evaluation of physical dependence/withdrawal has been required by FDA as part of an abuse liability evaluation. However, there are drugs and drug classes (e.g., selective serotonin reuptake inhibitors) that produce physical dependence but are not abused (Rosenbaum and Zajecka, 1997). Conversely, there are drugs of abuse that appear to produce only modest physical dependence or a protracted withdrawal syndrome but are used compulsively (e.g., cocaine). Thus, the development of physical dependence to a drug is neither necessary nor sufficient to conclude that the drug is likely to be abused.

Regardless of whether a drug is likely to be abused, it is important to determine whether the abrupt termination of the drug might lead to adverse effects. With regard to abuse liability assessment, an evaluation of physical dependence is important because withdrawal can lead to continued drug taking to avoid withdrawal symptoms. Avoidance of withdrawal symptoms appears to be an important mechanism underlying the compulsive use of some drugs such as opioids, nicotine, and caffeine. Given that the suppression of withdrawal effects might serve as an important negative reinforcement mechanism for the self-administration of some drugs, examination of how the reinforcing effects of a drug might change under conditions of withdrawal might be a valuable additional component of abuse liability assessment (e.g., Negus and Rice, 2009).

2.2.1.1. Experimental designs used in studies of physical dependence and withdrawal in non-human subjects: The probability that physical dependence will develop to a drug and the

severity of the subsequent withdrawal syndrome increase as a function of the dose, frequency of administration, and duration of treatment with that drug. However, studies of physical dependence can be time-consuming and expensive, and thus, it is usually not practical to evaluate a wide range of doses or dosing regimens in these studies. The selection of doses or frequency of administration that results in clear behavioral effects or supratherapeutic plasma levels of the drug is often considered to provide a sufficient assessment of a drug's potential for physical dependence. A pragmatic strategy is to administer high doses for an intermediate duration (e.g., 1 month). A longer duration of chronic treatment is thought to result in a more rigorous assessment of the potential for physical dependence than a shorter duration of chronic treatment (e.g., France et al., 2006).

The emergence of a withdrawal syndrome upon abrupt termination of chronic drug administration is referred to as spontaneous withdrawal and is thought to most closely model a withdrawal syndrome that might occur clinically with either medical or non-medical use. The pharmacokinetic characteristics of the drug being evaluated, including its elimination half-life and the presence of behaviorally active metabolites will affect the onset and duration of a spontaneous withdrawal syndrome. The experimental assessment of a spontaneous withdrawal syndrome should be based on the pharmacokinetics of the drug for the species that will be used, and should be of a long enough duration to include the biological elimination of the drug. Depending on the pharmacokinetics of the drug, long observation periods (e.g., up to one week or more) might be necessary after the abrupt discontinuation of a drug that is eliminated slowly. The frequency of observation should be greatest during the period immediately following drug discontinuation or, based on the pharmacokinetic profile, when the rate of elimination is highest.

If an antagonist for the receptor targeted by the test compound exists, withdrawal may be "precipitated" by administration of the antagonist. Antagonists at the mu opioid receptor (e.g., naloxone or naltrexone) or the benzodiazepine site on the GABA_A receptor (e.g., flumazenil) have been used extensively to study physical dependence to opioids and benzodiazepines, respectively (Young and Thompson, 1978; Gerak and France, 1999; McMahon et al., 2008). Precipitating withdrawal with an antagonist can afford greater experimenter control over the onset and the magnitude of withdrawal because withdrawal is not dependent on the metabolism and elimination of the chronically-administered drug.

2.2.1.2. Dependent measures used in studies of physical dependence and withdrawal:

Assessments of withdrawal should be as comprehensive as possible and include normal and abnormal behaviors as well as physiological measures. For drugs with a known mechanism of action, withdrawal behaviors and physiological effects known to be associated with the drug class should be assessed during discontinuation. Physiological measures that have been assessed in withdrawal studies include heart rate, blood pressure, rate of respiration, pupillary diameter, and body temperature. For drugs with a novel mechanism of action, a number of withdrawal behavior checklists derived for various classes of drugs might be useful in establishing which behaviors to assess after discontinuation (e.g., salivation, wet dog shakes, abnormal posture, tremor, rearing, yawning, scratching, vomiting, etc.). Unexpected behaviors might be observed during drug administration or discontinuation and video recording of animals during the study might be helpful. Having a video record of the study allows for the retrospective review of the video for new behaviors of interest, and also allows for repeated review by different observers who are blinded to the study conditions or trained to recognize specific behaviors. Such an approach can also be useful in assessing the accuracy and reliability of the behavioral scoring and providing information on subsequent behaviors of interest. In addition to observing changes in untrained behavior (e.g., rearing, scratching), studying changes in trained behavior (e.g., operant responding) after drug discontinuation broadens the scope of the assessment and can provide a more sensitive measure of a disruption of ongoing

behavior. Operant behavior (e.g., schedule-controlled responding, drug self-administration) is often extremely stable over long periods of time and is easily quantified. Operant behavior can also be readily assessed before, during, and after chronic drug treatment because the presence or absence of the behavior is not dependent on the presence or absence of physical dependence or withdrawal (Gerak and France, 1997; Ator et al., 2000; Negus and Rice, 2009).

3. Behavioral procedures for the assessment of the likelihood and consequences of abuse in human participants

Abuse liability assessments in human volunteers will often be more costly and time consuming than abuse liability assessments in non-human subjects. The current “gold standard” approach for initial human abuse liability testing of a novel compound is an acute dose-effect comparison study in volunteers with histories of drug abuse. The approach is thought to have high face validity and predictive validity of the likelihood of abuse by recreational drug abusers, and might be the only way to evaluate novel drug products such as abuse-deterrent formulations (see Griffiths et al., 2003; McColl and Sellers, 2006). Details of the experimental design and dependent measures assessed in the acute dose-effect abuse liability trial are described under the subsequent Sections 3.1. and 3.2.

Less well-developed alternatives to the classic acute dose-effect abuse liability trial are human drug discrimination and drug self-administration procedures. As with the non-human procedures, human drug discrimination procedures involve a period of initial acquisition of explicitly reinforced differential responding (i.e., drug discrimination) followed by a period of testing novel drug conditions. Although human drug discrimination procedures have been used to study various classes of abused drugs such as opioids, stimulants, and sedatives (Rush et al., 1995; Preston and Bigelow, 2000; Sevak et al., 2009), to our knowledge this approach has not been used to assess the abuse liability of novel compounds for regulatory purposes. That this approach requires time consuming training and does not provide a direct measure of reinforcing effects might account for its limited use for abuse liability assessment. With regard to human drug self-administration procedures, this methodology has been developed and employed to assess novel medications, particularly for therapeutics aimed to reduce drug taking behavior (see Comer et al., 2008; Hart et al., 2008; Haney and Spealman, 2008; Haney, 2009). As with the acute dose-effect studies, human self-administration studies often involve participants with histories of drug abuse. Human drug self-administration studies are usually less efficient to conduct than acute dose-effect studies because each dose that is assessed usually requires multiple sessions for initial exposure and subsequent assessment of self-administration. As with drug discrimination procedures, we are not aware that human drug self-administration procedures have ever been used for the assessment of relative abuse liability for regulatory purposes.

3.1. Experimental design

The most widely used procedure for assessing relative abuse liability is the classic acute dose-effect procedure, which is believed to have high predictive validity of the likelihood of abuse by recreational drug abusers. Most acute dose-effect abuse liability studies have used a similar experimental design, which includes within-subject, double-blind, placebo-controlled, administration of suprathreshold doses of drugs to recreational drug users. A sample size of about 14 participants will typically provide enough statistical power for comparisons between placebo and the novel drug condition; however, sample sizes of 20–40 participants might be necessary for making additional comparisons between different dose conditions (Griffiths et al., 2003; Schoedel and Sellers, 2008). In these studies, physiological, psychomotor, subjective, and cognitive (e.g., memory, attention) effects of a range of doses are characterized over the complete time course of the drug. Studies typically involve the assessment of several doses of

the novel compound compared to placebo and several doses of a positive control compound. The participant population selected for an abuse liability evaluation must be one in which the positive control comparison drug will test unequivocally positive. The most unambiguous results can be obtained from studies with participants who have extensive histories of polydrug abuse, including abuse of drugs from the same pharmacological class as the novel compound. There are several reasons for this. Participants with histories of polydrug abuse can use their prior drug abuse experiences with a variety of drugs and drug classes as a context from which to provide meaningful ratings of the drug experiences in the laboratory. This is thought to reduce the number of false positive and false negative results, measured by responses to placebo administration and known drugs of abuse, respectively (de Wit and Griffiths, 1991; Griffiths and Weerts, 1997). Also, participants with histories of drug abuse also provide the most face valid population for abuse liability assessment because they represent the population at greatest risk for illicit abuse of a novel compound.

In cases in which drug users with extensive histories of drug abuse are not readily available, it might be possible to conduct trials in less experienced volunteers by including a preliminary screening session (Busto et al., 1999). In this “qualifying” session, a test dose of a standard abused drug can be administered to volunteers. Volunteers who report positive effects (e.g., report liking) of the standard abused drug can then be enrolled into the abuse liability trial. Conducting a separate pilot study might be useful when relatively little information is available regarding the dose range or maximally tolerated dose of the novel drug in experienced drug users. Phase I and Phase II studies often provide little guidance for the selection of doses for an abuse liability trial because the clinical studies in a development program typically evaluate a narrow range of doses on a narrow range of measures in volunteers without extensive drug experience. Thus, a preliminary ascending dose run-up pilot study in participants with histories of drug abuse can be useful for selecting maximal doses and for matching doses between a novel compound and a comparator (Griffiths et al., 2003).

In most cases, drug abuse liability evaluations in recreational drug users should be conducted in controlled laboratory settings, which permit rigorous controlled assessment of outcome measures in the context of appropriate medical support while minimizing the risks and confounds of the use of other drugs. Ideally, and most often, such studies are conducted on a closed residential drug abuse pharmacology research unit, which provides a consistent routine and minimizes opportunities for obtaining drugs of abuse. Although such residential studies can also be conducted in other hospital settings such as a General Clinical Research Center, a lack of staff training in the management of these participants and a lack of structured visitation and inspection policies can be problematic. In some circumstances, studies can be conducted on an outpatient or ambulatory basis, with participants reporting to the clinical pharmacology laboratory and providing a drug-free urine sample that is tested at each session prior to drug administration (e.g., Schuh et al., 2000); however, the use of a dedicated inpatient unit is thought to reduce dropout rates, missed visits, and adverse events due to illnesses, accidents, and drug abuse, as well as help reduce variability with regard to sleep, nutrition, and drug use behaviors across participants.

Standard clinical pharmacology methods require that the participant and the staff who interact with the participant remain blind to the specific drug conditions administered on a given session (i.e., double-blind procedures). In addition, a placebo condition is included to control for effects of expectancy or accidental bias. Although ethical considerations and Institutional Review Boards require that some information about the drug conditions is communicated to the participants in the study (e.g., known adverse effects), specific details about the doses, number of active doses, and comparator compound(s) can remain unknown to the participants. Blinding can be enhanced, and the risk of expectancy biases can be minimized, by providing volunteers with more information than is applicable to the study in which they are participating (Griffiths

et al., 2003). For example, volunteers can be informed that the study will involve administration of one or more doses of the specific novel drug and, in addition, might involve administration of placebo and doses of a wide range of other mood altering drugs. The risks listed in the consent form may include the side effects of drugs that will not be administered in addition to the side effects of the drugs that will be studied.

Ideally, three or four doses of each drug are studied ranging from doses that have little or no effect to those that result in unequivocal subjective or behavioral effects. It should be recognized, however, that there might be cases in which doses of a CNS-acting drug many times larger than the high therapeutic dose appear to lack subjective and behavioral effects using traditional outcome measures (e.g., ramelteon; Johnson et al., 2006). Studying a range of doses of both the novel and the comparator drug(s) allows for the comparison of the slopes of the dose effect functions across different measures, which might be important for drawing conclusions regarding abuse liability (Carter et al., 2007). It is also essential to the validity of an abuse liability trial that a sufficiently high supratherapeutic dose of the novel compound is tested. With any new drug, it must be assumed that drug abusers will not be guided by the package insert in their selection of doses and will inevitably sample supratherapeutic doses of the drug. Thus, in determining the highest dose of the novel compound to be tested, the planned therapeutic dose of the compound is only marginally relevant (see Griffiths et al., 2003). Understandably, clinical investigators, industry sponsors, Institutional Review Boards, and FDA might be hesitant to approve the study of high doses for which limited safety data are available. Nonetheless, an adequate dose range is integral to a meaningful abuse liability evaluation and merits the extra precautions that might be required to gain approval for the protocol. Should a drug have untoward effects at higher doses, it is better to identify those effects early in the development process than fail to look for them and to risk patient health and liability from problems discovered through postmarketing surveillance.

The drug used as a positive control should have measurable abuse liability previously established through experimental studies and epidemiological data. If possible, the positive control should be an abused drug from the same pharmacological class and used for the same medical indication as that proposed for the novel compound. Within the study, the positive control compound should demonstrate dose-related statistically significant increases on the primary measures of abuse liability. Failure to demonstrate significant increases with the positive control drug invalidates the study. Given that speed of onset and duration of action can affect abuse liability (de Wit et al., 1993), interpretation of the results of an abuse liability study will be facilitated if the positive control and novel compound have a similar onset and duration of action. A drug that is quickly eliminated also allows for more frequent experimental sessions (e.g. daily sessions), thereby reducing the duration of the study.

3.2. Dependent measures in human abuse liability studies

Common dependent measures in a human abuse liability study include physiological measures, psychomotor performance tasks, self-report subjective effects, observer-rated measures of behavior, and cognitive (e.g., memory, attention, decision making) tasks. Ideally, the assessment procedures should be easily learned and rapidly completed by participants, and show good test-retest reliability so that measures can be assessed repeatedly within a session to track the time course of drug effects. Participants should be familiarized with and understand all measures, tasks, and questionnaires before administration of drug or placebo.

3.2.1. Dependent measures related to likelihood of abuse—Measures of drug liking (i.e. asking participants how much they like the drug) have face validity, have been used in most studies of abuse liability, and tend to be one of the most sensitive and reliable measures of likelihood of abuse (Griffiths et al., 2003; Schoedel and Sellers, 2008). Other participant

ratings that generally co-vary with liking include ratings of good effects, bad effects, degree to which you would like to take the drug again, estimations of the street value of the drug, and estimations of the amount of money the participant would personally be willing to pay for the drug (e.g., Sokolowska et al., 2008). In addition to assessing these measures repeatedly over the time course of drug action, the measures can also be assessed retrospectively after the drug effects have dissipated (e.g., in the form of a Next Day Questionnaire, given 24 h after drug administration). Retrospective ratings have the advantage of assessing the overall drug experience, or at least the remembered portion of that experience, under drug-free conditions and are thought to provide valuable indices of the likelihood that an individual, when sober and drug-free, would seek out an opportunity to re-administer the compound.

The Multiple-Choice Procedure was developed and validated as a tool to efficiently provide a behavioral measure of drug versus money choice in humans (e.g. Griffiths et al., 1993; Griffiths et al., 1996). With this procedure, after each session the participant makes a series of choices between receiving various amounts of money or receiving that dose of drug one more time in a final session at the end of the protocol. The data from the procedure is fundamentally different from simple participant ratings of liking or estimated street value because the choices have a tangible consequence to the participant in terms of money earned or drug received; typically, in the final study session, one of the participant's choices is randomly selected and the participant receives the consequence of that choice. The Multiple-Choice Procedure has been studied in the context of evaluating sedatives and stimulants (e.g., Lile et al., 2004), and a variation of the procedure also allows for the assessment of the punishing (i.e., aversive) effects of a drug (i.e., how much money a participant is willing to forfeit so as not to receive the drug again; Schuh et al., 2000).

Changes in subjective effects and mood might also have relevance to abuse liability. Adjective ratings scales are often used and may be grouped and scored as scales reflective of the symptom clusters commonly associated with prototypic effects of particular drug classes (e.g. sedation scale or an opioid agonist scale, Bigelow, 1991; Comer et al., 2008). Participants with extensive histories of recreational drug use can be asked to categorize the effects of the test drug as being most similar to one of several different classes of psychoactive drugs. This procedure provides information analogous to that from drug discrimination in preclinical studies, although in abuse liability studies (in contrast to human drug discrimination studies) human participants are not explicitly trained to recognize the effects of a drug or drug class.

The Profile of Mood States (POMS) questionnaire is a 65-item adjective rating scale that is considered to be a standardized mood state inventory with six subscales and a composite scale reflecting total mood disturbance (McNair et al., 1992). The POMS was shown to be sensitive to stimulants and sedatives in light drug users or nonusers and to intravenously administered cocaine in stimulant users (Foltin and Fischman, 1991), but has also been shown to be less sensitive than other measures of subjective effects in studies of sedative abusers (de Wit and Griffiths, 1991). The short form of the Addiction Research center Inventory (ARCI) is a 49 item true/false questionnaire with five empirically derived scales (Martin et al., 1971; Jasinski, 1977). The ARCI has been extensively used, both in studies with drug users and non-drug users, with the Morphine-Benzedrine Group (MBG) scale often cited as providing a measure of euphoria (Jasinski, 1977; Bigelow, 1991; Foltin and Fischman, 1991). However, the MBG scale was originally derived from a common pattern of items that were rated "true" after administration of morphine or benzedrine (Haertzen, 1966) and should not be assumed to be a valid predictor of abuse potential across other drug classes (de Wit and Griffiths, 1991).

3.2.2. Dependent measures related to consequences of abuse—Physiological measures are often considered a valuable supplemental measure in abuse liability studies because these measures are thought to be completely objective (e.g., pupillary diameter) and

often provide information on the safety profile of the drug (e.g., changes in blood pressure or blood oxygen saturation). Physiological measures can also be used to determine whether pharmacologically equivalent doses of the novel compound and positive controls were studied (e.g., Walsh et al., 2008). Similarly, behavioral and cognitive impairments represent well-recognized adverse effects of several classes of CNS drugs and behavioral and cognitive measures such as balance, hand-eye psychomotor speed and coordination, reaction time, divided attention, short-term or working memory, and long-term memory have been used in a number of abuse liability studies (e.g., Rush et al., 1999). Observer-rated measures of drug effect have also been included in most abuse liability studies, with the nature of the measures tailored to capture the expected effects of the class of drug being investigated (e.g., sedation and muscle-relaxation for anxiolytics and hypnotics; nodding and scratching for opioids; Walsh et al., 1995; Rush et al., 1999).

4. Procedures for the assessment of abuse liability during Phase I, II, and III clinical trials

Consideration should be given to incorporating measures that might be indicative of abuse liability in Phase I trials because Phase I trials are conducted relatively early in the clinical development process and the design of many Phase I trials are conducive to the addition of measures of abuse liability (see Brady et al., 2003; Schoedel and Sellers, 2008). Phase I trials often involve the testing of multiple doses of a drug under close medical supervision, which allows for an initial opportunity to collect data on the potential for abuse in human volunteers outside of a formal abuse liability study. Moreover, the psychomotor, subjective, and cognitive effects of a drug can be evaluated in healthy volunteers, without the potential confounds of a history of recreational drug use or the disease or disorder the drug is intended to treat. In addition, detailed information regarding the onset and duration of subjective effects can be collected during a Phase I trial, which might heighten or lessen concerns regarding the suspected abuse liability of the drug at a relatively early stage in the development process (see Mansbach et al., 2003 for review).

Many different measures could potentially be incorporated into Phase I clinical trials. Same day and next day measures of drug liking, good effects, bad effects, and the degree to which the participants would like to take the drug again (e.g., to relax or have fun) could be quickly and easily collected in a Phase I trial. However, it is important to note that most participants in Phase I trials will not have extensive histories of recreational drug use and therefore the risk of false positive and false negative reports from individuals without histories of recreational drug use might be greater than those from experienced recreational drug users. Thus, care must be taken to distinguish false positive and false negative signals from true positive and negative signals of abuse liability. This could be done by evaluating a positive control (i.e., drug with known abuse liability) in the Phase I population. For example, if a company was developing a compound for which there was concern that it might have stimulant-like abuse liability, then including amphetamine as a condition in a Phase I trial could be informative and useful for validating typical abuse liability measures in a healthy, non-drug using population.

In addition to providing information relevant to abuse liability for the purpose of de-risking a development program, prospective assessments of abuse liability in a Phase I trial might identify specific issues that should be more closely monitored in subsequent Phase II and III studies (e.g., positive subjective effects, reports of liking the drug). Data from all Phases of development contribute to the adverse event (AE) database and there are standard AE terms (e.g., euphoria, confusion, depersonalization) that could possibly be associated with abuse liability (Brady et al., 2003; Schoedel and Sellers, 2008). Spontaneous reports of these AE terms should be followed up with specific questions regarding the context in which the effects occurred and the qualitative nature of these effects. For example, a patient might report a

euphoria-like elevated mood after the therapeutic effect of the drug has alleviated his/her disorder or disease state. In the absence of a narrative, such a report could be recorded as an AE for euphoria, which would be misleading and a possible false positive with regard to abuse liability. It is also important to recognize that some AEs (e.g., sedation) are not invariably indicative of abuse liability. For example, with barbiturate-like sedative hypnotics, sedation often positively covaries with likelihood of abuse while this is not the case with antipsychotics. Thus, narratives about the AE could be critical to interpreting a possible signal of abuse potential.

Phase II trials are conducted in individuals with the disease or disorder that the drug is intended to treat, and therefore provide an opportunity to collect information relevant to abuse liability in the population that will be most frequently exposed to the drug. In addition to studying relevant subjective effects ratings (e.g., drug liking and good and bad effects), Phase II trials often involve the administration of the drug over a period of several weeks, which potentially allows one to study compliance, dose escalation, tolerance, and drug discontinuation. For example, participants can be given the flexibility to titrate their dose within a pre-determined, therapeutically-recommended range of doses (see Bigelow et al., 1980). Such a design is likely to more closely model the real world use of medications and also allows for the prospective monitoring of dose escalation and possible misuse. Again, care must be taken to document the context and rationale for a patient's titration to larger doses such that titration up for therapeutic efficacy is not misinterpreted as abuse liability. However, participants in a trial with an open titration schedule might be more likely to report increased use of the drug (within the accepted dose range) or other early signals of misuse because they could do so without admitting to non-compliance and risk being discontinued from the study.

Phase III clinical trials are also conducted in individuals with the disease or disorder that the drug is intended to treat and include a larger number of participants than Phase I or II clinical trials. This increases the likelihood of observing AEs that will affect a small proportion of patients; however, the inclusion/exclusion criteria for many Phase III clinical trials would typically exclude individuals with a history or current diagnosis of substance abuse or dependence. Nonetheless, medication compliance and spontaneous AE reporting are performed in these trials and can be tailored to closely monitor effects that are anticipated based on the results of the Phase I and II trials. Phase III trials also afford the possibility to study the possible emergence of withdrawal signs and symptoms upon termination of drug administration. This can be done at the end of a long-term study, or might involve a second randomization of patients to different treatment conditions that includes a placebo condition (e.g., U.S. Xyrem Multi-Center Study Group, 2003; U.S. Xyrem Multi-Center Study Group, 2004). The prospective monitoring of specific signs and symptoms of withdrawal (that have likely been identified and refined during preclinical and Phase I and II studies), can provide valuable information for manufacturers, physicians, and patients on the risks associated with abrupt discontinuation of a particular drug and how that drug should be gradually discontinued to prevent the emergence of AEs in clinical practice.

Although patients with histories or current drug abuse are often excluded from Phase I, II, and III clinical trials, drug abuse is frequently co-morbid with other psychiatric (e.g., depression, anxiety) and non-psychiatric (e.g., HIV) medical conditions (Chander et al., 2006). Thus, it is likely that patients with histories or current drug abuse who would have been excluded from clinical trials during development will be exposed to the new drugs in clinical practice. As such, consideration should be given to the conduct of abuse liability studies in individuals with a history of drug abuse and the co-morbid disorder that the drug in development is intended to treat. The results of such studies might appropriately confirm or dispel fears that there are different risks (perhaps necessitating different risk mitigation strategies) associated with prescribing a drug to patients with a history or current drug abuse.

5. Discussion

Abuse liability assessment provides critical information that is used in making decisions with regard to drug development, clinical use, and regulatory processes, including the implementation of risk management and postmarketing surveillance strategies. The most commonly used methodologies for the assessment of abuse liability for regulatory purposes include the evaluation of discriminative stimulus effects, reinforcing effects, and the potential for physical dependence in non-human subjects (Ator and Griffiths, 2003; Haney and Spealman, 2008), and the evaluation of subjective effects and measures or proxies of drug taking behavior in human participants (Griffiths et al., 2003; Comer et al., 2008). This review provides a general overview of the typical experimental designs and dependent measures employed in these procedures, and offers several suggestions for the incorporation of abuse liability measures into industry-sponsored Phase I, II, and III clinical trials. The inclusion of abuse liability measures in clinical trials might provide data that will complement or even preempt a formal human abuse liability trial, potentially speeding the process of drug development and allowing for the early planning and implementation of risk management strategies for a drug development program.

Over the last several decades validated methodologies for assessing the likelihood/severity of abuse and likelihood/severity of the consequences of abuse in non-human and human subjects in a laboratory setting have been developed (see Ator and Griffiths, 2003; Balster and Bigelow, 2003; Griffiths et al., 2003; Comer et al., 2008; Haney and Spealman, 2008). Moreover, these procedures have been shown to have good internal validity and predictive validity as drugs that are self-administered by non-human subjects are liked and self-administered by human recreational drug users (Griffiths and Balster, 1979; Griffiths and Ator, 1980; Panlilio and Goldberg, 2007; Haney and Spealman, 2008), and drugs that are liked and self-administered by human recreational drug users in the laboratory tend to be used and abused recreationally outside of the laboratory setting (Comer et al., 2008; Schoedel and Sellers, 2008; Haney, 2009).

Although the current preclinical and clinical methodology for abuse liability testing has been shown to have good predictive validity, a number of challenges remain that might stimulate the refinement or development of procedures for future studies. One challenge is the interpretation of “weak signals” or intermediate levels of drug-appropriate responding, self-administration, or subjective effects measures. In these cases, additional pharmacological and behavioral analyses might aid in the interpretation of the data. For example, if an intermediate level of drug-appropriate responding, drug self-administration, or subjective effects measures is receptor-mediated, administration of an antagonist (if available) should shift the dose effect curve rightward; if an intermediate level of drug-appropriate responding, drug self-administration, or subjective effects measures is a result of partial agonist activity, the partial agonist should attenuate the effects of a high efficacy agonist (Walsh et al., 1996).

Intermediate levels of responding in non-human and human studies might result from a non-specific impairment or disruption of behavior, or mixed positive and negative effects, respectively (e.g., Koek et al., 1993; Mumford et al., 1995b). Thus, the analysis of additional behavioral variables might also provide insight into the nature of an intermediate result. In drug discrimination studies, additional behavioral measures could include the latency to select a response option or the amount of responding on the non-selected option, which have allowed investigators to conclude that an intermediate response more closely resembled a non-specific disruption of behavior rather than partial generalization (Koek et al., 1993). In human studies, the evaluation of measures of bad effects and disliking, both during and after the period of drug effect, have been helpful in interpreting the intermediate effects of the drug on measures of good effects and liking (Mumford et al., 1995b; Carter et al., 2006).

There have also been discrepancies between drugs that are self-administered by non-human subjects and liked by human participants in the laboratory, but remain widely available and are not widely abused (e.g., diphenhydramine; Preston et al., 1992; Mumford et al., 1996). In some cases, the apparent discrepancies might be due to differences in the range of doses, route of administration, or experimental parameters between the laboratory assessments and real world abuse (cf., Woolverton et al., 1999 and Weerts et al., 2008); however, drug availability, ease of diversion, economic, and socio-cultural factors also influence patterns of drug abuse outside of the laboratory.

It is also important to recognize that the likelihood and consequences of abuse of a drug are affected by a wide range of non-pharmacological and pharmacological factors. For example, physical characteristics such as formulation (e.g., solid tablet, rapid-dissolving thin film, liquid solution, gas) might affect the desirability and ease by which a product is abused or diverted. Pharmacological properties of a product such as the availability of a large dose in an extended-release formulation that can be extracted or the absence of another potentially noxious ingredient (e.g., acetaminophen or an antagonist) might affect the relative attractiveness of a product (Fudala and Johnson, 2006). Economic factors might also play an important role; the actual abuse of a drug with high abuse liability might be low if there is an inexpensive and readily available alternative. Conversely, actual abuse of a drug with relatively low abuse liability might be high if alternatives are expensive or unavailable. Cultural and social factors such as the association of the drug with a particular activity (e.g., with sex) or the potential profitability of the illicit sale of the medication (e.g., HIV antiretroviral drugs) might affect a product's risk for abuse and diversion (Inciardi et al., 2007; Dart, 2009). In addition, some aspects of the attractiveness of a drug product might be difficult or impossible to evaluate in the laboratory setting. For example, publicized use of a drug by a celebrity, professional athlete, or other high-profile individual could influence the attractiveness of the drug for some individuals. Therefore, a positive abuse liability signal for a new medication indicates that the drug has potential for abuse, but not necessarily that the drug will be widely abused (e.g., tramadol; Epstein et al., 2006; Schoedel and Sellers, 2008).

One of the limitations of the procedures described in this review is that the laboratory-based procedures do not account for the social, cultural, and economic factors that influence drug abuse, and therefore cannot accurately predict the magnitude of the problems that a drug might cause. As a result, data from the procedures described in this review may be considered to be categorical in nature from a regulatory point of view. That is to say that the data might suggest that abuse liability is a concern that will require some form of regulation or control, or that abuse liability is not a concern and strict regulation and control are unwarranted. For cases in which intermediate results are observed, there might be a regulatory middle ground in which the drug could be made available with few restrictions for an initial probationary period of time during which post-marketing data could be gathered and appropriate risk mitigation strategies could be developed (e.g., tramadol; Cicero et al., 1999).

The recent increasing focus on the development of new formulations and molecules that are designed to be abuse-deterrent or tamper-resistant will require the comparative abuse liability evaluation of new products to existing products (Grudzinskas et al., 2006). Methods that have traditionally been used to compare drugs to placebo might require modifications to the design (e.g., a greater number of subjects required for greater statistical power) or procedures (e.g., greater reliance on choice procedures) for making comparisons between drugs or drug formulations. An approach that has been proposed as an alternative to the traditional measures associated with relative reinforcing efficacy is the examination of the elasticity of demand that is derived from behavioral economic studies (Bickel et al., 2000; Johnson and Bickel, 2006). The relationship between two reinforcers can be examined by comparing the characteristics of their demand curves and such an approach allows for comparisons to be made across

commodities (e.g., between drug and money or between different drug products; Hursh and Winger, 1995; Bickel et al., 2000; Johnson and Bickel, 2006). The development of such procedures to adequately compare different drugs and drug formulations is critically important for the evaluation of claims that a particular drug product has lower liability for abuse.

Drug safety and risks associated with drug-drug interactions have also received increasing attention. For drugs with known or suspected liability for abuse, drug interaction studies with other drugs of abuse (e.g., ethanol) might be warranted within the context of a new drug application. However, given the virtually unlimited number of possible drug interaction studies that might be proposed, it is important that the regulatory necessity of each drug interaction study is supported by rationale related to a specific public health risk. From an industry perspective, the effects of known or suspected drug-drug interactions on specific endpoints can be examined in non-human subjects at early stages in the drug development process. Isobolographic analyses allow for a statistically-based assessment of additive, antagonistic, and synergistic effects of dose combinations (see Tallarida, 2007 for review).

Although drug abusers are thought to be at greatest risk for abusing new drugs with abuse liability, there are compelling reasons to study the abuse related effects of new therapeutics in populations such as the intended therapeutic population, which will be most frequently exposed to the drug (e.g., Houtsmuller et al., 2002). The presence of an abuse liability signal in a drug abusing population together with the absence of an abuse liability signal in the therapeutic population might suggest that diversion, and not patient misuse, will be the most important area of concern (Wright et al., 2006). This information could guide the implementation of more stringent strategies to prevent diversion without enacting measures that limit patient access to the medication. A positive abuse liability signal in a patient population would suggest that greater patient education or medical supervision might be appropriate.

In addition to studying the effects of a novel drug in drug abusers in an abuse liability trial and in patients for whom the drug is intended to treat in Phase II and III studies, there are also likely to be patients with co-morbid substance use disorders who exhibit characteristics that overlap with both populations, but who are excluded from either type of study due to their substance abuse (a common exclusion from Phase II and III trials) or due to their co-morbid illness (a common exclusion from abuse liability trials). Patients with the illness the drug is intended to treat and a co-morbid substance use disorder might represent the group of individuals at highest risk for abuse and diversion because in addition to having a substance use disorder (and likely having friends or relatives with substance use disorders), these individuals will potentially obtain legitimate prescriptions for the drug. Studying the abuse liability of a drug in a patient population with co-morbid substance use disorders could be useful for examining medication compliance, medication tampering, and risk of diversion. Moreover, studies conducted in these patients might identify patient characteristics that are predictive of abuse or diversion in the broader patient population and could inform the development of differential prescribing practices (Dasgupta and Schnoll, 2009).

The implementation of well-designed pharmacovigilance and risk management strategies can substantially protect against the possibility that a drug with abuse liability is ever widely abused (Woody et al., 2003; Fuller et al., 2004; Wedin et al., 2006; Katz et al., 2007). Given that excessive regulation and restrictions can hinder the development and clinical use of safe and effective medications (Wright et al., 2006), efforts toward enhanced patient and physician education, pharmacovigilance, and postmarketing surveillance to prevent the diversion and misuse of drugs with abuse liability are likely to foster both the protection and the promotion of public health.

References

- Ator NA, Griffiths RR. Principles of drug abuse liability assessment in laboratory animals. *Drug Alcohol Depend* 2003;70:S55–S72. [PubMed: 12759197]
- Ator NA, Weerts EM, Kaminski BJ, Kautz MA, Griffiths RR. Zaleplon and triazolam physical dependence assessed across increasing doses under a once-daily dosing regimen in baboons. *Drug Alcohol Depend* 2000;61:69–84. [PubMed: 11064185]
- Balster RL, Bigelow GE. Guidelines and methodological reviews concerning drug abuse liability assessment. *Drug Alcohol Depend* 2003;70:S13–40. [PubMed: 12759195]
- Balster RL, Schuster CR. Fixed-interval schedule of cocaine reinforcement: effect of dose and infusion duration. *J Exp Anal Behav* 1973;20:119–129. [PubMed: 4197505]
- Bardo MT, Bevins RA. Conditioned place preference: what does it add to our preclinical understanding of drug reward? *Psychopharmacology* 2000;153:31–43. [PubMed: 11255927]
- Bickel WK, Marsch LA, Carroll ME. Deconstructing relative reinforcing efficacy and situating the measures of pharmacological reinforcement with behavioral economics: a theoretical proposal. *Psychopharmacology* 2000;153:44–56. [PubMed: 11255928]
- Bigelow GE. Human drug abuse liability assessment: opioids and analgesics. *Br J Addict* 1991;86:1615–1628. [PubMed: 1686196]
- Bigelow GE, Griffiths RR, Liebson R, Kaliszak JE. Double-blind evaluation of reinforcing and anorectic actions of weight control medications. *Arch Gen Psychiatry* 1980;37:1118–1123. [PubMed: 7000029]
- Brady KT, Lydiard RB, Brady JV. Assessing abuse liability in clinical trials. *Drug Alcohol Depend* 2003;70:S87–95. [PubMed: 12759199]
- Busto UE, Zawertailo LA, Kaplan HL, Sellers EM. Identifying appropriate subjects for abuse liability studies using prestudy pharmacological testing. *Can J Clin Pharmacol* 1999;6:103–110. [PubMed: 10519736]
- Carroll ME, Stotz DC, Kliner DJ, Meisch RA. Self-administration of orally-delivered methohexital in rhesus monkeys with phencyclidine or pentobarbital histories: effects of food deprivation and satiation. *Pharmacol Biochem Behav* 1984;20:145–151. [PubMed: 6694994]
- Carter LP, Flores LR, Wu H, Chen W, Unzeitig AW, Coop A, France CP. The role of GABAB receptors in the discriminative stimulus effects of gamma-hydroxybutyrate in rats: time course and antagonism studies. *J Pharmacol Exp Ther* 2003;305:668–674. [PubMed: 12606639]
- Carter LP, Griffiths RR, Suess PE, Casada JH, Wallace CL, Roache JD. Relative abuse liability of indiplon and triazolam in humans: a comparison of psychomotor, subjective, and cognitive effects. *J Pharmacol Exp Ther* 2007;322:749–759. [PubMed: 17502431]
- Carter LP, Richards BD, Mintzer MZ, Griffiths RR. Relative abuse liability of GHB in humans: A comparison of psychomotor, subjective, and cognitive effects of supratherapeutic doses of triazolam, pentobarbital, and GHB. *Neuropsychopharmacology* 2006;31:2537–2551. [PubMed: 16880774]
- Carter LP, Unzeitig AW, Wu H, Chen W, Coop A, Koek W, France CP. The Discriminative stimulus effects of gamma-hydroxybutyrate and related compounds in rats discriminating baclofen or diazepam: The role of GABA(B) and GABA(A) receptors. *J Pharmacol Exp Ther* 2004;309:540–547. [PubMed: 14742739]
- Cicero TJ, Adams EH, Geller A, Inciardi JA, Muñoz A, Schnoll SH, Senay EC, Woody GE. A postmarketing surveillance program to monitor Ultram (tramadol hydrochloride) abuse in the United States. *Drug Alcohol Depend* 1999;57:7–22. [PubMed: 10617309]
- Collins GT, Woods JH. Drug and reinforcement history as determinants of the response-maintaining effects of quinpirole in the rat. *J Pharmacol Exp Ther* 2007;323:599–605. [PubMed: 17675585]
- Colpaert FC. Drug discrimination: no evidence for tolerance to opiates. *Pharmacol Rev* 1995;47:605–629. [PubMed: 8746556]
- Comer SD, Ashworth JB, Foltin RW, Johanson CE, Zacny JP, Walsh SL. The role of human drug self-administration procedures in the development of medications. *Drug Alcohol Depend* 2008;96:1–15. [PubMed: 18436394]
- Czoty PW, Reboussin BA, Calhoun TL, Nader SH, Nader MA. Long-term cocaine self-administration under fixed-ratio and second-order schedules in monkeys. *Psychopharmacology* 2007;191:287–295. [PubMed: 17216155]

- Dart RC. Monitoring risk: Post marketing surveillance and signal detection. *Drug Alcohol Depend.* 2009;in press this issue
- Dasgupta N, Schnoll SH. Signal detection in post-marketing surveillance for controlled substances. *Drug Alcohol Depend.* 2009;in press this issue
- de Wit H, Dudish S, Ambre J. Subjective and behavioral effects of diazepam depend on its rate of onset. *Psychopharmacology* 1993;112:324–330. [PubMed: 7871037]
- de Wit H, Griffiths RR. Testing the abuse liability of anxiolytics and hypnotic drugs in humans. *Drug Alcohol Depend* 1991;28:83–111. [PubMed: 1679388]
- Epstein DH, Preston KL, Jasinski DR. Abuse liability, behavioral pharmacology, and physical-dependence potential of opioids in humans and laboratory animals: lessons from tramadol. *Biol Psychol* 2006;73:90–99. [PubMed: 16497429]
- Fantegrossi WE, Harrington AW, Eckler JR, Arshad S, Rabin RA, Winter JC, Coop A, Rice KC, Woods JH. Hallucinogen-like actions of 2,5-dimethoxy-4-(n)-propylthiophenethylamine (2C-T-7) in mice and rats. *Psychopharmacology* 2005;181:496–503. [PubMed: 15983786]
- Foltin RW, Fischman MW. Assessment of abuse liability of stimulant drugs in humans: a methodological survey. *Drug Alcohol Depend* 1991;28:3–48. [PubMed: 1679387]
- France CP, Weltman RH, Koek W, Cruz CM, McMahon LR. Acute and chronic effects of ramelteon in rhesus monkeys (*Macaca mulatta*): Dependence liability studies. *Behav Neurosci* 2006;120:535–541. [PubMed: 16768605]
- Fudala PJ, Johnson RE. Development of opioid formulations with limited diversion and abuse potential. *Drug Alcohol Depend* 2006;83:S40–S47. [PubMed: 16564141]
- Fuller DE, Hornfeldt CS, Kelloway JS, Stahl PJ, Anderson TF. The Xyrem risk management program. *Drug Saf* 2004;27:293–306. [PubMed: 15061684]
- Gerak LR, France CP. Repeated administration of flumazenil does not alter its potency in modifying schedule-controlled behavior in chlordiazepoxide-treated rhesus monkeys. *Psychopharmacology* 1997;131:64–70. [PubMed: 9181637]
- Gerak LR, France CP. Discriminative stimulus effects of flumazenil in untreated and in diazepam-treated rhesus monkeys. *Psychopharmacology* 1999;146:252–261. [PubMed: 10541724]
- Goldberg SR, Hoffmeister F, Schlichting UU, Wuttke W. A comparison of pentobarbital and cocaine self-administration in rhesus monkeys: effects of dose and fixed-ratio parameter. *J Pharmacol Exp Ther* 1971;179:277–283. [PubMed: 5002482]
- Griffiths RR, Ator NA. Benzodiazepine self-administration in animals and humans: a comprehensive literature review. *NIDA Res Monogr* 1980;33:22–36. [PubMed: 6798448]
- Griffiths RR, Balster RL. Opioids: similarity between evaluations of subjective effects and animal self-administration results. *Clin Pharmacol Ther* 1979;25:611–617. [PubMed: 108042]
- Griffiths RR, Bigelow GE, Ator NA. Principles of initial experimental drug abuse liability assessment in humans. *Drug Alcohol Depend* 2003;70:S41–54. [PubMed: 12759196]
- Griffiths, RR.; Bigelow, GE.; Henningfield, JE. Similarities in animal and human drug-taking behavior. In: Mello, NK., editor. *Advances in Substance Abuse*. Greenwich, CT: JAI Press, Inc; 1980. p. 1-90.
- Griffiths RR, Lukas SE, Bradford LD, Brady JV, Snell JD. Self-injection of barbiturates and benzodiazepines in baboons. *Psychopharmacology* 1981;75:101–109. [PubMed: 6119723]
- Griffiths RR, Rush CR, Puhala K. Validation of the multiple-choice procedure for investigating drug reinforcement in humans. *Exp Clin Psychopharmacol* 1996;4:97–106.
- Griffiths RR, Troisi JR II, Silverman K, Mumford GK. Multiple-choice procedure: an efficient approach for investigating drug reinforcement in humans. *Behav Pharmacol* 1993;4:3–13. [PubMed: 11224166]
- Griffiths RR, Weerts EM. Benzodiazepine self-administration in humans and laboratory animals – implications for problems of long-term use and abuse. *Psychopharmacology* 1997;134:1–37. [PubMed: 9399364]
- Grudzinskas C, Balster RL, Gorodetzky CW, Griffiths RR, Henningfield JE, Johanson CE, Mansbach RS, McCormick CG, Schnoll SH, Strain EC, Wright C. Impact of formulation on the abuse liability, safety and regulation of medications: the expert panel report. *Drug Alcohol Depend* 2006;83(Suppl 1):S77–82. [PubMed: 16580154]

- Haertzen CA. Development of scales based on patterns of drug effects, using the Addiction Research Center Inventory (ARCI). *Physiological Reports* 1966;18:163–194.
- Haney M. Self-administration of cocaine, cannabis and heroin in the human laboratory: benefits and pitfalls. *Addict Biol* 2009;14:9–21. [PubMed: 18855806]
- Haney M, Spealman R. Controversies in translational research: drug self-administration. *Psychopharmacology* 2008;199:403–419. [PubMed: 18283437]
- Hart CL, Haney M, Vosburg SK, Rubin E, Foltin RW. Smoked cocaine self-administration is decreased by modafinil. *Neuropsychopharmacology* 2008;33:761–768. [PubMed: 17568397]
- Houtsmuller EJ, Fant RV, Eissenberg TE, Henningfield JE, Stitzer ML. Flavor improvement does not increase abuse liability of nicotine chewing gum. *Pharmacol Biochem Behav* 2002;72:559–568. [PubMed: 12175452]
- Hursh SR, Winger G. Normalized demand for drugs and other reinforcers. *J Exp Anal Behav* 1995;64:373–384. [PubMed: 8551194]
- Inciardi JA, Surratt HL, Kurtz SP, Cicero TJ. Mechanisms of prescription drug diversion among drug-involved club- and street-based populations. *Pain Med* 2007;8:171–183. [PubMed: 17305688]
- Jasinski, DR. Assessment of the abuse potentiality of morphine-like drugs (methods used in man). In: Martin, WR., editor. *Drug Addiction I. Morphine, Sedative-Hypnotic and Alcohol Dependence Handbook of Experimental Pharmacology*. Vol. 45. Springer; Heidelberg: 1977. p. 197–258.
- Johnson MW, Bickel WK. Replacing relative reinforcing efficacy with behavioral economic demand curves. *J Exp Anal Behav* 2006;85:73–93. [PubMed: 16602377]
- Johnson MW, Suess PE, Griffiths RR. Ramelteon: a novel hypnotic lacking abuse liability and sedative adverse effects. *Arch Gen Psychiatry* 2006;63:1149–1157. [PubMed: 17015817]
- Justinova Z, Tanda G, Redhi GH, Goldberg SR. Self-administration of delta9-tetrahydrocannabinol (THC) by drug naive squirrel monkeys. *Psychopharmacology* 2003;169:135–140. [PubMed: 12827345]
- Katz NP, Adams EH, Benneyan JC, Birnbaum HG, Budman SH, Buzzeo RW, Carr DB, Cicero TJ, Gourlay D, Inciardi JA, Joranson DE, Kesslick J, Lande SD. Foundations of opioid risk management. *Clin J Pain* 2007;23:103–118. [PubMed: 17237659]
- Koek W, Carter LP, Lamb RJ, Chen W, Wu H, Coop A, France CP. Discriminative stimulus effects of gamma-hydroxybutyrate (GHB) in rats discriminating GHB from baclofen and diazepam. *J Pharmacol Exp Ther* 2005;314:170–179. [PubMed: 15774716]
- Koek W, Colpaert FC, Vignon J. Effects of phencyclidine-type drugs in rats discriminating fentanyl from saline: pharmacological and behavioral characterization of intermediate levels of drug lever selection. *J Pharmacol Exp Ther* 1993;264:746–756. [PubMed: 8437123]
- Leiderman DB. Risk Management of drug products and the FDA: Evolution and Context. *Drug Alcohol Depend*. 2009 in press this issue
- Lile JA, Stoops WW, Glaser PE, Hays LR, Rush CR. Acute administration of the GABA reuptake inhibitor tiagabine does not alter the effects of oral cocaine in humans. *Drug Alcohol Depend* 76:81–91. [PubMed: 15380292]
- Mansbach RS, Feltner DE, Gold LH, Schnoll SH. Incorporating the assessment of abuse liability into the drug discovery and development process. *Drug Alcohol Depend* 2003;70:S73–85. [PubMed: 12759198]
- Martin WR, Sloan JW, Sapiro JD, Jasinski DR. Physiologic, subjective, and behavioral effects of amphetamine, methamphetamine, ephedrine, phenmetrazine and methylphenidate in man. *Clin Pharmacol Ther* 1971;12:245–258. [PubMed: 5554941]
- McCabe SE, Cranford JA, West BT. Trends in prescription drug abuse and dependence, co-occurrence with other substance use disorders, and treatment utilization: Results from two national surveys. *Addict Behav* 2008;33:1297–1305. [PubMed: 18632211]
- McColl S, Sellers EM. Research design strategies to evaluate the impact of formulations on abuse liability. *Drug Alcohol Depend* 2006;83:S52–62. [PubMed: 16554125]
- McMahon LR, Li JX, Carroll FI, France CP. Some effects of dopamine transporter and receptor ligands on discriminative stimulus, physiologic, and directly observable indices of opioid withdrawal in rhesus monkeys. *Psychopharmacology* 2009;203:411–420. [PubMed: 18636243]

- McMillan DE, Sun WL, Hardwick WC. Effects of drug discrimination history on the generalization of pentobarbital to other drugs. *J Pharmacol Exp Ther* 1996;278:50–61. [PubMed: 8764335]
- McNair, DM.; Lorr, M.; Droppleman, LF. Edits manual for the profile of mood states. Edits/Educational and Industrial Testing Service; San Diego, CA: 1992.
- Meisch RA, Bell SM, Lemaire GA. Orally self-administered cocaine in rhesus monkeys: transition from negative or neutral behavioral effects to positive reinforcing effects. *Drug Alcohol Depend* 1993;32:143–158. [PubMed: 8508725]
- Mumford GK, Evans SM, Fleishaker JC, Griffiths RR. Alprazolam absorption kinetics affects abuse liability. *Clin Pharmacol Ther* 1995a;57:356–365. [PubMed: 7697954]
- Mumford GK, Rush CR, Griffiths RR. Abecarnil and alprazolam in humans: behavioral, subjective and reinforcing effects. *J Pharmacol Exp Ther* 1995b;272:570–580. [PubMed: 7853170]
- Mumford GK, Silverman K, Griffiths RR. Reinforcing, subjective, and performance effects of lorazepam and diphenhydramine in humans. *Exp Clin Psychopharmacol* 1996;4:421–430.
- Negus SS, Rice KC. Mechanisms of withdrawal-associated increases in heroin self-administration: pharmacologic modulation of heroin vs food choice in heroin-dependent rhesus monkeys. *Neuropsychopharmacology* 2009;34:899–911. [PubMed: 18704098]
- Panlilio LV, Goldberg SR. Self-administration of drugs in animals and humans as a model and an investigative tool. *Addiction* 2007;102:1863–1870. [PubMed: 18031422]
- Preston KL, Bigelow GE. Effects of agonist-antagonist opioids in humans trained in a hydromorphone/not hydromorphone discrimination. *J Pharmacol Exp Ther* 2000;295:114–124. [PubMed: 10991968]
- Preston KL, Wolf B, Guarino JJ, Griffiths RR. Subjective and behavioral effects of diphenhydramine, lorazepam and methocarbamol: evaluation of abuse liability. *J Pharmacol Exp Ther* 1992;262:707–720. [PubMed: 1501118]
- Rosenbaum JF, Zajecka J. Clinical management of antidepressant discontinuation. *J Clin Psychiatry* 1997;58:37–40. [PubMed: 9219493]
- Rowlett JK, Lelas S, Spealman RD. Transduction of the discriminative stimulus effects of zolpidem by GABA(A)/alpha1 receptors. *Eur J Pharmacol* 2000;406:R9–R10. [PubMed: 11011051]
- Rowlett JK, Spealman RD, Lelas S. Discriminative stimulus effects of zolpidem in squirrel monkeys: comparison with conventional benzodiazepines and sedative-hypnotics. *J Pharmacol Exp Ther* 1999;291:1233–1241. [PubMed: 10565847]
- Rush CR, Critchfield TS, Troisi JR, Griffiths RR. Discriminative stimulus effects of diazepam and buspirone in normal volunteers. *J Exp Anal Behav* 1995;63:277–294. [PubMed: 7751833]
- Rush CR, Frey JM, Griffiths RR. Zaleplon and triazolam in humans: acute behavioral effects and abuse potential. *Psychopharmacology* 1999;145:39–51. [PubMed: 10445371]
- Schoedel KA, Sellers EM. Assessing abuse liability during drug development: changing standards and expectations. *Clin Pharmacol Ther* 2008;83:622–626. [PubMed: 18212799]
- Schuh LM, Schuster CR, Hopper JA, Mendel CM. Abuse liability assessment of sibutramine, a novel weight control agent. *Psychopharmacology* 2000;147:339–346. [PubMed: 10672626]
- Schuster CR, Johanson CE. Relationship between the discriminative stimulus properties and subjective effects of drugs. *Psychopharmacol Ser* 1988;4:161–175. [PubMed: 3293041]
- Schuster CR, Thompson T. Self administration of and behavioral dependence on drugs. *Annu Rev Pharmacol* 1969;9:483–502. [PubMed: 4978013]
- Sevak RJ, Stoops WW, Hays LR, Rush CR. Discriminative stimulus and subject-rated effects of methamphetamine, d-amphetamine, methylphenidate, and triazolam in methamphetamine-trained humans. *J Pharmacol Exp Ther* 2009;328:1007–1018. [PubMed: 19104030]
- Sokolowska, M.; Setnik, B.; Jones, J.; Johnson, F.; Romach, M.; Sellers, E. Validation of assessments for abuse liability trials. Presented at the College on Problems of Drug Dependence 70th Annual Meeting; 2008 [(abstract accessed Aug 8, 2008)]. <<http://www.cpdd.vcu.edu/>>
- Tallarida RJ. Interactions between drugs and occupied receptors. *Pharmacol Ther* 2007;113:197–209. [PubMed: 17079019]
- Tanda G, Munzar P, Goldberg SR. Self-administration behavior is maintained by the psychoactive ingredient of marijuana in squirrel monkeys. *Nat Neurosci* 2000;3:1073–1074. [PubMed: 11036260]

- Thompson T, Schuster CR. Morphine self-administration, food-reinforced, and avoidance behaviors in rhesus monkeys. *Psychopharmacologia* 1964;5:87–94. [PubMed: 14137126]
- Thomsen M, Caine SB. Cocaine self-administration under fixed and progressive ratio schedules of reinforcement: comparison of C57BL/6J, 129X1/SvJ, and 129S6/SvEvTac inbred mice. *Psychopharmacology* 2006;184:145–154. [PubMed: 16369835]
- US Xyrem Multi-Center Study Group. The abrupt cessation of therapeutically administered sodium oxybate (GHB) does not cause withdrawal symptoms. *J Toxicol Clin Toxicol* 2003;41:131–135. [PubMed: 12733850]
- US Xyrem Multicenter Study Group. Sodium oxybate demonstrates long-term efficacy for the treatment of cataplexy in patients with narcolepsy. *Sleep Med* 2004;5:119–123. [PubMed: 15033130]
- Walsh SL, Nuzzo PA, Lofwall MR, Holtman JR Jr. The relative abuse liability of oral oxycodone, hydrocodone and hydromorphone assessed in prescription opioid abusers. *Drug Alcohol Depend* 2008;98:191–202. [PubMed: 18606504]
- Walsh SL, Preston KL, Bigelow GE, Stitzer ML. Acute administration of buprenorphine in humans: partial agonist and blockade effects. *J Pharmacol Exp Ther* 1995;274:361–372. [PubMed: 7542336]
- Walsh SL, Sullivan JT, Preston KL, Garner JE, Bigelow GE. Effects of naltrexone on response to intravenous cocaine, hydromorphone and their combination in humans. *J Pharmacol Exp Ther* 1996;279:524–538. [PubMed: 8930154]
- Wedin GP, Hornfeldt CS, Ylitalo LM. The Clinical Development of γ -Hydroxybutyrate (GHB). *Current Drug Safety* 2006;1:99–106. [PubMed: 18690919]
- Weeks JR. Experimental morphine addiction. Method for automatic intravenous injections in unrestrained rats. *Science* 1962;138:143–144. [PubMed: 14005543]
- Weerts EM, Ator NA, Griffiths RR. Comparison of the intravenous reinforcing effects of propofol and methohexital in baboons. *Drug Alcohol Depend* 1999;57:51–60. [PubMed: 10617313]
- Weerts, EM.; Goodwin, AK.; Kaminski, BJ.; Griffiths, RR.; Ator, NA. Self-administration of the gamma-hydroxybutyrate (GHB) prodrug gamma-butyrolactone (GBL) in baboons. Presented at the College on Problems of Drug Dependence 70th Annual Meeting; 2008 [(abstract accessed Aug 8, 2008)]. <<http://www.cpdd.vcu.edu/>>
- Winger G, Galuska CM, Hursh SR, Woods JH. Relative reinforcing effects of cocaine, remifentanyl, and their combination in rhesus monkeys. *J Pharmacol Exp Ther* 2006;318:223–229. [PubMed: 16571623]
- Winter JC. The stimulus properties of gamma-hydroxybutyrate. *Psychopharmacology* 1981;73:372–375. [PubMed: 6789360]
- Woody GE, Senay EC, Geller A, Adams EH, Inciardi JA, Schnoll S, Muñoz A, Cicero TJ. An independent assessment of MEDWatch reporting for abuse/dependence and withdrawal from Ultram (tramadol hydrochloride). *Drug Alcohol Depend* 2003;72:163–168. [PubMed: 14636971]
- Woolverton WL, Rowlett JK, Winger G, Woods JH, Gerak LR, France CP. Evaluation of the reinforcing and discriminative stimulus effects of gamma-hydroxybutyrate in rhesus monkeys. *Drug Alcohol Depend* 1999;54:137–143. [PubMed: 10217553]
- Wright C 4th, Kramer ED, Zalman MA, Smith MY, Haddox JD. Risk identification, risk assessment, and risk management of abusable drug formulations. *Drug Alcohol Depend* 2006;83:S68–76. [PubMed: 16581201]
- Wu LT, Pilowsky DJ, Patkar AA. Non-prescribed use of pain relievers among adolescents in the United States. *Drug Alcohol Depend* 2008;94:1–11. [PubMed: 18054444]
- Young AM, Thompson T. Effects of naloxone on schedule-controlled behavior in morphine-maintained pigeons. *J Pharmacol Exp Ther* 1978;205:236–245. [PubMed: 564955]
- Zhang L, Walker EA, Sutherland J 2nd, Young AM. Discriminative stimulus effects of two doses of fentanyl in rats: pharmacological selectivity and effect of training dose on agonist and antagonist effects of mu opioids. *Psychopharmacology* 2000;148:136–145. [PubMed: 10663428]