Title

Core outcome measures for opioid abuse liability laboratory assessment studies in humans: IMMPACT recommendations

Journal

Pain, 153(12)

ISSN

0304-3959

Authors

Comer, Sandra D
Zacny, James P
Dworkin, Robert H
et al.

Publication Date

2012-12-01

DOI

10.1016/j.pain.2012.07.035

Peer reviewed

Published in final edited form as:
Pain. 2012 December ; 153(12): 2315–2324. doi:10.1016/j.pain.2012.07.035.

Core outcome measures for opioid abuse liability laboratory assessment studies in humans: IMMPACT recommendations

Sandra D. Comer[a,*], **James P. Zacny**[b], **Robert H. Dworkin**[c], **Dennis C. Turk**[d], **George E. Bigelow**[e], **Richard W. Foltin**[a], **Donald R. Jasinski**[f], **Edward M. Sellers**[g], **Edgar H. Adams**[h], **Robert Balster**[i], **Laurie B. Burke**[j], **Igor Cerny**[j], **Robert D. Colucci**[k], **Edward Cone**[e], **Penney Cowan**[l], **John T. Farrar**[m], **J. David Haddox**[n], **Jennifer A. Haythornthwaite**[e], **Sharon Hertz**[j], **Gary W. Jay**[o], **Chris-Ellyn Johanson**[p], **Roderick Junor**[q], **Nathaniel P. Katz**[r,s], **Michael Klein**[j], **Ernest A. Kopecky**[t], **Deborah B. Leiderman**[u], **Michael P. McDermott**[v], **Charles O'Brien**[w], **Alec B. O'Connor**[x], **Pamela P. Palmer**[y], **Srinivasa N. Raja**[z], **Bob A. Rappaport**[j], **Christine Rauschkolb**[aa], **Michael C. Rowbotham**[ab], **Cristina Sampaio**[ac], **Beatrice Setnik**[ad], **Marta Sokolowska**[ae], **Joseph W. Stauffer**[af], and **Sharon L. Walsh**[ag]

[a]Division on Substance Abuse, New York State Psychiatric Institute/College of Physicians and Surgeons, Columbia University, New York, NY, USA

[b]Department of Anesthesia & Critical Care, University of Chicago, Chicago, IL, USA

[c]Department of Anesthesiology, University of Rochester, Rochester, NY, USA

[d]Department of Anesthesiology, University of Washington, Seattle, WA, USA

[e]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[f]Department of Medicine, Johns Hopkins University, Baltimore, MD, USA

[g]Kendle Early Phase Toronto, formerly Decision Line Clinical Research Corporation, Toronto, Ontario, Canada

[h]Covance, Princeton, NJ, USA

[i]Institute for Drug and Alcohol Studies, Virginia Commonwealth University, Richmond, VA, USA

[j]United States Food and Drug Administration, Silver Spring, MD, USA

[k]Colucci & Associates LLC, Newtown, CT, USA

[l]American Chronic Pain Association, Rocklin, CA, USA

[m]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, USA

[n]Purdue Pharma L.P., Stamford, CT, USA

[o]Pfizer, Inc., New London, CT, USA

[p]Psychiatry and Behavioral Sciences, Wayne State University, Detroit, MI, USA

[q]Eisai Limited, Woodcliff Lake, NJ, USA

[*]Address correspondence to: Sandra D. Comer, Ph.D., College of Physicians and Surgeons of Columbia University and the New York State Psychiatric Institute, 1051 Riverside Drive, Unit 120, New York, NY 10032, Ph: 212-543-5981, sdc10@columbia.edu.

[r]Analgesic Research, Needham, MA, USA

[s]Tufts University, Boston, MA, USA

[t]Endo Pharmaceuticals Inc., Chadds Ford, PA, USA

[u]CNS Drug Consulting LLC, McLean, VA, USA

[v]Departments of Biostatistics and Neurology, University of Rochester, Rochester, NY, USA

[w]Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

[x]Department of Medicine, University of Rochester, Rochester, NY, USA

[y]AcelRx Pharmaceuticals Inc., Redwood City, CA, USA

[z]Department of Anesthesiology and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

[aa]Johnson & Johnson Pharmaceutical Research & Development LLC, Raritan, NJ, USA

[ab]Research Institute, California Pacific Medical Center, San Francisco, CA, USA

[ac]Clinical Pharmacology and Therapeutics Laboratory, Faculdade de Medicina de Lisboa, Lisbon, Portugal

[ad]King Pharmaceuticals, Inc., Cary, North Carolina, USA

[ae]Grünenthal USA Inc., Bedminster, NJ, USA

[af]Durect Corporation, Cupertino, CA, USA

[ag]Department of Behavioral Science, Center on Drug and Alcohol Research, University of Kentucky, KY, USA

Abstract

A critical component in development of opioid analgesics is assessment of their abuse liability (AL). Standardization of approaches and measures used in assessing AL has the potential to facilitate comparisons across studies, research laboratories, and drugs. The goal of this report is to provide consensus recommendations regarding core outcome measures for assessing abuse potential of opioid medications in humans in a controlled laboratory setting. Although many of the recommended measures are appropriate for assessing the AL of medications from other drug classes, the focus here is on opioid medications because they present unique risks from both physiological (e.g., respiratory depression, physical dependence) and public health (e.g., individuals in pain) perspectives. A brief historical perspective on AL testing is provided and then those measures that can be considered primary and secondary outcomes and possible additional outcomes in AL assessment are discussed. These outcome measures include: (1) subjective effects (some of which comprise the primary outcome measures, including drug liking); (2) physiological responses; (3) drug self-administration behavior; and (4) cognitive and psychomotor performance. Prior to presenting recommendations for standardized approaches and measures to be used in AL assessments, the appropriateness of using these measures in clinical trials with patients in pain is discussed.

Keywords

opioids; opioid analgesics; abuse; abuse liability; abuse potential

1. Rationale for selection of core outcome measures

A critical component in the development of opioid analgesics is the assessment of their abuse liability (AL), 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" ([3]). A number of reviews have elucidated the methods and measures used in human abuse liability assessment (ALA), and some of those methods and measures will be elaborated upon in this article (cf. [1]; [3]; [4]; [7]; [9]; [16]; [22]; [25]; [35]; [36]; [37]; [46]; [53]). In addition to ALA, critical factors that inform evaluation of AL include, but are not limited to, the chemical composition of the drug, its formulation, data from preclinical behavioral pharmacology studies, data from clinical trials pertinent to abuse potential, safety and efficacy data, and epidemiological data on abuse when available ([17]).

Currently, a number of laboratories assess the AL of opioids. In some instances, the same instruments are used, but more often than not the ALA batteries are not identical. In some respects, this is good: if a drug with unknown AL is tested in several laboratories that use different ALA batteries, and the same conclusion is reached regarding the degree to which the drug has AL, this provides robust and convincing evidence that the drug indeed has that degree of AL. Inappropriate conclusions can be made, however, if new instruments are used that putatively assess AL without proper validation.

For over eight decades, the College on Problems of Drug Dependence (CPDD) has served a leadership role in the field of drug abuse, and its members have facilitated the development and refinement of methods for preclinical and clinical ALA of psychoactive drugs. In 2003 and 2006, CPDD convened conferences to address various issues pertaining to ALA. One of the recommendations made by panels of experts at both conferences was a call to standardize some psychometric scales (e.g., "drug liking") for human ALA in order to facilitate comparisons of AL testing across studies, research laboratories, and drugs.[1]

2. Methods

The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT; e.g., [13]; [14]; [62]; [63]) convened a consensus meeting to provide recommendations regarding key outcome measures for human opioid AL studies. The meeting included an international group of participants from universities, government agencies, industry, and a patient advocacy organization selected on the basis of research, clinical, or administrative expertise relevant to evaluating the efficacy and safety of analgesic medications, in particular, their AL. Background presentations that were delivered to facilitate discussion are available on the IMMPACT website: www.immpact.org (IMMPACT, 2009).

The objective of this article is to present the recommendations for standardizing, when possible, key outcome measures in opioid AL studies. The charge for the consensus meeting was to discuss the different measures that are used in opioid ALA and to come to an agreement on selecting key primary and secondary outcome measures that might comprise a standardized opioid ALA battery. This article describes the consensus reached on recommendations for core outcome measures. In response to the upsurge in prescription

[1]Expert panel recommendations in two special issues of *Drug and Alcohol Dependence*. The first issue (Volume 70, Issue 3, Supplement 1) was published in 2003 (Assessing the Abuse Liability of CNS Drugs) and the expert panel recommendations (Chair: Edward Sellers) pertaining to standardization of measures is on p.113. The second issue (Volume 83, Supplement 1) was published in 2006 (Impact of Drug Formulation on Abuse Liability, Safety and Regulatory Decisions); the expert panel recommendations (Chair: Charles Grudzinkas) again calling for measurement standardization is on p. S80.

opioid abuse, pharmaceutical companies and others have developed or will be developing formulations intended to minimize abuse of prescription opioids ([39]; [66]). This article is timely in that subsequent to the consensus meeting, the US Food and Drug Administration issued a draft guidance on the 'Assessment of abuse potential of drugs" ([17]).

3. Historical perspective

Human testing for AL of opioids was initiated in the 1930s to facilitate the government-sponsored efforts to develop "non-addicting" substitutes for morphine-like drugs in order to reduce the public health and social problems associated with the diversion and misuse of opioids ([19]; [31]; [32]). Initial studies of abuse potential at the Addiction Research Center in Lexington, Kentucky were based upon studies of the morphine withdrawal syndrome because it was thought at the time that the addiction potential of opioids was directly related to their ability to produce physical dependence. If a drug prevented the onset of physical withdrawal symptoms or signs in dependent subjects (substitution studies) or alleviated physical withdrawal symptoms (suppression studies), the drug was judged to have addiction potential. Subjective changes in mood and perception induced by opioids were recognized but did not contribute to the addiction potential assessment.

In the late 1940s and early 1950s, Isbell and colleagues demonstrated that opioid abusers could discriminate morphine from placebo as well as from other substances of abuse, and that the opioid abusers reported dose-related liking of the effects ([18]; [30]). In the early 1960s, the Addiction Research Center Inventory (ARCI) ([28]; [29]) and the Single Dose Questionnaire (SDQ) ([21]) were developed. These standardized instruments enabled systematic characterization of drug-induced changes in mood, feeling, thinking, and perception, as well as the degree to which drugs were "liked." These assessment measures, which will be described briefly in Section 5, allowed the assessment of changes in subjective experience as a function of dose and type or class of drug administered. Thus, subjective effects originally considered unimportant in AL testing of opioids became integral to the assessment.

4. Critical elements of an opioid AL study

The major focus of this article is on primary and secondary core outcome measures for opioid AL studies. However, other components or elements of an opioid AL study are important and they will be outlined briefly here because they are related to the outcome measures discussed subsequently ([22]; [46]). The design of the human abuse potential study should be based on the study objectives and the specific statistical model that will be used for data analysis. A prototypic opioid AL study is conducted in a controlled clinical pharmacology laboratory and a randomized, placebo-controlled, crossover design is typically used. The study includes a positive control opioid drug (i.e., one that has known AL, such as morphine) to which the test drug will be compared. Inclusion of a broad range of dose levels of the test compound and the positive control drug is typical. However, if the test opioid has known abuse potential but is a different formulation, fewer dose levels may be required than when the test opioid is novel and has not been previously examined. Supratherapeutic doses of both the test opioid and the positive control opioid should be included in the study design because drug abusers often use doses much higher than those used in clinical settings.

Subject population is a critical consideration. Experienced drug abusers (e.g., those having extensive and recent histories of non-medical use of drugs) are widely accepted as the appropriate and most sensitive clinical population for assessing AL. They have demonstrated by their behavior that they recognize and appreciate the pharmacological effects of certain drugs that lead them to be abused. Within the population of experienced

opioid abusers, a further distinction should be made. Specifically, a population of opioid abusers should be chosen that "matches" the test conditions. That is, if medications are being tested via the oral route, then the abusers who are enrolled in the study should prefer, or have some experience, using opioids via the oral route. The primary reason for this recommendation is that individuals who are accustomed to using drugs by routes that produce more rapid peak drug effects (e.g., intravenous or smoked) often find oral drug effects to be less appealing and therefore rate their subjective effects as being smaller in magnitude than individuals who are more accustomed to taking drugs by the oral route. Furthermore, the experienced drug abusers who are tested should not be physically dependent on opioids in order to avoid complications in interpretation of the results (e.g., if the test medication produces no increases in subjective effects, is it because of cross-tolerance to the maintenance medication or because it truly has minimal subjective effects?). Finally, ethical considerations also determine what type of opioid abusers should be recruited into an ALA study. First, they should not be seeking treatment for opioid abuse or currently in treatment. Second, if the test drugs are to be administered via the intravenous route, for example, the participant population should be one that has experience with and prefers using drugs intravenously. Introducing a more toxic route of drug administration or one associated with higher abuse liability to an individual who has never used drugs by that route is ethically problematic.

In contrast to experienced opioid abusers, individuals who do not abuse drugs are generally less able to reliably assess the AL aspects of drugs, and are not used in typical ALA studies of opioids. This does not necessarily preclude examination of opioids in those who do not abuse drugs, and some AL studies have been conducted in this population (e.g., [11]; [48]; [70]; [71]). However, individuals who do and do not abuse drugs may differ in important ways (e.g., [11]; [17]; [64]; [70]), and further research is warranted on the impact of these differences on ALA. Finally, a number of prescription opioid AL studies include prescreening procedures, initially determining that participants enrolled in the study can distinguish an abused opioid from placebo, and that they report positive subjective effects from the opioid ([6]; [17]; [38]; [57]; [61]). Failure to identify participants who can distinguish an opioid from placebo and report positive subjective effects prior to administering the test drug could result in a study that underestimates the abuse liability of the test drug. In short, while we recognize that individual variability in response to drugs exists (e.g., due to differences in drug use history, the presence or absence of pain, genetic polymorphisms, etc.), which may warrant the use of different endpoints under certain circumstances, our recommendations emphasize the use of standardized measures in experienced drug abusers.

5. Subjective effects

Drugs are abused because of the effects they produce, and many of these effects are uniquely and solely accessible to and assessable by the recipient of the drug. That is, many important effects of drugs may occur in the subjective realm and may not be detectable or quantifiable by objective observation and measurement. For these reasons, subjective reports from individuals to whom test medications have been administered constitute a core element of drug ALA.

5.1. Domains of subjective reports

The subjective effects assessed in drug ALA studies are typically within the following domains: qualitative and quantitative effects experienced (e.g., feel drug effect), drug liking, likelihood of taking the drug again, drug identification (i.e., similarity to known drugs of abuse), and direct or comparative value estimates (e.g., dollar value, preference, or choice). The domains are not mutually exclusive as specific assessments may contribute to multiple

domains. These domains can be viewed as steps along a continuum of behavioral specificity of the subjective reports – from having participants describe drug effects from which the investigator will infer the likelihood of drug-taking behavior to having participants themselves directly report their likely behavior.

5.1.1. Effects experienced—The subjective effects experienced are assessed by self-reported descriptive characterizations of symptoms, mood, and physiological, sensory, and behavioral experiences produced by a drug. Typically, these descriptive reports are interpreted by judging the extent to which a test drug produces a profile of effects similar to that of a known reference comparator drug. Drugs with similar profiles of effects generally have similar abuse liabilities. Instruments that have been used to assess subjective effects produced by opioid administration include the ARCI, the SDQ, adjective rating scales, and visual analog scales (VAS). These instruments have been described elsewhere (e.g., [4]; [21]; [22]; [31]; [44]; [52]) and so will be discussed only briefly here (Table 1). Finally, the subjective effects measured in opioid ALA studies should not just assess for positive effects but should also assess for negative or “bad” effects, because an opioid that produces only positive effects is likely to have greater AL than an opioid that produces both positive and negative effects ([12]; [15]; [22]; [46]).

The original ARCI consisted of 550 true-false statements ([23]; [24]; [28]). In 1971, Martin and colleagues described a shorter version (49 true-false statements and five scales) that is now commonly used in opioid AL testing ([44]). Three of the scales especially pertinent to opioid ALA are the Morphine-Benzedrine Group (MBG) scale, a measure of euphoria; Pentobarbital-Chlorpromazine-Alcohol Group (PCAG) scale, a measure of sedation; and the Lysergic Acid Diethylamide (LSD) scale, a measure of dysphoria, as well as somatic and sensory disturbances ([4]). In opioid abusers, full mu agonists tend to increase scores (but not always) on the MBG scale without increasing scores on the PCAG scale and the LSD scale ([4]; [31]). Mixed action opioids, such as nalbuphine and pentazocine, have also been shown to increase MBG scores, but at high doses they have also been shown to increase scores on the PCAG and LSD scales, which differs from the profile of effects produced by full mu opioids (e.g., [33]; [34]; [50]).

The SDQ was designed to assess whether or not a participant felt the effects of a drug, the participant’s identification of the substance as belonging to a particular pharmacological class, what symptoms were being experienced, and the degree of liking of the drug ([21]). Like the ARCI, different profiles of effects have been found between full mu agonists and mixed-action agonists on the SDQ (e.g., [33]; [34]). Although the SDQ in its original form is rarely used nowadays, each component of it is typically included in ALA of opioids via other formats, including adjective rating scales or checklists, and VASs.

5.1.2. Liking—The measurement of drug liking is considered to be one of the most sensitive and reliable assessments of the likelihood of abuse of a drug ([4]; [7]; [22]) and is a key component of opioid ALA. It also has excellent face and predictive validity to other key measures of abuse liability, including drug self-administration in both the laboratory and in real-world settings (e.g., [9]; [22]). The SDQ originally measured this domain by use of an ordinal scale but a VAS is now typically used to assess drug liking. The VAS can be unipolar in nature (0=not at all; 100=extremely) or bipolar (e.g., 0=dislike a lot; 50=neutral; 100=like a lot). There has been some discussion that the bipolar VAS may have advantages over a unipolar VAS in that the bipolar VAS presumably reflects the net balance of liked and disliked effects ([22]). Although some research groups have been evaluating the merits of a unipolar versus bipolar scale empirically, there is as yet no clear consensus within the ALA community about which type of scale should be preferred; some groups prefer unipolar scales because they may better reveal when mixed effects occur. If both unipolar and bipolar

scales are used as different outcomes in a single study, participants should be carefully instructed on their differences to avoid confusion.

5.1.3. Likelihood to take again—In this assessment, participants rate their desire, willingness, or disposition to take a drug again. This measure is often highly correlated with drug liking ([22]), but is useful as a key outcome measure because it may provide an estimate of future behavior rather than simply drug liking in the moment. For some individuals, opioids produce both positive and negative effects (such as nausea) that manifest at different points in time after drug administration. Therefore, while participants may report liking the drug during the experimental session, their desire to take the drug again either during or after the session may not differ significantly from placebo, depending on when the aversive effects emerge ([11]; [71]).

5.1.4. Drug identification—This is a domain in which the participant assesses the similarity of a test drug to known drugs with which the individual has experience. This is commonly done with either categorical drug-class identification questions such as “Is it an opioid?” or with graded similarity rating questions such as “Is it similar to [Drug X] (e.g., morphine, heroin, cocaine, alcohol)?” The SDQ assesses this domain, and since its origin other investigators have added other drugs relevant to the drug classes listed on the SDQ (e.g., [49]; [64]).

5.1.5. Drug high—Most experienced opioid abusers report feeling “high” after opioid administration. While we believe that the term “high” is often a valuable measure to aid in characterizing the AL of drugs ([46]), including opioids, one problem is that the meaning of the term can be different for different individuals. For some abusers, “high” may be equivalent to “good effect,” but for others, “high” may be any feeling other than “normal.” For example, studies have shown that opioid abusers can report feeling high from an array of mixed action opioids while simultaneously reporting significant dysphoria (e.g., “feel bad”) and no evidence of liking or good effects from the drug ([49]; [65]). However, we should emphasize that this latter finding is more the exception than the rule. Recent clinical pharmacology laboratory studies examining putative abuse deterrent formulations have considered the subjective effect of “high” as either a primary outcome measure designed to assess AL ([67]) or as a term denoting positive effects ([55]). While assessment of feeling “high” is useful and we recommend it as a key outcome measure, it should not be used as a sole outcome measure because of the variability between participants in how the term is defined and its sensitivity to CNS effects that are unrelated to euphoria.

5.1.6. Street value—This domain assesses an estimate of the street value of a drug and/or how much money a person would be willing to pay for the drug. There is typically a high degree of correlation between how much individuals are willing to pay for a dose of drug and the degree to which they report liking the effects of that drug ([22]). However, until recently, most opioid AL studies involved subjects who were currently abusing heroin, which could be considered a prototypic “street” drug that would be purchased with money. In prescription opioid AL studies, the subject pool typically includes people who are abusing prescription opioids. In some geographical regions, the predominant means of obtaining these opioids is less likely to be by purchasing them on the street. In the 2010 National Survey on Drug Use and Health, for example ([54]), respondents who admitted to nonmedical use of prescription opioids in the previous 12 months were queried on their primary source of the drugs: the predominant source was from a friend or relative for free (55.0%). Only 4.4% of respondents reported they obtained prescription opioids from a drug dealer or other stranger. If a subject in a prescription opioid AL study has limited experience purchasing opioids on the street, their self-report on how much they would be willing to pay

for a prescription opioid would be of questionable validity. For this reason we are not recommending estimated street value as a core outcome measure, but rather a potentially useful secondary outcome measure (e.g., if most participants in a prescription AL study obtain their drugs from a street dealer).

5.2. Prioritizing AL outcome measures

The FDA's Drug Abuse Advisory Committee initiated development of the draft guidelines on ALA in the 1990s ([3]), and the FDA published a draft guidance in January 2010 that included prioritization of various data sources that might contribute to ALA ([17]). These ranged from chemical structure to epidemiological experience, with the latter designated as having the highest priority as it is the actual outcome of public health interest. Clinical trial results, such as human clinical pharmacology laboratory assessments, as described in this section, were designated as having the next highest priority and as providing one of the most useful and valid approaches to ALA in the absence of epidemiological data (which is typically the circumstance with non-approved medications).

5.3. Limitations of descriptive effect and similarity assessments

Clinical research methods that focus on assessing subjective reports of effects experienced and of similarity to comparison opioids of known AL have a long and proven history of being sensitive, informative, and useful in assessing AL (e.g., [22]; [32]). These tests focus on capturing the qualitative aspects of an opioid's subjective effects and include the ARCI, some items on the SDQ, VASs, and adjective rating checklists. However, because these methods are based on assessing similarity to known opioids of abuse, they may be unresponsive or less sensitive for detecting AL of drugs with novel pharmacologies, such as those with different mechanisms of action or with different profiles of effects than the traditional drugs of abuse commonly used as reference standards in opioid ALA studies. For assessing such novel drugs, indices reflecting liking or likelihood to take again may be more sensitive and have broader applicability because they do not depend to the same degree on similarity to a known reference standard.

5.4. Strengths, limitations, and applicability of subjective effects measures

Human laboratory subjective effects assessment methodologies are excellent for characterizing direct drug action and assessing the relative similarities and differences among pharmacologically similar and different compounds. The methods are effective in assessing and characterizing test compounds relative to the known reference compound(s) included in a study. It is important to recognize and emphasize this "relative to" function -- these methods are relativistic tools, not absolute measurements. They provide a gradation and rank ordering of AL. Historically, these methods appear to be quite valid, with good correspondence between the human laboratory methods and assessments provided by other methods such as animal laboratory behavioral studies of drug self-administration and epidemiological estimates of abuse.

The public health phenomenon of drug abuse is strongly influenced by the pharmacology of drugs and also by the social context in which they are taken, including availability, laws, and cultural norms; these latter features cannot be readily incorporated into the laboratory setting. Thus, a drug of substantial AL from a pharmacological perspective might have limited abuse if it is available only in highly controlled inpatient hospital settings, whereas a drug exhibiting only modest apparent AL from a pharmacologic perspective might have substantial adverse public health impact if it is widely available in outpatient medical settings.

5.5. Is there a single best assessment?

The consensus among drug ALA experts is that the subjective rating of liking is probably the most informative single AL outcome measure ([7]; [17]; [22]; [46]). In AL studies that specify *a priori* a single primary outcome measure, subjective liking is the recommended choice. However, there is variability among investigators regarding exactly how liking is best assessed – Likert-rating scales, VASs, unipolar liking versus bipolar disliking-liking scales, peak liking, area-under-the-curve liking, next-day or end-of-session liking. One would expect all of these to be highly inter-correlated. At present, we do not have strong data for arguing that any one is better than the others. Most common as a primary outcome measure in current practice is peak (maximum) VAS liking, assessed using either a unipolar liking scale or a bipolar disliking-liking scale over an appropriate time course based upon the drug's duration of action.

In practice, no study would be conducted with only a single measure, and measures are typically included from most or all of the subjective report domains described above. The ultimate judgment about AL will be guided by both the pre-specified primary or co-primary outcome measures and two or three possible secondary outcome measures along with the overall profile of outcomes and their coherence. Usually coherence and high intercorrelation of indices will be the case, and subjective report assessments can then contribute to a persuasive characterization of relative AL.

5.6. Recommendations for subjective ALA measures

Many subjective report indices have been utilized to assess the AL of new opioid medications relative to existing medications. Although a convergence of multiple indices is most persuasive in determining the likelihood that a medication will be abused, the use of a few key standardized measures in ALA will allow comparisons to be made across different studies in different laboratories. For this purpose, measures of liking, likelihood to take again, and drug identification appear most useful. Another measure that could be considered as a key outcome measure in opioid AL studies is drug "high," especially when used with other measures such as drug "liking" and desire to "take the drug again" to ensure that elevated ratings are reflective of positive effects. These effects should be assessed prior to drug administration, as well as at specific time points afterward. At the very least, post-drug administration assessments should be made before the peak drug effect is observed (the ascending portion of the time-effect curve), at the time of peak effect, during the descending portion of the time-effect curve, and when the effects have completely dissipated. Ideally, additional post-drug assessments would be made in order to minimize the chances of missing the ascending, peak, and descending portions of the time-effect curve. For example, in some studies, frequent assessments are made during the first 1–2 hours after drug administration (e.g., every 15 min), after which less frequent assessments occur (e.g., every 30 or 60 min). In this manner, the complete time course of drug effects would be captured.

6. Physiological responses

Physiological responses are assessed in opioid AL studies for at least two reasons: to ensure the safety of the subject and to provide a more comprehensive profile of the effects of an opioid. The monitoring of such responses is an integral component of opioid ALA, and such physiological responses can be considered key secondary outcome measures.

6.1. Medical supervision

Because opioid AL studies typically are conducted in healthy recreational drug using volunteers who can expect no therapeutic benefit, often with supratherapeutic doses, special attention should be given to medical safety issues. The study setting should be a medical one

with continuous supervision by appropriately trained and licensed medical personnel (e.g., ACLS certified) and appropriate resuscitative equipment.

6.2. Drug safety

Typical physiological safety measures in opioid AL studies should include cardiovascular monitoring (blood pressure, heart rate, electrocardiography), respiratory monitoring (pulse oximetry for oxygen saturation, respiratory rate, and/or expired carbon monoxide), urine drug testing, and in some cases, clinical laboratory tests (e.g., hepatic and renal function electrolytes, lipids related to dosing). Urine drug testing (UDT) is typically performed at the beginning of ALA as a safety precaution because of the possibility of drug interactions and can also be used to confirm that a person is actually using opioids. Measurement of plasma drug and metabolites can play a supplementary role in the evaluation of safety during an AL study. However, these measurements and subsequent pharmacokinetic and pharmacodynamic analyses are not routinely done.

6.3. What physiological endpoints should be measured?

Physiological effects that are typically produced by opioids include cardiorespiratory depression (as assessed by respiration rate and oxygen saturation) and miosis. Measuring cardiorespiratory depression is particularly important for safety assessment. Miosis (pupillary constriction) is an objective biological index of opioid action that has been employed as a standard measure in ALA for over five decades ([4]; [31]). Miosis is usually highly correlated with plasma opioid concentration, ability to suppress abstinence, incidence of side effects, and intensity of euphoria in abusers (e.g., [20]; [31]; [40]).

6.5. Recommendations for physiological and safety measures

Physiological and safety measures are an integral aspect of human AL studies. Specific physiological measures recommended for opioid ALA in humans include pupillary constriction, respiration rate, and oxygen saturation. Such measures can be associated with an opioid's positive subjective effects, similarity to known drugs of abuse, and public health risk. For safety, monitoring of respiration rate and oxygen saturation should be initiated prior to drug administration and should continue throughout the laboratory session. Recording of respiration rate and oxygen saturation data should be made at intervals that are guided by the expected time course of drug effects based on the best available data (e.g., recordings would be made for longer periods of time for medications that are contained within sustained-release formulations). Pupil diameter should be measured prior to drug administration, during the ascending portion of the time-effect curve, at the time of peak effect, during the descending portion of the time-effect curve, and when the effects have completely dissipated.

7. Drug self-administration behavior

During assessments of the AL of a drug, a primary interest is measuring the likelihood that it will be taken for non-medical purposes. As described above, measurements of subjective effects generally provide good estimates of the likelihood that a drug will be abused. However, in some instances, such as when the profile of subjective effects of a novel drug is ambiguous or when the AL of certain putative abuse-deterrent formulations of known drugs of abuse need to be ascertained, measuring drug-taking behavior itself may be useful.

7.1. Factors affecting the reinforcing effects of drugs

A drug is said to serve as a "reinforcer" if behaviors leading to its consumption increase in probability after the drug is experienced. In studies of drug self-administration, these behaviors have included verbal requests for the drug, bicycle pedaling, pulls or presses on a

lever, and finger presses on a computer mouse (e.g., [10]; [41]; [59]). The exact behavior is not critical; it just needs to be one that can be reliably detected and quantified. Factors affecting the reinforcing effects of drugs include duration of drug action, route of drug administration, participant characteristics, and the specific type of self-administration procedure utilized. The issue of participant characteristics (e.g., intravenous versus oral opioid abuser) has already been discussed in Section 3 and will not be discussed in this section.

The duration of action of the medication is an important variable because potential carry over effects from one session to another may alter the choice to self-administer the drug. That is, if participants are still experiencing effects of the first dose, responding for subsequent doses may be reduced. When trying to determine the appropriate interval for self-administration, the investigator should rely on available pharmacodynamic and pharmacokinetic data collected in humans to guide selection of an appropriate intersession interval.

Another important consideration is the expected route by which the drug will most likely be abused. Opioids can be abused by multiple routes, including but not limited to the oral, intranasal, injected (intravenous, intramuscular, subcutaneous), and smoked routes. Because delivery of a drug by a route that produces a faster onset of action and/or higher plasma concentrations is generally believed to produce greater abuse potential, the route by which participants are allowed to self-administer the drugs is a critical variable to consider in designing a self-administration study.

A final important consideration is the type of self-administration procedure that should be used. Many different methods for assessing the reinforcing effects of drugs have been used in humans. Broadly speaking, these procedures can be divided into verbal and non-verbal operant procedures. We have recommended a few procedures, described in more detail in Table 2. The choice of procedure depends in part on the characteristics of the test drug and the specific research question.

7.2. Strengths and limitations of self-administration procedures

Drug self-administration procedures have the advantage of providing direct information about the AL of a novel compound or formulation compared to a drug with known AL (e.g., drug versus drug choice). In addition, these procedures model the behavior (namely, drug taking) that is of interest to those who are concerned with a medication's potential for abuse. However, self-administration studies typically are not used in first-line assessments of the AL of a medication for two reasons: (1) they are more difficult to design and conduct than the standard AL study, and (2) self-administration procedures have not been validated to the same extent as the subjective effects measures described above.

7.3. Recommendations for self-administration procedures

In some instances, measuring drug-taking behavior itself may be useful in order to obtain a more complete profile of the AL of a drug, such as when a drug with known abuse liability is re-formulated into a tablet that is ostensibly "abuse-deterrent" or "tamper-resistant". Thus, IMMPACT recommends that drug self-administration should be considered as a potential outcome measure in opioid ALA. We have provided recommendations for different self-administration procedures (Table 2) because different research objectives may determine which of these procedures is most appropriate. Regardless of which procedure is used, investigators should pay particular attention to variables that may influence drug-taking behavior, such as duration of drug action and route of drug administration as well as the characteristics of the study population. Ideally, each dose of the test drug should be assessed

in the self-administration paradigm in order to obtain a complete dose-effect curve of its reinforcing effects, because the slope of the dose-effect curve gives important information about the efficacy of the drug as a reinforcer (i.e., shallower dose-response curves typically are obtained with drugs that have lower AL). The time points at which reinforcing effects are assessed should be driven by the drug's pharmacodynamic effects. That is, participants should be given the opportunity to self-administer drug when the effects of the previous dose have completely dissipated.

8. Cognitive and psychomotor performance

When assessing AL, it may be useful to incorporate cognitive and psychomotor testing into the study as a secondary measure in all ALA studies in order to: (1) quantify dose-response functions directly on multiple measures of drug effect, thus resulting in a more complete characterization of the psychoactive effects of the drug, and (2) provide information on the likelihood that abuse of the drug will produce harmful effects ([4]; [22]; [46]; [53]). In non-drug-abusing healthy volunteers, a number of studies have shown that prescription opioids at supratherapeutic doses impair performance on various cognitive and psychomotor tests, including the Digit Symbol Substitution Test (DSST) ([68]), the Baddeley Logical Reasoning Test ([2]), and the Maddox Wing Test ([26]) (e.g., [11]; [48]; [69]; [70]; [71]). In contrast, a number of studies primarily using a computerized version of the DSST ([47]) showed that performance was not impaired after parenteral administration of various full mu agonists, partial mu agonists, and mixed action agonists in non-dependent heroin users (e.g., [27]; [49]; [50]; [51]). Whether such lack of impairment from opioids would generalize to prescription opioid abusers should not be assumed.

Several recent ALA studies have incorporated cognitive and psychomotor testing as secondary measures in their study design and have detected impairment on some tests ([11]; [56]; [60]; [64]). Those tests included a divided attention task, the computerized version of the DSST, the Maddox Wing Test, and the critical flicker-fusion test ([58]). We therefore suggest that such testing would be informative as secondary outcome measures in ALA testing of opioids and novel analgesic compounds. However, it would be premature to recommend preferential use of one of these tests over others.

8.1. Recommendations for the assessment of cognitive and psychomotor performance

We encourage investigators who conduct AL testing to include one or more cognitive and psychomotor measures in their studies, as ultimately this may inform the research community regarding which tests are the most sensitive in detecting impairment. We recommend that measures be used that have demonstrated good test-retest reliability, are not prone to practice effects, and assess different aspects of performance (cognitive as well as psychomotor processes). Ideally, performance should be measured prior to drug administration, during the ascending portion of the time-effect curve, at the time of peak effect, during the descending portion of the time-effect curve, and when the effects have completely dissipated.

9. Extrapolation to people with chronic pain

A final topic for which there is limited evidence involves the appropriateness of using the measures described above in individuals with pain. If these subjective and behavioral measures are used to assess the AL of an opioid in patients with pain, the results will be difficult to interpret. For example, patients may report drug liking, willingness to take the drug again, and good drug effects not because of feelings of elation or elevated mood, but rather because the drug reduced their pain. One of the few studies to address this issue was conducted in the early 1950's by Lasagna and colleagues ([42]). Amphetamine,

pentobarbital, heroin, morphine, and placebo were administered to 30 chronically ill patients, who were then asked a series of questions about their drug experience. Eight of the 30 patients reported no differences among the drugs, five could not remember the effects well enough to compare, and 16 reported some drugs as more pleasant. Four patients reported that morphine was pleasant only because it relieved symptoms. Thus, it is difficult to determine whether the patients believed the drugs were pleasant because they reduced pain or because they produced euphoria. Future research should examine factors that may influence the subjective effects of opioids in patients with pain (e.g., history of drug abuse, presence of aberrant drug-related behaviors, and pain severity).

Most assessments of opioid AL are conducted in individuals without pain who abuse opioids, and the instruments were tailored for that population. Refinements of existing measures with questions such as “I like the dose because it reduces my pain” or “I like the dose because it reduces my anxiety” or “I like the dose because it makes me feel high” will need to be tested in patients with pain before the measures can be used for ALA in such individuals. Such questions get at the issue of motives for use, and have been employed successfully in other populations (e.g., adolescents who use prescription opioids for nonmedical purposes; [5]; [45]). Because abuse of opioid medications in patients with chronic pain is not well characterized and may be more prevalent than previously recognized (e.g., [8]; [43]), additional research is needed to gain a better understanding of factors underlying opioid abuse among patients with chronic pain.

10. General recommendations for conducting ALA studies

We recommend that four subjective effects measures should be included in ALA studies of opioid analgesics: drug liking, likelihood to take again, drug identification, and drug high (Table 3). These can be considered core primary outcome measures, that is, core in the sense that these measures should be part of any opioid ALA battery, and primary in the sense that they are important measures in determining the AL of a drug. This does not preclude the use of other subjective effects measures (e.g., good effects), and in many cases other tests should be used in order to more fully characterize AL-related effects as well as other effects. Indeed, these other effects might include those that would be considered negative in nature (e.g., bad effects) and could play a role in reducing the overall AL of a drug.

We consider the measures of respiration rate, oxygen saturation, and miosis as core secondary measures in opioid ALA studies (Table 3) for safety assessments and because they have demonstrated sensitivity to opioids with substantial AL (e.g., heroin and other full mu-opioid agonists). Although not integral to opioid ALA, cognitive and psychomotor measures are recommended for inclusion in opioid ALA in order to characterize more fully their psychopharmacologic effects and potential public health risks. It is premature, however, to recommend a specific test or tests at this time. In some circumstances, self-administration studies may be useful or necessary in ALA of opioids and novel analgesics. Whether to include cognitive, psychomotor performance, and/or self-administration measures and which test(s) to include should be determined by the study aims and the specific drug/formulation under evaluation.

Acknowledgments

The views expressed in this article are those of the authors, none of whom have financial conflicts of interest related to the issues discussed in this manuscript. No official endorsement by the US Department of Veterans Affairs, US Food and Drug Administration, US National Institutes of Health, or the pharmaceutical companies that provided unrestricted grants to the University of Rochester Office of Continuing Professional Education should be inferred.

The authors thank Paul J. Lambiase and Mary Gleichauf for their invaluable assistance in the organization of the IMMPACT meeting. Unrestricted grants were provided to the University of Rochester Office of Continuing

Professional Education to support the activities of IMMPACT, including the consensus meeting on which this article is based, by AcelRx, Durect, Eisai, Endo, Grünenthal, Johnson & Johnson, King, Pfizer, Purdue.

References

1. Ator NA, Griffiths RR. Principles of drug abuse liability assessment in laboratory animals. Drug Alcohol Depend. 2003; 70(3 Suppl):S55–72. [PubMed: 12759197]
2. Baddeley AD. A three-minute reasoning test based on grammatical transformation. Psychon Sci. 1968; 10:341–342.
3. Balster RL, Bigelow GE. Guidelines and methodological reviews concerning drug abuse liability assessment. Drug Alcohol Depend. 2003; 70(3 Suppl):S13–40. [PubMed: 12759195]
4. Bigelow GE. Human drug abuse liability assessment: opioids and analgesics. Br J Addict. 1991; 86(12):1615–1623. [PubMed: 1686196]
5. Boyd CJ, McCabe SE, Cranford JA, Young A. Adolescents' motivations to abuse prescription medications. Pediatrics. 2006; 118(6):2472–2480. [PubMed: 17142533]
6. 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(2):103–110. [PubMed: 10519736]
7. Carter LP, Griffiths RR. Principles of laboratory assessment of drug abuse liability and implications for clinical development. Drug Alcohol Depend. 2009; 105 (Suppl 1):S14–25. [PubMed: 19443137]
8. Chou R, Fanciullo GJ, Fine PG, Miaskowski C, Passik SD, Portenoy RK. Opioids for chronic noncancer pain: prediction and identification of aberrant drug-related behaviors: a review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. J Pain. 2009; 10(2):131–146. [PubMed: 19187890]
9. 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(1–2):1–15. [PubMed: 18436394]
10. Comer SD, Collins ED, Fischman MW. Choice between money and intranasal heroin in morphine-maintained humans. Behav Pharmacol. 1997; 8(8):677–690. [PubMed: 9832953]
11. Comer SD, Sullivan MA, Vosburg SK, Kowalczyk WJ, Houser J. Abuse liability of oxycodone as a function of pain and drug use history. Drug Alcohol Depend. 2010; 109(1–3):130–138. [PubMed: 20079977]
12. Comer SD, Sullivan MA, Whittington RA, Vosburg SK, Kowalczyk WJ. Abuse liability of prescription opioids compared to heroin in morphine-maintained heroin abusers. Neuropsychopharmacology. 2008; 33(5):1179–1191. [PubMed: 17581533]
13. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005; 113(1–2):9–19. [PubMed: 15621359]
14. Dworkin RH, Turk DC, Peirce-Sandner S, Baron R, Bellamy N, Burke LB, Chappell A, Chartier K, Cleeland CS, Costello A, Cowan P, Dimitrova R, Ellenberg S, Farrar JT, French JA, Gilron I, Hertz S, Jadad AR, Jay GW, Kalliomaki J, Katz NP, Kerns RD, Manning DC, McDermott MP, McGrath PJ, Narayana A, Porter L, Quessy S, Rappaport BA, Rauschkolb C, Reeve BB, Rhodes T, Sampaio C, Simpson DM, Stauffer JW, Stucki G, Tobias J, White RE, Witter J. Research design considerations for confirmatory chronic pain clinical trials: IMMPACT recommendations. Pain. 2010; 149(2):177–193. [PubMed: 20207481]
15. Evans SM, Funderburk FR, Griffiths RR. Zolpidem and triazolam in humans: behavioral and subjective effects and abuse liability. J Pharmacol Exp Ther. 1990; 255(3):1246–1255. [PubMed: 2262904]
16. Fischman MW, Johanson CE. Ethical and practical issues involved in behavioral pharmacology research that administers drugs of abuse to human volunteers. Behav Pharmacol. 1998; 9(7):479–498. [PubMed: 9862072]

17. Food and Drug Administration USDHHS, Center for Drug Evaluation and Research. Guidance for Industry: Assessment of Abuse Potential of Drugs. 2010. http://www.fda.gov/cder/guidance/index.htm
18. Fraser HF, Isbell H. Addiction liabilities of morphinan, 6-methyldihydromorphine and dihydrocodeinone. J Pharmacol Exp Ther. 1950; 100(2):128–135. [PubMed: 14784948]
19. Fraser, HF.; Jasinski, DR. The assessment of the abuse potentiality of sedative/hypnotics(depressants) (methods used in animals and man). In: Martin, WR., editor. Drug Addiction I Morphine, sedative-hypnotic and alcohol dependence Handbook of Experimental Pharmacology. Heidelberg: Springer-Verlag; 1977. p. 589-612.
20. Fraser HF, Nash TL, Vanhorn GD, Isbell H. Use of miotic effect in evaluating analgesic drugs in man. Arch Int Pharmacodyn Ther. 1954; 98(4):443–451. [PubMed: 13208336]
21. Fraser HF, Van Horn GD, Martin WR, Wolbach AB, Isbell H. Methods for evaluating addiction liability. (A) “Attitude” of opiate addicts toward opiate-like drugs. (B) a short-term “direct” addiction test. J Pharmacol Exp Ther. 1961; 133:371–387. [PubMed: 13701509]
22. Griffiths RR, Bigelow GE, Ator NA. Principles of initial experimental drug abuse liability assessment in humans. Drug Alcohol Depend. 2003; 70(3 Suppl):S41–54. [PubMed: 12759196]
23. Haertzen CA. Development of scales based on patterns of drug effects, using the addiction Research Center Inventory (ARCI). Psychol Rep. 1966; 18(1):163–194. [PubMed: 5908477]
24. Haertzen CA, Hill HE, Belleville RE. Development of the Addiction Research Center Inventory (ARCI): Selection of items that are sensitive to the effects of various drugs. Psychopharmacologia. 1963; 4:155–166. [PubMed: 14054658]
25. Haney M. Self-administration of cocaine, cannabis and heroin in the human laboratory: benefits and pitfalls. Addict Biol. 2009; 14(1):9–21. [PubMed: 18855806]
26. Hannington-Kiff JG. Measurement of recovery from outpatient general anaesthesia with a simple ocular test. Br Med J. 1970; 3(5715):132–135. [PubMed: 5431082]
27. Higgins ST, Preston KL, Cone EJ, Henningfield JE, Jaffe JH. Supersensitivity to naloxone following acute morphine pretreatment in humans: behavioral, hormonal and physiological effects. Drug Alcohol Depend. 1992; 30(1):13–26. [PubMed: 1591977]
28. Hill HE, Haertzen CA, Wolbach AB Jr, Miner EJ. The Addiction Research Center Inventory: Appendix. I. Items comprising empirical scales for seven drugs. II. Items which do not fifferentiate placebo from any drug condition. Psychopharmacologia. 1963; 4:184–205. [PubMed: 14048540]
29. Hill HE, Haertzen CA, Wolbach AB Jr, Miner EJ. The Addiction Research Center Inventory: Standardization of scales which evaluate subjective effects of morphine, amphetamine, pentobarbital, alcohol, LSD-25, pyrahexyl and chlorpromazine. Psychopharmacologia. 1963; 4:167–183. [PubMed: 14048539]
30. Isbell H. Methods and results of studying experimental human addiction to the newer synthetic analgesics. Ann N Y Acad Sci. 1948; 51(Art 1):108–122. [PubMed: 18890121]
31. 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. Heidelberg: Springer-Verlag; 1977. p. 197-258.
32. Jasinski DR. History of abuse liability testing in humans. Br J Addict. 1991; 86(12):1559–1562. [PubMed: 1786487]
33. Jasinski DR, Mansky PA. Evaluation of nalbuphine for abuse potential. Clin Pharmacol Ther. 1972; 13(1):78–90. [PubMed: 4550324]
34. Jasinski DR, Martin WR, Hoeldtke RD. Effects of short- and long-term administration of pentazocine in man. Clin Pharmacol Ther. 1970; 11(3):385–403. [PubMed: 4910901]
35. Johanson CE. Assessing the reinforcing properties of drugs. NIDA Res Monogr. 1989; 95:135–145. [PubMed: 2701296]
36. Johanson CE. Behavioral pharmacology, drug abuse, and the future. Behav Pharmacol. 1990; 1(4):385–393. [PubMed: 11175423]
37. Johanson CE, de Wit H. The use of choice procedures for assessing the reinforcing properties of drugs in humans. NIDA Res Monogr. 1989; 92:171–210. [PubMed: 2512492]

38. Kaplan HL, Busto UE, Baylon GJ, Cheung SW, Otton SV, Somer G, Sellers EM. Inhibition of cytochrome P450 2D6 metabolism of hydrocodone to hydromorphone does not importantly affect abuse liability. J Pharmacol Exp Ther. 1997; 281(1):103–108. [PubMed: 9103485]
39. Katz N. Abuse-deterrent opioid formulations: are they a pipe dream? Curr Rheumatol Rep. 2008; 10(1):11–18. [PubMed: 18457606]
40. Lalovic B, Kharasch E, Hoffer C, Risler L, Liu-Chen LY, Shen DD. Pharmacokinetics and pharmacodynamics of oral oxycodone in healthy human subjects: role of circulating active metabolites. Clin Pharmacol Ther. 2006; 79(5):461–479. [PubMed: 16678548]
41. Lamb RJ, Preston KL, Schindler CW, Meisch RA, Davis F, Katz JL, Henningfield JE, Goldberg SR. The reinforcing and subjective effects of morphine in post-addicts: a dose-response study. J Pharmacol Exp Ther. 1991; 259(3):1165–1173. [PubMed: 1762068]
42. Lasagna L, Von Felsinger JM, Beecher HK. Drug-induced mood changes in man. I. Observations on healthy subjects, chronically ill patients, and postaddicts. J Am Med Assoc. 1955; 157(12): 1006–1020. [PubMed: 14353639]
43. Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, Fiellin DA. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. Ann Intern Med. 2007; 146(2):116–127. [PubMed: 17227935]
44. Martin WR, Sloan JW, Sapira JD, Jasinski DR. Physiologic, subjective, and behavioral effects of amphetamine, methamphetamine, ephedrine, phenmetrazine, and methylphenidate in man. Clin Pharmacol Ther. 1971; 12(2):245–258. [PubMed: 5554941]
45. McCabe SE, Cranford JA, Boyd CJ, Teter CJ. Motives, diversion and routes of administration associated with nonmedical use of prescription opioids. Addict Behav. 2007; 32(3):562–575. [PubMed: 16843611]
46. McColl S, Sellers EM. Research design strategies to evaluate the impact of formulations on abuse liability. Drug Alcohol Depend. 2006; 83 (Suppl 1):S52–62. [PubMed: 16554125]
47. McLeod DR, Griffiths RR, Bigelow G, Yingling J. An automated version of the Digit Symbol Substitution Test (DSST). Behav Res Meth Instr. 1982; 14:463–466.
48. Petry NM, Bickel WK, Huddleston J, Tzanis E, Badger GJ. A comparison of subjective, psychomotor and physiological effects of a novel muscarinic analgesic, LY297802 tartrate, and oral morphine in occasional drug users. Drug Alcohol Depend. 1998; 50(2):129–136. [PubMed: 9649964]
49. Preston KL, Bigelow GE, Bickel WK, Liebson IA. Drug discrimination in human postaddicts: agonist-antagonist opioids. J Pharmacol Exp Ther. 1989; 250(1):184–196. [PubMed: 2473187]
50. Preston KL, Bigelow GE, Liebson IA. Comparative evaluation of morphine, pentazocine and ciramadol in postaddicts. J Pharmacol Exp Ther. 1987; 240(3):900–910. [PubMed: 3559982]
51. Preston KL, Liebson IA, Bigelow GE. Discrimination of agonist-antagonist opioids in humans trained on a two-choice saline-hydromorphone discrimination. J Pharmacol Exp Ther. 1992; 261(1):62–71. [PubMed: 1373189]
52. Preston, KL.; Walsh, SL. Drug Abuse Handbook. Boca Raton, Florida: CRC Press; 1998. Evaluating abuse liability: methods and predictive value; p. 276-306.
53. Roache JD. Performance and physiological measures in abuse liability evaluation. Br J Addict. 1991; 86(12):1595–1600. [PubMed: 1786492]
54. SAMHSA. NSDUH Series H-41, HHS Publication No (SMA) 11–4658. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2011. Results from the 2010 National Survey on Drug Use and Health: Summary of National Findings.
55. Schoedel KA, McMorn S, Chakraborty B, Potts SL, Zerbe K, Sellers EM. Positive and negative subjective effects of extended-release oxymorphone versus controlled-release oxycodone in recreational opioid users. J Opioid Manag. 2011; 7(3):179–192. [PubMed: 21823549]
56. Schoedel KA, McMorn S, Chakraborty B, Zerbe K, Sellers EM. Reduced cognitive and psychomotor impairment with extended-release oxymorphone versus controlled-release oxycodone. Pain Physician. 2010; 13(6):561–573. [PubMed: 21102969]
57. Shram MJ, Sathyan G, Khanna S, Tudor IC, Nath R, Thipphawong J, Sellers EM. Evaluation of the abuse potential of extended release hydromorphone versus immediate release hydromorphone. J Clin Psychopharmacol. 2010; 30(1):25–33. [PubMed: 20075644]

58. Simonson E, Brozek J. Flicker fusion frequency; background and applications. Physiol Rev. 1952; 32(3):349–378. [PubMed: 12983227]
59. Spiga R, Huang DB, Meisch RA, Grabowski J. Human methadone self-administration: effects of diazepam pretreatment. Exp Clin Psychopharmacol. 2001; 9(1):40–46. [PubMed: 11519634]
60. Stoops WW, Hatton KW, Lofwall MR, Nuzzo PA, Walsh SL. Intravenous oxycodone, hydrocodone, and morphine in recreational opioid users: abuse potential and relative potencies. Psychopharmacology (Berl). 2010; 212(2):193–203. [PubMed: 20665209]
61. Tompkins DA, Lanier RK, Harrison JA, Strain EC, Bigelow GE. Human abuse liability assessment of oxycodone combined with ultra-low-dose naltrexone. Psychopharmacology (Berl). 2010; 210(4):471–480. [PubMed: 20386884]
62. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS, Hewitt DJ, Jadad AR, Katz NP, Kramer LD, Manning DC, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robinson JP, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Witter J. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003; 106(3):337–345. [PubMed: 14659516]
63. Turk DC, Dworkin RH, McDermott MP, Bellamy N, Burke LB, Chandler JM, Cleeland CS, Cowan P, Dimitrova R, Farrar JT, Hertz S, Heyse JF, Iyengar S, Jadad AR, Jay GW, Jermano JA, Katz NP, Manning DC, Martin S, Max MB, McGrath P, McQuay HJ, Quessy S, Rappaport BA, Revicki DA, Rothman M, Stauffer JW, Svensson O, White RE, Witter J. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials. Pain. 2008; 139(3):485–493. [PubMed: 18706763]
64. 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(3):191–202. [PubMed: 18606504]
65. Walsh SL, Strain EC, Abreu ME, Bigelow GE. Enadoline, a selective kappa opioid agonist: comparison with butorphanol and hydromorphone in humans. Psychopharmacology (Berl). 2001; 157(2):151–162. [PubMed: 11594439]
66. Webster L. Update on abuse-resistant and abuse-deterrent approaches to opioid formulations. Pain Med. 2009; 10 (Suppl 2):S124–133. [PubMed: 19691683]
67. Webster LR, Johnson FK, Stauffer J, Setnik B, Ciric S. Impact of intravenous naltrexone on intravenous morphine-induced high, drug liking, and euphoric effects in experienced, nondependent male opioid users. Drugs R D. 2011; 11(3):259–275. [PubMed: 21902287]
68. Wechsler, D. The Measurement and Appraisal of Adult Intelligence. Baltimore: Williams & Wilkins; 1958.
69. Zacny JP, Drum M. Psychopharmacological effects of oxycodone in healthy volunteers: roles of alcohol-drinking status and sex. Drug Alcohol Depend. 2010; 107(2–3):209–214. [PubMed: 19948383]
70. Zacny JP, Gutierrez S. Characterizing the subjective, psychomotor, and physiological effects of oral oxycodone in non-drug-abusing volunteers. Psychopharmacology (Berl). 2003; 170(3):242–254. [PubMed: 12955305]
71. Zacny JP, Gutierrez S. Within-subject comparison of the psychopharmacological profiles of oral hydrocodone and oxycodone combination products in non-drug-abusing volunteers. Drug Alcohol Depend. 2009; 101(1–2):107–114. [PubMed: 19118954]

Summary Recommendation

We recommend that four subjective effects measures should be included in ALA studies of opioid analgesics: drug liking, likelihood to take again, drug identification, and drug high. These can be considered core primary outcome measures, that is, core in the sense that these measures should be part of any opioid ALA battery, and primary in the sense that they are important measures in determining the AL of a drug.

Table 1

Instruments used to measure self-reported subjective effects of opioids and other psychoactive drugs

Instruments	Comments
ARCI-Short Form (ARCI)	49 items are grouped into five scales • PCAG (measure of apathetic sedation) • MBG (measure of euphoria) • LSD (measure of dysphoria and somatic and sensory disturbances) • BG and A (measures of stimulant-like effects)
Single Dose Questionnaire (SDQ)	Measures four aspects of a drug's effects • effect detected (yes/no) • drug identification (e.g., opiate, cocaine, marijuana) • symptoms experienced (e.g., skin itchy, sleepy, nervous) • drug liking
Adjective Ratings	Consists of a list of symptoms • can be a checklist (yes/no) or an ordinal scale (e.g., 0–4) • responses are sometimes grouped into subscales (e.g., Agonist, Antagonist/Withdrawal)
Visual Analog Scale (VAS)	Typically a 10 cm (or 100 mm line) • endpoints are usually labeled with such descriptors as "not at all" and "extremely" • allows subject to make a more graded response in terms of intensity to which a certain effect (e.g., feel good) is perceived than a checklist

PCAG = Pentobarbital-Chlorpromazine-Alcohol Group; MBG = Morphine-Benzedrine Group; LSD = Lysergic Acid Diethylamide; BG = Benzedrine Group; A = Amphetamine.

Table 2

Recommended drug self-administration procedures

Procedures	Advantages	Disadvantages
Discrete trial choice: Drug versus drug Participants sample Dose A and Dose B on separate occasions (Dose A and B could be active and placebo, low and high doses of the same drug, or different doses of different drugs) and then are later asked to choose between Dose A and Dose B Response: Verbal response “Dose A” or “Dose B” Dependent measures: Number or percentage of doses chosen	**1** Simple **2** Experimenter controls the interval at which doses are available, which enhances safety **3** Shows good sensitivity to dose conditions (active drug is generally chosen more than placebo, and high doses more than low doses) **4** It is possible to test the relative reinforcing effects of two or more different drugs	**1** Few dependent measures **2** Requires many sessions in order to examine dose-response relationships using a within subjects (crossover) design
Non-verbal operant choice procedure: Modified progressive ratio (PR) schedule Participants sample a dose of drug and money and are later given the opportunity to “work” for that dose and/or money by responding on a manipulandum; the ratio requirement increases after completion of the previous ratio requirement, but drug and money not delivered until all responding has been made Response: Finger presses on a manipulandum (computer mouse, joystick), pedals on a bicycle, etc. Dependent measures: Amount of drug consumed, amount of money earned, rate of responding, maximum ratio value completed (breakpoint)	**1** Provides an estimate of the reinforcing efficacy of a drug, which allows comparisons to be made across different drugs (theoretically, lower efficacy drugs maintain lower breakpoints than higher efficacy drugs) **2** Generally sensitive to dose manipulations	**1** In the absence of an alternative reinforcer, the slope of the dose-response function can be shallow (i.e., very low doses of drug, including those that produce minimal subjective effects, can maintain breakpoint values that are significantly different from placebo)

Table 3

Domains of opioid human abuse liability laboratory testing and some recommended endpoints to assess those domains

Domain	Endpoints
Subjective effects (Primary endpoint)	Liking Likelihood to take again Drug identification Drug high
Physiological responses (Secondary endpoint)	Pupillary constriction (miosis) Respiration rate and oxygen saturation
Drug self-administration behavior (Optional endpoint)	Discrete trial choice: drug vs. drug Nonverbal operant procedure: modified progressive ratio schedule
Cognitive and psychomotor performance (Optional endpoint)	No specific recommendations at this time