EDITORIAL

# Striving for Consensus on Approaches to Category 1 Testing of Abuse-Deterrent Formulations of Opioids: Discussions from the First Category 1 Focus Group Meeting

The development of opioid formulations with potentially abuse-deterrent (AD) properties has gained considerable momentum over the last decade with the introduction of opioid products with unique properties that are designed to limit unintended routes or methods of drug administration. The science of assessing AD formulations (ADFs) is being refined as new ADFs are developed and submitted for regulatory approval. Recent guidance (Guidance for Industry: Abuse-Deterrent Opioids–Evaluation and Labeling, 2015) from the US Food and Drug Administration (FDA) serves as the framework for the testing of ADFs of opioids.[1] The guidance describes 3 categories of premarket studies that should be performed to provide a thorough characterization of AD features of novel ADFs.[1] Category 1 consists of laboratory-based in vitro manipulation and extraction studies to evaluate how easily the AD characteristics of a formulation can be compromised or defeated. Category 2 studies evaluate pharmacokinetics, and Category 3 studies assess clinical abuse potential. A final category, Category 4 studies, uses postmarketing data to analyze the impact of an ADF on actual abuse.[1]

From the regulatory perspective, a flexible and adaptive approach that takes into account the properties of the test product, the particular active pharmaceutical ingredient (API) or ingredients, and the anticipated routes of unintended use is needed. The totality of the evidence from all categories of testing is weighed when reviewing data assessing AD properties of a specific test product. Therefore, the more comprehensive the Category 1 data and results are, the better will be the understanding of the AD properties. Further, findings from Category 1 studies can provide useful information for the design of Category 2 and 3 studies and are relevant for the interpretation of data from Category 3 clinical abuse potential studies (Figure 1). The final "to-be-marketed" formulation should be used in all Category 1 testing, and this formulation needs to be well understood from both a composition and a manufacturing process because both may affect its AD properties. The agency defined the final "to-be-marketed formulation" as the formulation having the final composition, using production-scale equipment, and final process parameters. It is the experience of the agency that minor changes in the production setup not affecting efficacy and safety of the product (changes within the Scale-up and Postapproval Change Modified Release guide) might influence the product's AD properties.

Assessment of the degree of effort needed to defeat AD properties of a formulation is essential to Category 1 characterization (Figure 2).[1] The FDA, however, does not provide any specific recommendations on how to measure the level of effort or incorporate the results of such measurements into the overall profile of ADFs. The design of the testing methodologies is highly dependent on the specific nature of the opioid, the AD features of the formulation (eg, mechanism of abuse deterrence), and methods anticipated to be utilized by abusers to manipulate and administer the drug. ADFs can be diverse, with AD properties consisting of one or more of the following features: physical/chemical barriers, agonist/antagonist combinations, aversion, delivery systems, new molecular entities and prodrugs, or novel approaches.[1] Furthermore, companies are obliged to take into consideration a concern of the FDA that those who manipulate ADFs will adapt to new products and

DOI: 10.1111/papr.12488

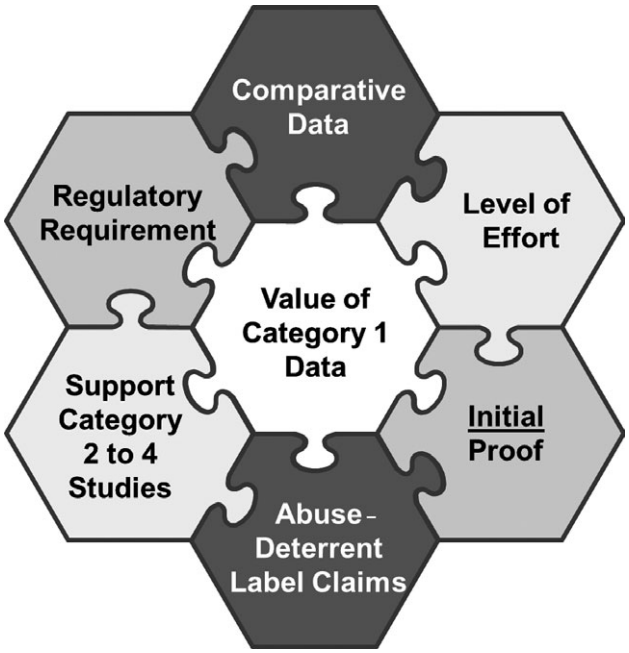


**Figure 1.** Key areas of value added by Category 1 data in the development of abuse-deterrent formulation drugs.

find new ways of defeating them. In cases when Category 4 postmarketing studies show that an AD technology no longer impedes unintended use, the FDA may require labeling revisions.

Category 1 methods and data are rarely presented or published, in large part because of concerns about providing a recipe for individuals seeking to manipulate prescription opioids to defeat specific AD technologies or ADFs. Thus, there is a need for a forum between representatives of the pharmaceutical industry, regulatory authorities, and academia to discuss the design, types of data, data interpretation, and current best practices for Category 1 studies. Another rationale for generating a group focusing solely on Category 1 is that, although several other forums already exist, such as the Cross Company Abuse Liability Consortium (CCALC) meetings, ExL Events, and CBI conferences on abuse deterrence, their emphasis is on more clinical and general regulatory aspects. In these meetings, Category 1 studies are generally discussed in light of their support for Category 2/3 studies rather than as independent sources of abuse-deterrence data. As the laboratory environment and study designs of Category 1 in vitro studies started to grow into its own field of science, the need for having discussions focused on Category 1 was identified. With the initiative of Egalet Corporation (Wayne, PA, U.S.A.; represented by author KL) and in collaboration with the CCALC (represented by author MS) and a leading Category 1 expert (author EJC), a core group of stakeholders was identified, leading to the initiation of discussions and an initial meeting of the Category 1 Focus Group.

The charter of the Category 1 Focus Group is to (1) share experiences and best practices in in vitro (Category 1) assessments of ADFs; (2) discuss standardization of in vitro (Category 1) studies and limits associated with established standard practices; (3) provide an open forum for discussion on the FDA Guidance for Industry: Abuse-Deterrent Opioids–Evaluation and Labeling; and (4) provide an opportunity for presentations by academia, regulatory, and industry experts focused on Category 1 studies. The purpose of this communication is to present a high-level summary of important issues that were discussed at the first meeting of the Category 1 Focus Group held on November 4, 2015. The main topics of discussion were standardization of testing, methods to incorporate and account for the level of effort during manipulation in the overall AD profile of a formulation, and real-world experiences from industry and regulatory perspectives applying the FDA guidance for regulatory submissions.

Category 1 studies consist of physical and chemical manipulation procedures with a variety of commonly used household tools and solvents to determine whether the AD features of a test product are compromised by such manipulations. Physical manipulation of ADFs is

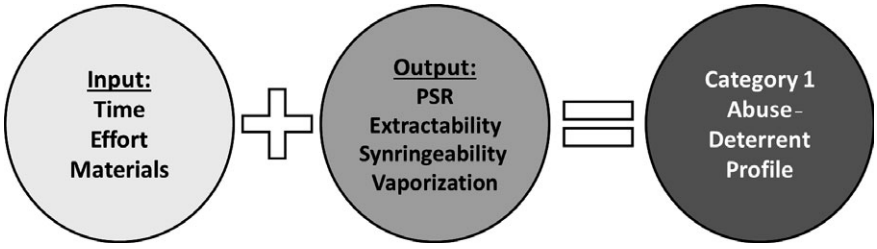


**Figure 2.** Elements of a thorough Category 1 characterization of an abuse-deterrent formulation. PSR, particle size reduction.

attempted with common tools (eg, hammer, pill crusher, knife, razor blade, file, and grater) or electrical appliances (eg, spice grinder and coffee grinder) that crush, cut, grate, or grind the test product. The goal of physical manipulations is to reduce a tablet to a small particle size for either intranasal use or as the starting point for more complex tampering procedures that might render a product amenable to other routes of administration. Photography is used frequently as documentation of the results of Category 1 testing. The outcomes from the physical manipulation are typically described in terms of particle fractions are analyzed to evaluate the homogeneity of opioids across the bands of particle size fractions. Additional descriptive elements can provide documentation of the state of the test product (eg, "completely powdered", "some flaking", or "remains intact"). Products that can readily be reduced to a "snortable" powder would be the most desirable, and products that remain intact or largely intact would be the least desirable.

Chemical extraction tests are usually performed on the intact and manipulated products using solvents that are commonly available (eg, water, ethanol, and vinegar) and ingestible. These tests help determine whether the product can be prepared for injection and for administration by other routes (eg, oral and rectal). More complex singlestep and multistep procedures with immiscible organic solvents (noningestible and toxic) may also be evaluated. In addition to the importance of choosing the methods of physical and chemical manipulations to be used in a Category 1 study, it is also important to include an appropriate comparator product. Generally, procedures that do not include appropriate comparator(s) are less informative and results are difficult to interpret. However, even comparisons with the most appropriate comparator may prove challenging with respect to interpreting the results of the manipulations of the 2 products in relation to outcome data from all of the Category 1 through Category 3 studies. It is also important to relate the different manipulations to the different anticipated routes of administration. For example, extraction in large volumes of solvent would be best suited to the oral route but not to intravenous use (unless significantly more work was expended to reduce volume); generally, use by the injection route requires extraction in small volumes.

Category 1 studies should be, by design, iterative in nature. Because of the unique nature of test procedures employed (many of which have never been applied to a product by the innovator), there is occasional "discovery" of formulation properties that were not anticipated (eg, loss of the API). Consequently, as new knowledge about a product emerges, there may be need for further exploration of the phenomenon. Thus, careful data review should occur as studies progress. Not uncommonly, test results may reveal the need for further investigation of the "boundaries" of a specific formulation beyond what was originally planned.

On the surface, it may seem relatively straightforward to use a standardized set of tools (ie, same type and model), solvents (acidic/basic, organic, or alcohols, depending on the solubility of the API), or procedures to perform Category 1 studies. However, there are many factors (eg, the force of the action, the strength of the person, the time and duration of the action, and the sharpness of the tool [in the case of mechanical procedures]) that can contribute to variability between individual testers and laboratories. Therefore, reproducibility of test procedures between laboratories may be difficult. Although FDA guidance provides the basis for the type of studies and data encompassing Category 1 evaluations, methodological diversity exists between studies with different ADFs because of the need for "tailoring" study designs to fit each specific product. For example, the AD features of a specific opioid formulation may be designed to resist manipulation resulting in forms that are most commonly administered by certain routes of administration (eg, snorting, intravenous, and chewing). It is known that specific opioids are often taken by a variety of routes of administration that vary from compound to compound.[2,3] Although standardization of Category 1 testing is desired, this desire must be tempered by the reality that the testing should also be tailored to the specific AD features of each novel test product.

This need to tailor the individual tests for each specific test product was illustrated by the approaches used for Category 1 testing of 3 different products (Hysingla® Extended Release [ER], Targiniq™ ER, and OxyContin®) from the same pharmaceutical company (Purdue Pharma L.P., Stamford, CT, U.S.A.). Two of the test products (Hysingla ER and OxyContin) have physical and chemical barriers designed to resist manipulation, and the other contains an agonist/antagonist combination. The same type of testing was performed on the 2 products with the physical and chemical barriers; however, because of differences in the API, size of the tablets, and ratio of excipients to API, the testing procedures had to be modified for the second product. For the agonist/antagonist product (Targiniq ER),

testing procedures needed to be designed for both components and to determine whether the agonist could be separated from the antagonist and whether differential degradation of the 2 components could be achieved. Thus, each ADF has unique AD properties and characteristics that must be explored with specifically designed testing procedures.

Some examples of how it might be possible to standardize approaches were presented by Sebastian Schwier (Grünenthal GmbH, Stolberg, Germany) and Torben Elhauge (Egalet Corporation). Schwier presented the development of a standardized hammer apparatus with the goal of developing a method that excluded the "human factor" and established a flexible test method reflecting real-life conditions in a reproducible manner. Questions were raised as to whether other test methods could be standardized in a similar manner. It also remains to be demonstrated that the apparatus is useful for testing other ADFs.

Elhauge focused on the outcomes of small volume extraction and the test for syringeability/injectability. A correlation between viscosity and syringeability (eg, whether a tablet dissolved in a small volume of solvent could be expelled through a certain needle size) could potentially lead to a standardized way of translating viscosity into interpretation of syringeability/injectability tests. The test for syringeability/injectability is an example of a test that most likely will be dependent on the person performing the test and the equipment used. The findings also highlighted the difficulty in reaching conclusions from syringeability/injectability studies because it is unclear whether an abuser will deem the resulting product suitable for injection, even though it may "pass" a laboratory-based test. Data from Category 4 postmarketing studies, which provide insights into what is relevant for those who abuse opioids, are needed to improve Category 1 test methods. Indeed, Category 4 data will inform the design across categories of testing, which will serve to strengthen the relationship between results of premarket studies and the ability of ADFs to deter real-world abuse.

An example of an instrument that provides a subjective rating of the degree of effort required for physical manipulation of ADFs containing "hardened" tablet features was also discussed. The ALERRT® instrument (PinneyAssociates, Bethesda, MD, U.S.A.), which uses 100-mm visual analog scales (0 = "very easy"; 100 = "extremely difficult") and was developed to assess the labor, effort, and resources required for tampering, allows for a quantitative measurement of the degree of effort required for physical manipulation using 10 tools and methods commonly employed by drug abusers.[4] Scores on the high end of the scale denote "extremely difficult to defeat," while scores on the low end of the scale denote "very easy to defeat." Ideally, and even after the maximum amount of time allowed for manipulation had been attained, the formulation would remain intact or largely intact in a form that would not be amenable to alternative routes of administration such as intravenous or intranasal. Although the ALERRT instrument is an example of an approach that can help standardize the assessment of the amount of effort needed to physically manipulate ADFs, its use may not be appropriate for all types of ADFs, such as nonhardened tablets. To develop a comprehensive profile of the properties of an ADF, it may be necessary to combine different assessment approaches and test methods.

## SUMMARY

A total of 34 participants responded to a postmeeting survey; all indicated that they found the meeting informative and, as a consequence, a second meeting is being planned for 2016. As part of the survey, a number of topics for the next meetings have been suggested and will be considered for future meetings:

1. How should a standardization program be established, if possible?
2. Both from a regulatory and industry perspective, how can the Category 1 test programs be optimized to expand testing in areas of special relevance and reduce testing in less relevant areas?
3. How can data from testing that goes beyond "common abuser practice" be interpreted and can the data be used in proposed label text?
4. How can potency be included in the AD evaluation?
5. How do Category 1 data relate to epidemiological evidence of abuse deterrence?
6. How can use of Category 1 data in educational material be optimized to further add value to Category 2/3 data?
7. How can the abuser community be followed to adjust criteria on a regular basis? Can Category 1 data be used to support that?
8. Is there a relation between AD features and diversion of certain drugs?

Additional suggestions for topics and/or speakers for the next meeting can be sent to Karsten Lindhardt (klindhardt@egalet.com).

A website has been developed for the focus group that provides details about the group and the initial meeting.[5] The website can be used as an avenue for information about past and future meetings and serves as a forum to generate and foster discussions about key issues affecting Category 1 testing.

In conclusion, the development and testing of ADFs of opioids in accordance with FDA guidance is relatively new and evolving. Category 1 testing requires a flexible and adaptive approach to tailor procedures for the unique AD features of novel ADFs; however, there may be areas that are suitable for standardization of testing for all ADFs. The Category 1 Focus Group is clearly well suited to provide leadership and key information and to serve as a forum that meets the needs of the stakeholders in the development of ADFs. Furthermore, as new AD technologies are developed, the methods used for Category 1 testing will likely change, making comparison of data to other ADFs challenging. The Category 1 Focus Group will monitor the impact of new AD technologies on Category 1 testing methodologies and consider when standardized methods are appropriate. By involving participants from the pharmaceutical industry, regulatory authorities, and academia, an open dialogue can be maintained as the Category 1 testing evolves. Finally, by sharing experiences and knowledge about Category 1 testing and interpretation of data, the Category 1 Focus Group can gain a better understanding of how the results of Category 1 testing can become a more integral part of the totality of evidence used for the evaluation of ADFs.

## DISCLOSURES

E.J.C. Employee of PinneyAssociates, which has received funding from Egalet Corporation, Purdue Pharma (focused on pharmaceutical risk management and abuse-deterrent drug formulation evaluation), GlaxoSmithKline Consumer Healthcare on their stop-smoking medications (Nicorette and NicoDerm CQ in the U.S.), for NJOY, Inc., a developer and marketer of electronic nicotine delivery systems, and since February 2015, for Reynolds American, Inc. (RAI) on tobacco harm minimization. Our work for RAI focuses on products, regulations, and policies related to smoking cessation and harm minimization; we do not work on combustible conventional cigarettes. E.J.C. also owns intellectual property for an as-yet not-commercialized nicotine gum, an option for which has been acquired by Niconovum, a subsidiary of RAI. Our work with Purdue Pharma has focused on pharmaceutical risk management and abuse-deterrent drug formulation evaluation. M.S. Employee of Grunenthal U.S.A. (developer and manufacturer of abuse deterrent technologies). K.L. Employee of Egalet Corporation; owns Egalet stock or stock options.

## ACKNOWLEDGEMENTS

The authors would like to thank Patrick Little, PhD, of Complete Publication Solutions, a CHC Group Company (North Wales, PA, U.S.A.), for providing medical writing support, half of which was funded by Egalet Corporation (Wayne, PA, U.S.A.), and half of which was funded from attendee fees and sponsors' contributions for the Category 1 Focus Group meeting. The authors would also like to thank all the speakers and meeting sponsors for their valuable support. Lastly, the authors would like to thank Torben Elhauge from Egalet Corporation (Wayne, PA, U.S.A.) for his hands-on help in making this a successful meeting and continued support for the Category 1 Focus Group work.

Edward J. Cone, PhD*
Marta Sokolowska, PhD†
Karsten Lindhardt, MSc, PhD, DBE‡
*PinneyAssociates, Bethesda, Maryland;
†Grünenthal USA Inc., Morristown, New Jersey;
‡Egalet Corporation, Wayne, Pennsylvania, U.S.A.
E-mail: KLindhardt@egalet.com

## REFERENCES

1. US Food and Drug Administration. *Guidance for Industry: Abuse-Deterrent Opioids – Evaluation and Labeling*. Silver Spring, MD: US Department of Health and Human Services; 2015.

2. Butler SF, Budman SH, Licari A, et al. National addictions vigilance intervention and prevention program (NAVIPPRO): a real-time, product-specific, public health surveillance system for monitoring prescription drug abuse. *Pharmacoepidemiol Drug Saf*. 2008;17:1142–1154.

3. Butler SF, Black RA, Cassidy TA, Dailey TM, Budman SH. Abuse risks and routes of administration of different prescription opioid compounds and formulations. *Harm Reduct J*. 2011;8:29.

4. Cone EJ, Buchhalter AR, Lindhardt K, Elhauge T, Dayno JM. Extreme work requirement of EG-001, an abuse-deterrent extended-release morphine product, as demonstrated with the ALERRT[SM] visual analog scales. *Postgrad Med*. 2015;127:S26.

5. Category 1 Focus Group Meeting. http://www.cat1meeting.com/ (accessed January 21, 2016).