Development of Novel Endpoints for Clinical Trials in Substance Use Disorders Working Group Manuscript Draft Discussion

Chair: Tanya Ramey MD, PhD
Co-Chair: Martin Mumenthaler, PhD

ISCTM Annual Meeting Feb 21, 2025, Washington DC

Disclosures

Tanya Ramey – Medical Officer, Program Director, NIH/NIDA, Division of Therapeutics and Medical Consequences (DTMC)

Martin S Mumenthaler -Adjunct Professor, Stanford University, Department of Psychiatry and Behavioural Sciences; Consultant for AfaSci, Inc

Welcome!

INTRODUCTIONS

Goals of the session:

- **Discuss the manuscript draft. Current title**:
  "Development of Novel Clinical Trial Endpoints for Opioid Use Disorder and Stimulant Use Disorder"
- **Review current sections and discuss streamlining options**
- **Work on the manuscript sections**

At the conclusion of the session:

1page MINUTES with Next Steps.

Manuscript Current Outline

1. **Introduction**
2. **Endpoints in RCTs for Opioid Use Disorder (OUD) and Stimulant Use Disorder (StUD)**
3. **Regulatory: FDA Guidance for OUD and Draft Guidance for StUD. Added the DDT Qualification Letter received on Feb 14, 2025 for the Alcohol Use Disorder. &Heddie Martynowicz contribution**
4. **Heterogeneity of the SUD population**
5. **Novel behavioral endpoints, new assessments Brian Kiluk**
6. **Assessment instruments for novel endpoints:** (PROs such as **TLFB** and **EMA** vs. objective measures)
7. **Other potential novel endpoints related to patient functioning- social functional domains (Recovery Capital)- PENDING, Discuss interest level. Economic endpoints could be suggested.**
8. **Novel statistical approaches: Vlad Dragalin text awaited to be added after ISCTM meeting**
9. **Beatrice Setnik's tables – where we put them? Addendum is suggested**
10. **Conclusions**

1. Introduction

- Opioid & other drug epidemic – Opioid Crisis in the United States that is being described as a public health catastrophe of modern times (2017). Overdose (OD) deaths are increasing exponentially as well, e.g., more than 10 times increase from 2009 to 2019, according to the CDC data. The overdose situation is compounded by stimulants that are frequently taken with opioids, such as heroin, or fentanyl, or when stimulants are laced with these drugs and users are unaware.
- Urgency of a public health crisis requires rapid efforts from all of us to develop new medications and therapeutics new and expanded treatment options for Opioid Use Disorder (OUD) and Stimulant Use Disorder (StUD).
- Despite many recent technological and biological advances, the development of new medications and therapeutics in SUD hasn't caught up.
- **None** of the ten largest pharmaceutical companies have active addiction medicine programs or drug candidates, and the pharmaceutical industry as a whole has only pursued minimal substance use-focused drug development According to the Biotechnology Innovation Organization (BIO)

Introduction cont.

- **Currently** there are **3** FDA-approved pharmacotherapies to treat OUD – methadone, buprenorphine and naltrexone (not taking into account the opioid overdose reversal medications naloxone & nalmefene),
- There are no FDA-approved pharmacotherapies to treat StUD.

- Drug and therapeutics development can only be successful as long as and endpoint – a targeted outcome in a feasible randomized controlled trial (RCT) can be statistically analyzed to determine efficacy and safety of the drug under study. Clinically endpoint should reflect the subject's symptoms, adhere to the FDA's definition of "how patient feels, functions, or survives" and be sensitive to change.
- As a result of FDA's direct meetings with the public on patient-focused drug development for OUD (2018),medical devices for OUD(2022), StUD (2020) the FDA accepted the harm reduction approach to drug use.
  - Total abstinence is still an ideal outcome, though harm reduction allows for a decreased substance use, as an endpoint, and can potentially include novel endpoints detecting salutary benefits of decreased drug use on physical health, health resources utilization, social and societal roles.
- Novel, non-abstinence - focused endpoints for OUD and StUD need to be developed and accepted by the regulatory in order to progress the addiction field and save human lives, counter the US drug use epidemic.

2. Endpoints in RCTs for Opioid Use Disorder (OUD) and Stimulant Use Disorder (StUD): a historical perspective, challenges with **binary** endpoint & limitations. (Tanya)

- Dichotomous (use vs. no use) abstinence outcome assessed by self-report and confirmed by UA is the most common historical and current primary endpoint in a clinical trial for OUD and StUD.
- Dichotomous endpoint present a major challenge in assessing efficacy in an RCT, especially in heterogeneous substance use population. Unfeasible sample size needed to achieve adequate statistical power. As a result, underpowered 'pilot' clinical trials are abounded in heterogenous study populations, and sometimes simplified conclusions can be made that may not be unequivocally supported from a statistical standpoint. Results are not replicable.

WHILE

- Other mental disorders typically have a continuous endpoint and corresponding assessment
- Innovation in statistical approach to data analysis in SUD RCT is urgently needed. There are precedents from other therapeutic areas that deemed to be acceptable by the regulatory

Endpoints in RCTs for Opioid Use Disorder (OUD) and Stimulant Use Disorder (StUD) cont. (Tanya)

- Recently regulatory stance had changed re the requirement of absolute abstinence as a primary endpoint for SUD indication. For OUD, for example, in the FDA Guidance for the Industry (2020) **drug use patterns other than abstinence** can be used as thresholds to define treatment response.
- For the StUD FDA prefers the phrase "change in the pattern of stimulant use", as opposed to the phrase "reduction in stimulant use" and recommends its use to emphasize that **within-subject responses** are of interest. (This opens up a potential innovation portal for statistical strategies)
- The **proportion of subjects** achieving a target pattern of use days per specific period is currently stated as an acceptable endpoint.

AND LONG-AWAITED

- Major breakthrough for Alcohol Use Disorder: Feb 14, 2025 the FDA's CDER had announced that Alcohol Use Disorder have now a new Drug Development Tool qualified endpoint: two-level reduction in risk drinking level (RDL) of alcohol consumption and was validated as clinically meaningful. As a result, it can be used as an acceptable primary endpoint in studies of medications to treat adults with moderate to severe AUD.

3.FDA Guidance for OUD and Draft Guidance for StUD: features that impact the clinical trial design for clinical trialists – -suggestions for improvement

- **FDA's 2019 guidance addresses clinical endpoints acceptable for demonstrating the effectiveness of drugs to treat OUD**.
- In general, clinical trials evaluating the effectiveness of drugs for treating OUD have used reduction in drug-taking behavior (drug use patterns) as an endpoint. FDA acknowledges in their guidance great interest in expanding the primary and secondary endpoints used in clinical trials of drugs for treating OUD, including other outcome measures important to patients and their families, clinicians, and the public. The following are outcome measures that could potentially be used as primary and/or secondary endpoints in clinical trials and included in FDA-approved labeling:
- **Reductions in adverse outcomes related to OUD are desirable endpoints** for study, Examples of these adverse outcomes include: mortality (overall mortality or overdose mortality), need for emergency medical interventions and hepatitis C virus infection or reinfection. Sponsors can propose other adverse outcomes and can study several of these endpoints in the same trial, selecting one as the primary endpoint, selecting one or more as secondary endpoints, or combining outcomes in a composite endpoint.

FDA Guidance for OUD and Draft Guidance for StUD cont.

- **Change in Disease Status Using Diagnostic Criteria for OUD**: FDA recommends that, if all trial patients meet the DSM-5 criteria for moderate-severe OUD at baseline, sponsors could use the **proportion of patients meeting DSM-5 criteria for remission of OUD at the end of the trial as a primary or secondary efficacy endpoint.**
- **Patient-Reported Outcomes**: Using input from patients and family members to determine the most concerning symptoms/experiences associated with OUD, sponsors could develop a patient-reported outcome (PRO) instrument to evaluate a direct effect on how patients feel or function (e.g., improvement **in sleep or mood**). Outcomes on this measure could be used as a secondary endpoints in trials that use behavioral change, such as change in drug use patterns, as a primary endpoint.
- **Change in Drug Use Pattern**: Change in drug use pattern is the most commonly used endpoint in registration trials for drugs in development to treat OUD. Sponsors have used it successfully to provide support of efficacy for all approved products for treating OUD. **A commonly used definition for a responder - is abstinence, defined as no detected or self-reported use during the specific assessment window. Sponsors can employ drug use patterns other than abstinence to define response to OUD treatment.**

FDA Guidance for OUD and Draft Guidance for StUD cont.

- **FDA's 2023 draft guidance** provides the FDA's current recommendations regarding the overall development program and clinical trial designs for the development of drugs to support indications for treatment of **moderate to severe cocaine use disorder, treatment of moderate to severe methamphetamine use disorder, or treatment of moderate to severe prescription stimulant use disorder.**
- **DSM-5 has a single diagnosis, stimulant use disorder**, defined as "a pattern of amphetamine-type substance, cocaine, or other stimulant use leading to clinically significant impairment or distress," ranging from mild to severe. The definition lists various symptoms of impairment or distress, but notably, it does not include any criteria related to amount or frequency of stimulant use. The group of individuals meeting DSM-5 criteria for stimulant use disorder is very heterogeneous, with individuals using different stimulants in a range of different settings and for different reasons. This heterogeneity may contribute to the difficulty in identifying medications that are efficacious for the entire subset of patients diagnosed with cocaine use disorder or methamphetamine use disorder, and even more for all patients meeting the broader criteria for stimulant use disorder. Cocaine, methamphetamine, and other stimulants have different mechanisms and effects, and this may lead to differences in clinical presentation and responses to treatment.
- **Currently there are no approved treatments for stimulant use disorder, thus sponsors should engage the division early in the drug development process**.

FDA Guidance for OUD and Draft Guidance for StUD Excerpts

- **Sponsors may consider demonstrating an effect in one or more of the following endpoint options**: change in the pattern of stimulant use, change in disease status using diagnostic criteria, and other potential outcome assessments.
- **Change in Pattern of Stimulant Use:** frequency measures are more feasible to measure. FDA guidance recognizes the need for a certain amount of reliance on self-report it highlights that it is not persuasive by itself. With regards to urine drug tests, the guidance states that there is no evidence to support a frequency of biological testing and suggests a balance between burden to the subject and some degree of biological confirmation of drug use self-report. It also states that a "sustained period of negative urine toxicology findings, indicating abstinence, could be a valid surrogate for clinical benefit"
- **Change in Disease Status Using Diagnostic Criteria**: Sponsors should enroll trial subjects who meet DSM-5 criteria for moderate to severe stimulant use disorder at baseline, based on clinical interview. These criteria include a variety of symptoms and reflect how subjects feel and function.
  - remission to be used as a specifier. After criteria for stimulant use disorder were previously met, early remission is defined as meeting none of the criteria for stimulant use disorder for between 3 and 12 months, and sustained remission is defined as meeting none of the criteria for at least 12 months.
  - **A suitable primary endpoint could be the proportion of subjects meeting criteria for early remission from stimulant use disorder at the end of the trial.**
- **Use of Other Clinical Outcome Assessments**: A suitably developed, fit-for-purpose measure that assesses relevant aspects of a subject's health status, functioning, and/or symptoms may be appropriate as a primary endpoint for a clinical trial and may be the most suitable approach for some investigational drugs.
- https://www.fda.gov/media/172703/download
- **Both FDA Guidances discuss potential acceptability of a fit-for-purpose measure of CRAVING.**

4. Heterogeneity of the SUD population: ways to reduce noise in substance use disorders clinical trials: transdiagnostic deep phenotyping battery use for enrichment & anomalous subjects' identification (Tanya, Bruce)

- **Heterogeneity factors**:
  - genetics
  - psychiatric comorbidities
  - Somatic illness comorbidities
  - behaviors, including patterns and contexts of substance abuse
- SUD diagnosis as per the DSM-V TR, with the determination of its severity, relies on a numerical amount out of 11 criteria that the subject must meet in any combination in the last 12 months: mild – 2-3 symptoms are present; moderate- 4- 5 symptoms present, and severe- presence of 6 or more symptoms.
- There are, for example, close to a 1,000 combinations of 3 criteria out of 11, that could be made. This situation alone brings about extreme heterogeneity in populations under study in any SUD clinical trial.
- How can we reduce heterogeneity "noise" in such a situation?
- One way is to use either an increased number or strive for a strict set of exclusion criteria to have a more homogenous study population. This could lead to a clinical trial that is exceedingly difficult to enroll in, if not impossible, let alone that can have a generalizability problem.

Heterogeneity of SUD population & RCT Methodology (Tanya, Bruce)

- **Enrichment approach:** of the study population with the desired patient profile from the outstart, using a screener deep phenotyping battery, for example. NIDA had developed one such battery (NIDA PhAB) that is feasible for deep phenotyping during a screening visit with ~ 3 hours to complete, and there is also a very brief version (PhAB-B) that takes much lesser amount of time ~ 23 to 35 min to complete (Keyser-Marcus et al., 2021; Parlier-Ahmad et al., 2023). **See the poster at the ISCTM 2024 Autumn meeting.**

- **Precision medicine approach: that targets available treatments to homogeneous characteristics** possessed by patient subgroups may offer more effective treatment to the individual patient

- To reduce the noise in the heterogeneity of treatment effect in sub-populations, an **adaptive enrichment design** or a **basket trial design** might be appropriate.

5. Novel behavioral endpoints in pipeline: continuous measures, new assessments: 1) Craving 2) Change in addiction severity, 3) Change in use frequency

• 1) Craving (Cecilia Bergeria)

- Craving is a symptom/endpoint fundamental to substance use disorder in general and OUD in particular, that it is included in the Diagnostic and Statistical Manual (DSM-5-TR), but there is no agreed-upon way to assess craving. **Craving represents a critical unmet treatment endpoint and an intermediate variable that contributes to poor treatment outcomes among individuals with OUD**.
- Accurately measuring and effectively treating opioid craving has the potential to improve treatment retention and prevent non-prescribed opioid use among individuals in treatment for OUD (Vafaie and Kober, 2022).
- There is little scientific consensus on how to assess opioid craving (Kleykamp et al., 2019). For example, there is variability across assessments in how frequently opioid craving is assessed and how many dimensions are included in an opioid craving assessment.
- **The FDA formally accepted into the Drug Development Tool Qualification program a NIDA-funded grant – a program for opioid craving assessment (DDT#000138) (Cecilia Bergeria).**

Novel behavioral endpoints in pipeline-continuous measures, new assessments cont.

- **Craving cont.**
- The resultant preliminary assessment has been designed to assess tonic craving over a two-week period. The preliminary **Opioid Craving Assessment** examines 6 dimensions, including (1) Preoccupation, obsessive thoughts about opioids, (2) Anticipation of negative reinforcement, (3) Anticipation of positive reinforcement, (4) Drive, intent to use, (5) Feeling lack of control, and (6) Tension and uneasiness related to opioid use.
  - Within each dimension are 4-5 items that are based on existing opioid craving assessments which query similar dimensions and/or patient feedback gathered in structured cognitive interviews.
- The preliminary version of the assessment will soon undergo exploratory and confirmatory factor structure analyses which will reduce the number of items. Subsequent studies are planned to test the reliability and validity of the assessment.

Novel behavioral endpoints in pipeline-continuous measures, new assessments

- **2) Change in addiction severity (Brian Kiluk)**
- A biological measure of drug use is considered only a surrogate marker and not the disorder under treatment. Rather, **the most meaningful outcome of treatment includes improvement in the key physical and psychosocial problems/consequences that characterize the disorder**.
- DSM criteria are used to define, diagnose, and determine the severity of SUD and to indicate the need and type of treatment required, but are rarely repeated or used to assess treatment response in clinical trials.
- FDA strongly encourages sponsors interested in using endpoints reflecting disease severity to select, modify, or develop suitable instruments.
- Clinician-administered, structured interviews for assessment of SUD diagnosis and severity are costly, require extensive training, and involve lengthy administration times, and thus impractical for repeated assessment in clinical trials.

Novel behavioral endpoints in pipeline-continuous measures-cont.

- Patient-reported DSM-5 symptom checklists (e.g., Substance Use Symptom Checklist) appear to be practical tools to aid SUD diagnosis in clinical settings but assess the presence of symptoms (yes/no) over the past year. There is a need to evaluate variation in the severity of specific DSM-5 symptoms/criteria rather than a simple criterion count as a measure of SUD severity and need for a shorter timeframe to be more responsive to change.
- **An instrument measuring the severity of SUD symptoms** would provide a direct indicator of change in the disorder's defining characteristics. Funds from NIDA are supporting **a project to develop a novel Patient Reported Outcome instrument measuring the severity of OUD according to DSM-5 symptom criteria (the Opioid Use Disorder Severity Scale; OUDSS).**
- .

Novel behavioral endpoints in pipeline-continuous measures-cont.

- Following PROMIS methodology for item development items were written in a first-person, past-tense format, with 5-response options reflecting frequency experiencing each item over the past 30 days (“0 – never”, “1 – rarely”, “2 – sometimes”, “3 – often”, “4 – almost always”); consistent with the PROMIS item bank

- Following item generation, two rounds of 1:1 cognitive interviews have been conducted with a heterogeneous sample of patients with OUD.
- The OUDSS was reduced to 50 items based on patient input. An observational study is ongoing to evaluate the OUDSS reliability, validity, and ability to detect change among a racially and ethnically diverse sample of n=200 patient seeking or enrolled in treatment for OUD.
- A Letter Of Intent (LOI) to the FDA Center for Drug Evaluation and Research COA Qualification Program was submitted to initiate the qualification process for the OUDSS. **The LOI has been accepted into the CDER COA Qualification Program to pursue qualification as a Drug Development Tool**.

Novel behavioral endpoints in pipeline-continuous measures-cont.

- **3) Change in use frequency: Summary of the new 2-shift endpoint for AUD (Martin -in progress)**
- Given that reduction in WHO drinking risk level has been shown to be a meaningful and useful non-abstinence endpoint among individuals receiving treatment for alcohol use disorder, a similar approach could be useful for establishing a non-abstinence reduction-based clinical trial endpoint among individuals receiving treatment for cocaine use disorder.
- Roos et al. (2019) - Using a pooled dataset of clinical trials for cocaine use (N=716), three cocaine frequency levels were evaluated: abstinence (no cocaine use in the past month), low-frequency use (1-4 days of cocaine use in the past month), and high-frequency use (5 or more days of cocaine use in the past month). These frequency levels at baseline and end of treatment (EOT) were based on the self-reported days of cocaine use during the 28-day period prior to the assessment.

.
.

Novel behavioral endpoints in pipeline-continuous measures- Cocaine Use Disorder

- Results of analyses indicated **those who achieved either a one-level or two-level reduction in cocaine use frequency category from baseline to end-of-treatment had more favorable outcomes during a 12-month follow-up** period as compared to those who showed no change or an increase in cocaine frequency category.
- Those who reduced from high to low-frequency categories had similar outcomes at follow-up as those who reduced to abstinence. Thus, reducing to the 'low frequency' cocaine use category can be a clinically meaningful outcome.
- **These findings parallel those showing categorical reductions in levels of alcohol use (using WHO alcohol risk categories) is linked with substantial clinical benefit following treatment. Reduction in cocaine use frequency may be a practical and useful endpoint for clinical trials**.
-

6. Assessment instruments for novel endpoints: (PROs such as **TLFB** and **EMA** vs. objective measures such as biomarkers)(Martin)

- A major reason why the development of medications for OUD and StUD has been challenging is the difficulty in accurately assessing the substance use of the participants during the clinical trials.
- **The current gold standard for human substance use assessment**, particularly in pivotal trials for drug approval is the well-studied and highly validated **TLFB**, a thirty-year-old method even though more modern, but less validated assessment methods such as EMAs have since been developed.
- The TLFB has demonstrated both reliability and validity when administered as an in-person interview, by telephone, or with a computer-based format.
- Discuss advantages/disadvantages of the different administration methods as well as time durations between assessments. Spanning longer intervals may lead to reduced accuracy in documenting the timing of events, and potentially to underreporting of use (recall bias)

Assessment instruments for novel endpoints – cont.

- More recently, many different designs of EMAs have been developed for and used in clinical trials specifically targeted to the context of the specific studies, making it difficult to validate.
- **High-resolution (HR) vs. low-resolution (LR) designs**:
  - **HR designs**: closely spaced assessments in a short period of time, collecting fine-grained data but placing a higher burden on the participant. Better suited to investigate how substance use relates to variables that frequently change during a given day (i.e., mood, cravings, social and environmental context). Short periods (i.e., 5-7 days) with multiple prompts throughout the day (i.e., 8-10 prompts) / self-initiated reports.
  - **LR designs**: better suited for investigations of day-to-day variability in substance use in combination with other domains.
- **Burst designs**: multiple short bursts (i.e., 2 weeks) of EMA surveys over a longer period as a compromise may lead to reduced accuracy in documenting the timing of events, and potentially to underreporting of use (recall bias)

Assessment instruments for novel endpoints – cont.

- (Maybe discuss some specific EMAs, advantages, disadvantages…)

- Self-report assessments such as TLFB and EMA can affect the behavior of participants depending on how detail-oriented they are and have the potential of becoming an undesired intervention if they are too intrusive.
- While validity is still being established for EMAs relative to the TLFB, recall bias observed across studies showed a trend of underreporting on the TLFB/retrospective reports relative to EMA.
- The validity of EMAs can be influenced by individuals' increased awareness of their substance use patterns, which may result in them catering their reports to the expected outcome that researchers may seek from the study
- ----------
- **Biomarker assessments**, such as biological samples (urine drug screens, blood tests, saliva tests), with the main advantage of being objective, have the disadvantage that they are typically intrusive.
- Despite fast advances in technology and research, biomarkers often still are less precise and less accurate than the established subjective measures when applied as assessments of drug use and still not sensitive enough to detect subtle but clinically significant changes in drug use.
- Discuss tradeoff between accuracy and invasiveness

7.Other potential novel endpoints related to patient functioning- social functional domains **(Recovery Capital)**

- **Recovery capital includes** Physical capital (transportation, employment, housing, and income); social capital (professional support, social activities, friends, and family); cultural capital (rituals, activities, and cultural values); and human capital (skills, abilities, attitudes, knowledge).
    - Examine intrinsic and extrinsic factors in Recovery capital and their interaction; examine the trajectory of recovery capital during treatment and beyond.

8. Statistical approaches. Evaluating several endpoints for establishing treatment effect (Vlad, Martin D)

- **Longitudinal endpoints using within-individual summary statistics**: clinically relevant time-frame (change from baseline), rate of change (slope), area under the curve, mixed effects repeated measurements (MMRM) analysis
- **Multiple primary endpoints: dual vs co-primary**
- **Hierarchical composite endpoints combining several outcomes** of different types (binary, ordinal, continuous, time-to-event) in a prioritized order: win ratio, win-odds, net-benefit as measures of the treatment effect

9. Conclusions

- **Development of novel endpoints and corresponding COAs are urgently needed to counter the US opioid crisis and increasing stimulant drug use.**
- **Clinical drug development in SUD has a constellation of methodological problems that make it daunting and extremely challenging to develop new therapeutics**
  - **historic focus on abstinence,**
  - **dichotomous endpoints lens**
  - **necessity to prove that the proposed proximal in-treatment endpoint and COA correlates with a distal (after treatment in a long-term follow-up) salutary social function or health-related benefit (possible only via meta-analytical approach)**
  - **disconnect between some of the ways to ascertain a signal in early drug development phase (e.g., via biomarkers, neuroimaging) with the later phases of drug development where such technologies are no longer used**
  - **population heterogeneity (DSM-5 problems)**
- **There is an urgency to develop suggestions and ways to make clinical drug development for OUD and StUD more approachable, to de-risk and discharge technical risks, and to find and develop endpoints that are reflective of the real life of the patient suffering from SUD.**
- **This article addresses presents progress, challenges and future steps and limitations in SUD drug development and suggests innovations to overcome endpoint-related difficulties to increase chances for success.**

THANK YOU
DISCUSSION, Q&A,
NEXT STEPS

Novel behavioral endpoints in pipeline-continuous measures-cont. Excerpt from ACTIVE (Alcohol Clinical Trials Initiative) public-private partnership consortium with NIAAA

Table 1.	WHO risk drinking levels (grams of alcohol consumption)
Risk level	Definition of each level, in grams and US standard drinks
Very high	>100 g (>7.1 drinks) for men; >60 g (>4.3 drinks) for women
High	60–100 g (4.3–7.1 drinks) for men; 40–60 g (2.9–4.3 drinks) for women
Moderate	40–60 g (2.9–4.3 drinks) for men; 20–40 g (1.4–2.9 drinks) for women
Low	1–40 g (<2.9 drinks) for men; 1–20 g (<1.4 drinks) for women

The European Medicines Agency (EMA) currently endorses a **2-level reduction in WHO risk drinking levels as one potential outcome** in the regulatory evaluation of new drug applications for AUD pharmacotherapy trials. Based on our work, we propose that the FDA also consider adding a reduction in the WHO level as a primary endpoint to the existing endpoints of total abstinence and the percent of subjects with no heavy drinking days in their guidance offered for regulatory approvals of AUD trials.