A Reduction in the World Health Organization (WHO) Risk Levels of Alcohol Consumption as an Efficacy Outcome in Alcohol Use Disorder (AUD) Clinical Trials

Compiled By:

Alcohol Clinical Trials Infinitive (ACTIVE)

National Institute on Alcohol Abuse and Alcoholism (NIAAA)

U.S. Food & Drug Administration

Critical Path Innovation Meeting

November 09, 2018

1:30 PM EST – 3:00 PM EST

TABLE OF CONTENTS

Meeting Attendees & Contributors 1

Executive Summary 3

Description of the Alcohol Clinical Trials Initiative (ACTIVE) 4

Introduction and Rationale for a Reduction in the Who Health Organization (WHO) Risk Levels of Alcohol Consumption as an Efficacy Outcome in Alcohol Use Disorder (AUD) Clinical Trials 5

Objective of the Meeting 7

The Measurement of Alcohol Consumption 8

Rationale for Evaluating Epidemiological Data in Relationship of Drinking Reduction to Morbidity and Mortality 10

WHO Risk Drinking Levels in a Nationally Representative General Population Sample – Relationship to Clinical Diagnoses and Function 11

Effect of WHO Risk Drinking Level Change on Various Feel and Function Measures 14

WHO Risk Levels Reductions in an AUD Clinical Trial Sample – Relationship to Behavioral and Physical Function 21

Reductions in WHO Risk Drinking Levels and Health Care Costs in an Alcohol Use Disorder Clinical Trial Sample 29

Reductions in WHO Risk Drinking Levels as Outcomes in Alcohol Pharmacotherapy Trials 33

Stability of WHO Risk Drinking Reduction Outcomes in a 6-Month Clinical Trial 37

Review of Important Points and Observations 38

References 39

FDA
Critical Path Innovation Meetings (CPIM) Attendees

Raymond F. Anton, M.D. | Thurmond Endowed Chair - Distinguished Professor of Psychiatry and Behavioral Science
Chair of ACTIVE
Scientific Director- Alcohol Research Center
Medical University of South Carolina
67 President Street, MSC 861
Charleston, South Carolina 29425
843-792-1226
antonr@musc.edu

Arnie Aldridge | Research Economist
ACTIVE Member
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709
919-990-8389
aaldridge@rti.org

Daniel Falk, Ph.D. | Health Scientist Administrator
ACTIVE Member
Division of Medications Development
National Institute on Alcohol Abuse and Alcoholism
National Institutes of Health
6700B Rockledge Dr, Room 1427
Bethesda, MD 20892-6902
301-443-0788
falkde@mail.nih.gov

Deborah S. Hasin, Ph.D. | Professor of Epidemiology
ACTIVE Member
Columbia University
1051 Riverside Drive #123
New York, New York 10032
646-774-7909
dsh2@cumc.columbia.edu
deborah.hasin@gmail.com

Raye Z. Litten, Ph.D. | Acting Director
ACTIVE Member
Division of Medications Development
Division of Treatment and Recovery Research
National Institute on Alcohol Abuse and Alcoholism
National Institutes of Health
6700B Rockledge Drive, Room 1320, MSC 6902
Bethesda, Maryland 20892-6902
301-443-0636
rlitten@mail.nih.gov

Stephanie S. O'Malley, Ph.D. | **Professor and Deputy Chair of Psychiatry**
ACTIVE Member
Yale School of Medicine
34 Park Street
New Haven, CT 06517
203-974-7590
stephanie.omalley@yale.edu

Katie Witkiewitz, Ph.D. | Professor
ACTIVE Member
Department of Psychology
Center on Alcoholism, Substance Abuse, and Addictions
MSC03-2220
University of New Mexico
Albuquerque NM 87131
505-277-5953
katiew@unm.edu

Patricia A. Powell, Ph.D. | Deputy Director
Observer
National Institute on Alcohol Abuse and Alcoholism
National Institutes of Health
6700B Rockledge Drive, Room 1207, MSC 6902
Bethesda, MD 20892-6902
301-443-5106
ppowell@mail.nih.gov

Executive Summary

For many individuals, AUD is a chronic relapsing disease where maintaining abstinence is often difficult. However, significant improvement in drinking, short of full abstinence, can have many health, social, and economic benefits. Analyses done by a Public-Private Partnership workgroup, the Alcohol Clinical Trials Initiative (ACTIVE), herein demonstrate that a reduction in a relatively new metric of alcohol harm reduction, the World Health Organization (WHO) risk drinking level, has validity in real-world settings and could be a viable AUD clinical trial endpoint. For example, we show that individuals drinking at WHO very high-risk levels, who subsequently reduce their WHO risk drinking level, report clinically significant improvements in how they feel and function. These improvements include a lower risk for alcohol dependence, less severe alcohol-related consequences, reduction in health care costs, improved mental health, improved quality of life, improved liver function and lower systolic blood pressure. In addition, the WHO risk drinking level endpoint is as sensitive or, in some instances, more sensitive than the standard FDA guided "total abstinence" and "percent of subjects with no heavy drinking" endpoints in detecting differences between an experimental medication and placebo in alcohol treatment clinical trials. Moreover, this endpoint identifies more individuals as improved compared to "full abstinence" or "percent subjects with no heavy drinking". Finally, the WHO risk drinking reduction endpoint appears to maintain its efficacy in predicting improved function after extended periods of time (at least 1 year for function and up to 3 years for health costs). Furthermore, this outcome displays considerable stability by the third or fourth month of clinical trial participation, suggesting that a clinical trial length of 3-4 months might be an adequate duration for AUD phase three clinical trials using this outcome as a primary endpoint.

International publicly-available data suggest that reducing drinking from very high levels to more moderate levels can reduce long-term mortality and overall disease burden. Also, the use of the same outcome by regulatory agencies in the United States and Europe, where a 2-level reduction in WHO risk drinking is accepted, would harmonize regulatory requirements and provide efficiency in medication development for this undertreated disorder. In addition, the WHO risk drinking level outcomes capture reductions in drinking, the preferred goal of most patients, and are more readily achieved as a measure of success than "abstinence" or "no heavy drinking days". Substantial reductions in drinking, if agreed upon as a suitable goal of alcohol treatment, could increase the desirability and acceptability of treatment to patients and caregivers, and enhance the drug development process by providing additional outcomes for clinical trials of this underserved and costly disorder.

In sum, it is felt that data provided by ACTIVE in conjunction with the National Institute of Alcohol Abuse and Alcoholism, as detailed in this report, support the use of the reduction in WHO risk drinking levels as a primary-efficacy endpoint for phase-3 AUD pharmacotherapy clinical trials in the United States. It is with great interest that we request FDA discussion and guidance on this proposal.

Description and Purpose of the Alcohol Clinical Trials Initiative (ACTIVE) (2009-2018)

ACTIVE was formed nearly 10 years ago at the request of the American College of Neuropsychopharmacology (ACNP), one of the most prestigious societies focusing on the biological causes and cures for psychiatric brain disease. Despite great clinical and public health need, the ACNP recognized that few drugs were submitted and/or approved by regulatory authorities for the treatment of alcohol and other substance use disorders. As such, after an initial planning meeting (attended by the Directors of the National Institute on Alcohol Abuse and Alcoholism and the National Institute on Drug Abuse (NIAAA), representatives from the FDA, academics, and 23 Pharmaceutical Companies), the decision was made to focus initially on Alcohol Use Disorder (AUD). From its start, ACTIVE was envisioned as a Public-Private partnership, funded by the Pharma Industry but receiving in-kind support from NIH (e.g. for ancillary data analysis) and FDA (by attendance). Recently, ACTIVE applied for and received "Public-Private Partnership" approval from the FDA and has complied with the required public disclosure guidelines.

ACTIVE has held 2 meetings per year focusing on the improvement and standardization of AUD clinical trial methodology. The goals, mission, and questions to be addressed by ACTIVE were laid out in its initial publication (Anton et al., 2012). The overarching goal has been to inform AUD clinical trials design, analysis, and outcomes that would not only meet regulatory approval, but would also increase clinical trial participation, and ultimately treatment to a wider spectrum of AUD affected individuals.

To this end, ACTIVE has sought to better understand the efficacy, and clinical relevance, of various drinking outcomes (i.e., endpoints) for AUD trials. ACTIVE's current focus has been to develop a new drinking reduction outcome for AUD trials that reflects what has been observed by clinicians and researchers to date—that the vast majority of individuals with AUD would like to reduce their drinking and that reductions in drinking, short of abstinence, can result in improvements in feeling and functions. Such a reduction-oriented drinking outcome, if appropriately validated, could significantly improve both patient satisfaction, treatment seeking, and ultimately public health (see the section below for the rationale underlying this focus).

ACTIVE has requested a meeting with the FDA-CPIM to share data analysis that supports this concept of a drinking reduction and request guidance and consideration as to whether a reduction in the World Health Organization (WHO) risk drinking levels of alcohol consumption could be used as a primary efficacy endpoint in Phase III clinical trials for this underserved and costly disorder.

Introduction and Rationale for a Reduction in the World Health Organization (WHO) Risk Drinking Levels as an Efficacy Outcome in Alcohol Use Disorder (AUD) Clinical Trials

Alcohol use disorder (AUD), previously called alcohol dependence, is a devastating brain disease that causes a myriad of medical, psychological, social, and economic problems, affecting over 15 million adults and over 620,000 adolescents in the United States (U.S.) (https://pubs.niaaa.nih.gov/publications/AlcoholFacts&Stats/AlcoholFacts&Stats.htm). It estimated that over 88,000 Americans die from alcohol-related causes annually, costing society over $249 billion a year in medical, economic, and social costs. Moreover, 3.3 million deaths, or 5.9 percent of all global deaths annually are attributable to alcohol consumption, one of the highest ranked causes of death. Globally, alcohol misuse is the fifth leading risk factor for premature death and disability. Heavy alcohol use is known to increase heart disease, stroke, cancer, gastrointestinal problems, and is the leading cause of liver transplantation. In the U.S. and elsewhere, heavy alcohol consumption contributes to many social problems such as accidental injury, domestic violence, rape, assaults, and murders, while more than 10 percent of children live with a parent with identifiable alcohol problems. Despite this immense burden of disease and associated social costs, and despite the advanced neuroscience identification of pharmacological targets, there are a limited number of FDA approved medications to treat AUD.

Currently, only 3 medications have been approved by the FDA for "alcohol dependence": disulfiram, oral and injectable long-acting naltrexone, and acamprosate (Litten et al., 2016), and as seen in the table below the last one being over 12 years ago.

COMMERCIAL NAME	GENERIC NAME	DATE OF FDA APPROVAL
ANTABUSE	Disulfiram	08/28/1951
REVIA	Naltrexone, oral	11/20/1984
Campral	Acamprosate	07/29/2004
VIVITROL	Naltrexone for extended-release injectable suspension	04/13/2006

However, partially because of the heterogeneity of AUD, these medications are not universally effective, especially when applying the conservative "success criteria" of "total abstinence" or "no heavy drinking days" currently recommended by the FDA. Since developing new and more effective medications is a priority for the National Institute on Alcohol Abuse and Alcoholism (NIAAA), in conjunction with the ACTIVE workgroup, it has engaged in an effort to evaluate better ways to evaluate the effectiveness of novel medications, including evaluating methods to measure and validate a new efficacy outcome of drinking harm reduction, a reduction in the WHO risk levels of alcohol consumption.

To date, the FDA recommends 2 primary dichotomous outcomes for defining a successful response to treatment in Phase 3 alcohol pharmacotherapy trials: a) total abstinence and b) the percent of subjects with no heavy drinking days (where a heavy drinking day is defined as 5 or more drinks for men and 4 or more drinks for women) (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM433618.pdf). These 2 measures are excellent at capturing incidences of abstinence and low risk-drinking; however, they fail to capture large numbers of individuals with AUD who might achieve significant clinical benefit from lesser, but potentially clinically meaningful, reductions in drinking.

Currently, less than 10 percent of individuals with AUD seek treatment for their drinking problems (Grant et al., 2015), and the majority of individuals with AUD who do not seek alcohol treatment report not wanting

to stop drinking as the primary reason for not seeking treatment (Figure 25 in https://www.samhsa.gov/data/sites/default/files/NSDUH-DR-FRR3-2014/NSDUH-DR-FRR3-2014/NSDUH-DR-FRR3-2014.htm). Many individuals with AUD who do not initially accept an abstinence goal may find treatment more appealing if it is focused on reductions in drinking (Probst et al., 2015; van Amsterdam and van den Brink, 2013). This is partially based on the perception/anticipation that a reduction in drinking will lead to a reduction in alcohol-induced harm. Recent evidence suggests that reduction in drinking is an important goal for many individuals with AUD. In 3 recent AUD multisite pharmacotherapy clinical trials, individuals seeking participation who had a non-abstinence goal of drinking reduction ranged from 72 to 91 percent (Falk et al., 2018; Litten et al, 2013; Ryan et al., 2017). Additionally, many individuals with AUD report fear of the "stigma" associated with being identified as having AUD and are therefore reluctant to stop drinking completely (Probst et al., 2015) since abstinence is often interpreted by others has having an alcohol problem. Moreover, AUD is a chronic relapsing disorder for many people, whereby focusing on the reduction of harmful drinking can be an important aspect of disease management (Maremmani et al., 2015) similar to the management of other chronic diseases like diabetes, obesity, and hypertension where objective measures of reduced-harm (e.g. lowered blood sugar, weight, and blood pressure) are acceptable goals. Indeed, there are numerous benefits to reductions in drinking, including decreases in morbidity and mortality (Laramee et al., 2015), lower healthcare costs (Kline-Simon et al., 2014), decreased alcohol-related consequences (Falk et al., 2010; Witkiewitz, 2013), and improved psychosocial functioning (Kline-Simon et al., 2013; Witkiewitz, 2013).

Objective of the Meeting

Given the perceived utility and need for an alcohol consumption reduction measure for clinical trials, the objective of this meeting is to discuss a new endpoint for alcohol medications development that we believe successfully identifies individuals with AUD whose reductions in drinking are associated with clinically meaningful improvements in how they feel and function. Through careful research, NIAAA in conjunction with ACTIVE has conducted a series of analyses to validate the World Health Organization (WHO) risk levels of alcohol consumption as a clinically meaningful indicator of drinking reduction for AUD clinical trials. The WHO risk drinking levels include very high-risk drinking, high-risk drinking, moderate-risk drinking, and low-risk drinking based on grams of ethanol consumed per day (see below).

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.

In the following sections, we will demonstrate and discuss the following topics: 1) how drinking is measured; 2) how reductions in the WHO risk drinking level results in improvements in how individuals with AUD feel and function; 3) how reductions in WHO risk drinking level translates to long-term reduction in health care costs; 4) how this drinking endpoint performs and compares to the current primary drinking endpoints in several prominent AUD pharmacotherapy trials; and 5) the degree to which this drinking endpoint remains stable during AUD pharmacotherapy trials.

We welcome the opportunity to discuss these data as well as the advantages of using the reduction in WHO risk drinking levels as one of the primary endpoints for evaluating candidate medications in pivotal alcohol clinical trials.

The Measurement of Alcohol Consumption

To place the data presented in this report in perspective, it is necessary to understand how and why alcohol consumption (i.e. drinking) is measured in clinical trials. Alcohol use disorder (AUD) is a chronic disorder, often with varying periods of recovery interspersed with relapses to drinking. Given the complex pathogenesis and time course of AUD, the main aim of alcohol treatment should be to improve how a patient "feels and functions". One way of assessing the efficacy of a treatment is to directly measure its effects on physical or psychosocial consequences. However, experience has shown that this often is impractical, necessitating trials that are too long and large (FDA, 2015). Thus, rather than demonstrating a direct effect on these measures in AUD clinical trials, the FDA has allowed the use of "alcohol consumption" itself as a suitable "surrogate/estimate" for assessing how a patient feels and functions, provided "the pattern of [drinking] behavior can be reasonably predictive of clinical benefit…" (FDA, 2015, line 68).

Over the past several decades, alcohol consumption typically has been measured using 2 methods: the calendar-based estimate of daily drinking, the Timeline Follow-Back (TLFB), and the more global quantify/frequency (Q/F) alcohol consumption measure. In alcohol clinical trials, the TLFB is considered the gold standard by alcohol researchers and is recommended by the FDA.

The TLFB is a semi-structured interview that uses a calendar prompt and other memory aids (e.g., holidays, payday, and other personally relevant dates) to facilitate accurate recall of the number of drinks consumed each day during a given target period (e.g.,7- 28 days) (Sobell et al., 1996). The TLFB has demonstrated good validity and reliability in many research settings (Carey, 1997; Sobell, et al., 1988; Sobell, et al., 1996) for periods up to 90 days. In alcohol clinical trials, the TLFB is administered at baseline (e.g. with a 30- to 90-day recall period) and then multiple times during the treatment period (often weekly, biweekly, or monthly, depending on the length of the trial). A variety of drinking endpoints can be derived from the daily number of drinks captured by the TLFB; these include the FDA-recommended endpoints (percentage of subjects abstinent and the percentage of subjects with no heavy drinking days), as well as the new World Health Organization (WHO) risk drinking endpoints, among others (e.g. percent drinking days, percent heavy drinking days, drinks per day, and drinks per drinking day).

The Q/F measure is used commonly in epidemiological studies and survey research. The measure queries the number of drinks consumed on a typical drinking day (quantity) and the typical number of drinking days (frequency) over a given period (e.g., a week, month, or year). Studies in the general population have shown that measuring past-year drinking by Q/F correlate well with a modified TLFB conducted by mail over the same time period (Sobell et al., 2003). In clinical trial data, the correlation between the Q/F and TLFB (using a 30-day recall period) was less robust. Q/F consistently underestimated the quantity of alcohol consumed when compared with the TLFB (NIAAA, unpublished data). Although the Q/F measure is easier to administer and requires less time than the TLFB, the TLFB is considered the more accurate assessment instrument for clinical trials.

Finally, the amount of alcohol, measured by the TLFB and Q/F, is often expressed as the number of standard drinks consumed. However, the amount of alcohol in a standard drink varies among the different countries, ranging from 8 grams in the UK to 14 grams in the U.S. (Hasin et al., 2017). An important advantage of the WHO risk drinking categories, which measure the amount of drinking in grams, is that information provided can be translated readily into standard drink equivalents per day across countries that use different gram to standard drinks conversions. This utility should provide a universal/global metric, allowing easier cross-border interpretation and comparisons of clinical trial results and medication efficacy.

While perhaps less germane for the current discussion but alluded to in subsequent data analyses, biological estimates of alcohol consumption can be made using laboratory tests performed on urine and blood. A few sensitive and reliable lab tests/biomarkers of alcohol consumption are available (Litten et.

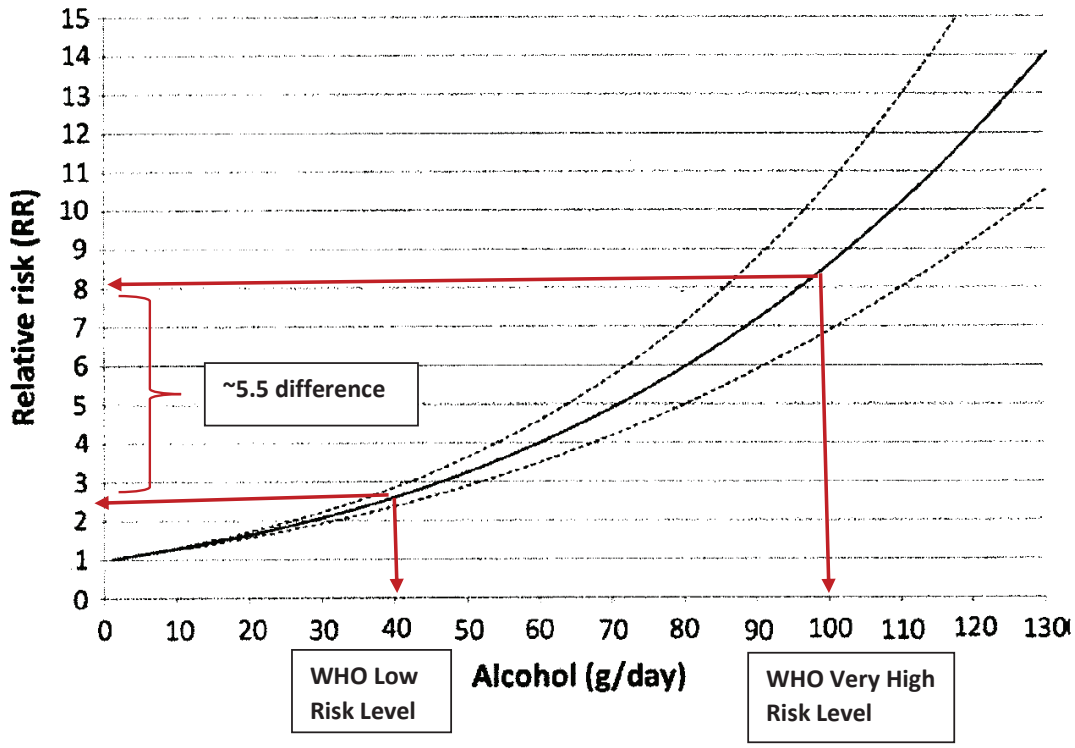
al. 2010). These range from measurement of low amounts of acute consumption (measured by urinary Ethylglucuronide -EtG), moderate amounts of almost daily consumption (measure by red blood cell Phosphatidylethanol – PETh) and heavy/harmful alcohol consumption (measured by serum Carbohydrate Deficient Transferrin – CDT or gamma glutamyl transferase -GGT). It is possible, and even likely, that the combined use of these lab tests could enhance the accuracy of verbal reporting, but these tests have not been used widely in AUD clinical trials an may be more vulnerable to missing data compared to the TLFB.

Rationale for Evaluating Epidemiological Data in Relationship of Drinking Reduction to Morbidity and Mortality

Epidemiological evidence shows positive linear and curvilinear relationships between alcohol consumption (gm/day) and mortality and various chronic diseases, including cancers of the mouth, oropharynx, esophagus, larynx, colon, liver, and breast; liver cirrhosis; pancreatitis; hemorrhagic stroke; hypertension; and epilepsy (Rehm et al., 2017; Shield et al., 2013; Tramacere et al., 2010). Wood et al. (2018) recently published the results of 83 prospective studies in 599,912 current drinkers in *The Lancet* showing a relationship between alcohol intake (drinks/week) and mortality, as well as cardiovascular disease, stroke, and heart failure.

Although these prospective and retrospective studies (Rehm et al., 2017; Shield et al., 2013; Wood et al., 2018) have potential limitations, current guidelines on drinking are based, in part, on these types of studies. With regard to the WHO risk drinking categories, it is readily observable that, in many cases, moving from the WHO "very high/high risk" level to a "medium/low risk" level is associated with a decreased risk of mortality and chronic disease from alcohol consumption. For example, in the dose-response curve by Tramacere et al (2010) (see Figure below), going from alcohol consumption corresponding to WHO "very high" risk drinking level (100+ gm/day for men, for example) to a "low" risk drinking level (<40 gm/day) corresponds to difference in relative risk reduction of about 5.5 times for oral and pharyngeal cancers.

Figure 1. Reduction of oral and pharyngeal cancer risk in individuals reducing alcohol consumption on average from greater than 100 grams/day to 40 grams per day.



Given the wealth of epidemiological data supporting the link between reduction of alcohol consumption and improved longevity and health, it made sense to evaluate available epidemiological data in the United States to evaluate the link between reductions in WHO risk drinking levels and changes in function in the general population.

WHO Risk Drinking Levels in a Nationally Representative General Population Sample - Relationship to Clinical Diagnoses and Function

Over the last 3 years, the ACTIVE group has been working on examining whether reductions in the WHO risk drinking levels are associated with benefits to drinkers in the general population, i.e., improvements in how they feel and/or function. The ACTIVE group decided to examine these issues in general population data for 3 main reasons: (1) Most clinical trial datasets had relatively small sample sizes, and lacked follow-up periods of sufficient duration to produce meaningful results. (2) Patients in clinical trials are not likely to be representative of all individuals who are heavy drinkers. Participants in a large, national study of adult household residents would be more representative. (3) Earlier work on the meaning and implications of reduction in WHO risk drinking levels emphasized the importance of examining reduction in WHO levels by initial drinking level, as a 1-level reduction could mean something entirely different for an initially very heavy drinker than for a moderate or light drinker (Rehm J, Roerecke M. Alcohol Alcohol 2013; 48: 509–13). For this reason, a large sample that included drinkers across varying levels of the WHO risk drinking levels was needed. Such data were found in a large, nationally representative survey of US adults who participated in a 3-year follow-up interview (Hasin et al., Lancet Psychiatry 2017).

Methods

The survey was the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) in which face-to-face interviews were conducted in participants' homes. The NESARC target population consisted of adults age at least 18 years living in households and group quarters. The baseline and follow-up interviews are referred to as Wave 1 and 2, respectively. Wave 1 was conducted in 2001-2002, and Wave 2 in 2004-2005. Wave 1 included 43,093 participants; Wave 2 re-interviewed 34,653 of them, a follow-up response rate of 86.7% after excluding deceased or incapacitated individuals. Of the individuals re-interviewed at Wave 2, 22,005 were drinkers at Wave 1 (at least 1 drink in the prior 12 months). These constituted the analytic sample for the study the ACTIVE group undertook of WHO risk drinking level reduction.

NESARC measures of WHO risk drinking levels. The baseline and follow-up interviews included state-of-the-art detailed measures of drinking, including quantity and frequency of drinking by types of alcohol (beer, wine, liquor). These measures were used to calculate the WHO risk drinking levels and evaluate change in these levels between baseline and follow-up interviews (Dawson et al., 2005). Drinking was measured in grams of ethanol and then converted into drinks per day. The 4 WHO risk levels are shown in Table 1 (above). As shown, the levels are defined separately for men and women.

NESARC outcome measures. The surveys also included measures of several conditions that were used as outcomes for the study. These all indicated aspects of how participants were feeling and/or functioning at Wave 2. They included the following:

1. DSM-IV Alcohol Dependence. This was diagnosed using the criteria in the Diagnostic and Statistical Manual, 4th edition (DSM-IV). To receive the diagnosis, 3 out of the 7 DSM-IV criteria were required. The DSM-IV alcohol dependence measure used in the NESARC has been shown to be highly reliable in clinical and general population samples in the US and internationally (Grant et al., 1995; 2003; Ruan et al., 2008; Hasin et al., 1997a; 1997b; Cottler et al., 1997).

2. Functioning. This was measured on a widely-used functioning scale, the Short-Form 12, 2nd version (SF12v2; Ware et al., 2002). We dichotomized this variable to represent impairment among participants scoring ≥1 SD below the population mean (Hasin et al., 2017).

3. AUDIT-C screening scale. Developed initially for the World Health Organization (Saunders et al., 1993), this screening scale for risky drinking is validated (Dawson et al., 2005; 2012) and widely used in clinical settings, e.g., in yearly screenings across all VA patients with primary care visits (Bradley et

al., 2007). AUDIT-C scores were dichotomized at the breakpoint commonly accepted as indicating increased risk (Knox et al., in press).

4. Liver cirrhosis/other liver disease. This is self-reported in the medical conditions' module of the interview. For this study, liver disease was counted as positive only if a doctor made the diagnosis (Knox et al., in press).

5. DSM-IV-defined common psychiatric disorders. In this study, common psychiatric disorders were included that are associated with heavy drinking and are indicators of how individuals feel and function. These included depressive disorders (major depression, dysthymia), and anxiety disorders (panic, generalized anxiety, social or specific phobia). These disorders were also diagnosed using DSM-IV criteria. Given the strong comorbidity between depressive and anxiety disorders that we and others have found (e.g., Hasin et al., 2004; 2018) we created a binary variable for the study indicating that any of these disorders were present (Knox et al., under review).

6. Other DSM-IV substance use disorders. These were abuse or dependence diagnoses for sedatives/tranquilizers, painkillers, marijuana, cocaine/crack, stimulants, club drugs, hallucinogens, inhalants and heroin. These were also diagnosed using DSM-IV criteria (Knox et al., in preparation).

Analysis. The method of analysis was the same for all 6 outcomes. Logistic regression was used to test the relationship between reductions in WHO risk drinking level between Waves 1 and 2 with the status of each of the 6 outcomes at Wave 2. Within each overall model, reductions were examined by initial WHO risk drinking level at Wave 1. Participants who were very high-risk drinkers at Wave 1 could show 3 levels of reduction in drinking: 1 level (to high risk), 2 levels (to moderate risk) and 3 levels (to low risk). High-risk drinkers could reduce 1 or 2 levels (to moderate or low risk). Table 2 shows the prevalence of changes in WHO risk drinking levels at Wave 2 for Wave 1 very high and high-risk drinkers. Change in risk for the outcome by change in drinking level was indicated with adjusted odds ratios (aOR). aOR indicate the effect of the predictor variable (change in WHO risk drinking level) on the outcome (e.g., Wave 2 alcohol dependence), relative to the comparison group consisting of participants whose WHO risk drinking level did not change between Waves 1 and 2, adjusted for covariates. aOR >1.00 indicate increased odds relative to the comparison group, aOR=1.00 indicate no difference in risk, and aOR <1.00 indicate decreased risk relative to the comparison group. Statistical significance of an aOR is indicated by a 95% confidence interval that does not include 1.00 between its lower and upper limit. All models controlled for sex, age, education, race/ethnicity, smoking, body-mass index, health insurance, and Wave 1 status of the outcome. While all models included the participants in all 4 WHO risk drinking categories at Wave 1, in interpreting the results, in this report we focus on the very high-risk and high-risk drinkers, since they are the ones most like patients entering AUD pharmacotherapy clinical trials and thus of most interest for evaluating reduction in WHO risk drinking levels and improvements in how a patient may feel and/or function.

Table 2. Overall prevalence of WHO risk drinking level change observed at Wave 2 for Wave 1 very high and high-risk drinkers

Wave 1 WHO risk level	Wave 2 change in WHO level	Prevalence (%)
Very High (n=512)		
	No change	26
	Decreased 1 level	14
	Decreased 2 levels	11
	Decreased 3 levels	41
	Became abstainer	7
High (n=546) *		
	No change	15
	Decreased 1 level	19
	Decreased 2 levels	46
	Became abstainer	6
* numbers do not total 100% because 14% had increased to the very high-risk level by Wave 2		

Results

Effect of Who Risk Drinking Level Change on Various Feeling and Function Measures

Alcohol dependence. Among very-high-risk drinkers at Wave 1, each decrease in WHO risk drinking level was associated with significantly lower prevalence and adjusted odds of alcohol dependence at Wave 2, compared to respondents with no change in drinking level (the reference group; Table 3). Specifically, in individuals whose drinking remained at the very high-risk level, 37 % were alcohol dependent at Wave 2, whereas among those who decreased their drinking by 1, 2, or 3 WHO risk levels, 14%, 9%, and 4% were alcohol dependent, respectively. The risk of alcohol dependence among abstainers was zero. Similarly, among high-risk drinkers at Wave 1, each decrease in risk drinking level by Wave 2 was associated with significantly lower prevalence and adjusted odds of Wave 2 alcohol dependence. An additional set of analyses among the subset of very-high-risk drinkers who were alcohol dependent at Wave 1 (not shown) also indicated that each decrease in WHO risk level by Wave 2 was associated with significantly lower odds of alcohol dependence at Wave 2, indicating that reduction in WHO risk level not only reduced the prevalence but also the persistence of alcohol dependence by Wave 2 (Hasin et al., 2017).

Table 3. Risk for DSM-IV alcohol dependence at Wave 2 by Wave 1 WHO risk drinking level and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in WHO level	% with Wave 2 alcohol dependence	aOR (95% CI)
Very High (n=512)			
	No change	37	Reference
	Decreased 1 level	14	**0.27 (0.18-0.41)**
	Decreased 2 levels	9	**0.17 (0.10-0.27)**
	Decreased 3 levels	4	**0.07 (0.05-0.10)**
	Became abstainer	0	---
High (n=546)			
	No change	27	Reference
	Decreased 1 level	19	**0.67 (0.56-0.80)**
	Decreased 2 levels	4	**0.12 (0.09-0.15)**
	Became abstainer	0	---

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

SF-12 functioning. Results were clearest for participants whose WHO risk drinking level was very high at Wave 1 (Table 4). Among all very high-risk drinkers at Wave 1, reductions in WHO risk drinking levels predicted significantly lower risk of poor functioning at Wave 2 compared to respondents with no change in drinking level (the reference group). Additional analyses (not shown) indicated that this was also the case for the subset of Wave 1 very high-risk participants who were also alcohol dependent.

Table 4. Risk for impaired functioning (at least 1 SD below population mean on SF-12v2) at Wave 2 by Wave 1 WHO risk drinking level and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in risk level	% with Wave 2 impaired functioning	aOR (95% CI)
Very High (n=512)			
	No change	22	Reference
	Decreased 1 level	8	**0.34 (0.25-0.45)**
	Decreased 2 levels	12	**0.51 (0.38-0.67)**
	Decreased 3 levels	12	**0.51 (0.35-0.75)**
	Became abstainer	23	1.10 (0.73-1.67)
High (n=546)			
	No change	18	Reference
	Decreased 1 level	11	**0.56 (0.43-0.74)**
	Decreased 2 levels	19	1.09 (0.84-1.41)
	Became abstainer	14	0.76 (0.52-1.10)

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

Positive AUDIT-C screening test. Among Wave 1 very high-risk drinkers (Table 5), compared to those with no change, each decrease in WHO risk level predicted significantly lower prevalence and adjusted odds of a Wave 2 positive AUDIT-C. Additional analyses (not shown) indicated that among those who were very high-risk drinkers at Wave 1 *and* positive on the AUDIT-C, each decrease in WHO risk drinking level by Wave 2 predicted significantly lower risk of a persistent Wave 2 positive AUDIT-C score. Among those who were high-risk drinkers at Wave 1, all reductions in WHO risk levels were associated with significant decreases in risk of a positive Wave 2 AUDIT-C score, compared to respondents with no change in drinking level (the reference group). Additional sensitivity analyses (not shown) that additionally controlled for Wave 1 alcohol dependence were very similar to the main results (Knox et al., in press).

Table 5. Positive AUDIT-C score prevalence at Wave 2 by Wave 1 WHO risk drinking level (drinks per drinking day) and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in risk level	% with Wave 2 positive AUDIT-C scores	aOR (95% CI)
Very High (n=512)			
	No change	93	Reference
	Decreased 1 level	78	**0.27 (0.20-0.36)**
	Decreased 2 levels	55	**0.09 (0.07-0.12)**
	Decreased 3 levels	27	**0.03 (0.02-0.04)**
	Became abstainer	0	---
High (n=546)			
	No change	68	Reference
	Decreased 1 level	56	**0.61 (0.54-0.69)**
	Decreased 2 levels	35	**0.25 (0.23-0.29)**
	Became abstainer	0	---

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

Liver Disease. As shown in Table 6, among Wave 1 very high-risk drinkers, each decrease in WHO risk level predicted significantly lower adjusted odds of Wave 2 liver disease, compared to respondents with no change in drinking level (the reference group), although a decrease to complete abstention was predictive of increased risk for liver disease. This latter finding may be due to the "sick quitter" phenomenon, i.e., those who stop drinking entirely due to illness. Among Wave 1 high-risk drinkers, reductions in risk for liver disease by reduction in WHO risk drinking levels were not significant (Knox et al., in press).

Table 6. Liver disease prevalence at Wave 2 by Wave 1 WHO risk drinking level (drinks per drinking day) and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in risk level	% with Wave 2 liver disease	aOR (95% CI)
Very High (n=512)			
	No change	2	Reference
	Decreased 1 level	0.5	**0.34 (0.21-0.54)**
	Decreased 2 levels	0.4	**0.23 (0.15-0.36)**
	Decreased 3 levels	0.3	**0.17 (0.10-0.29)**
	Became abstainer	3	**2.03 (1.18-3.51)**
High (n=546)			
	No change	0.8	Reference
	Decreased 1 level	0.6	0.71 (0.36-1.41)
	Decreased 2 levels	0.5	0.63 (0.34-1.14)
	Became abstainer	0.1	**0.16 (0.05-0.54)**

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

Common psychiatric disorders. As shown in Table 7, among Wave 1 very high-risk drinkers, each non-abstinent decrease in WHO risk level predicted significantly lower odds of Wave 2 DSM-IV depression and/or anxiety disorders compared to respondents with no change in drinking level (the reference group). Results for high-risk drinkers were less consistent, with a decrease of 1 level in WHO risk drinking curiously predicting an increased risk for depression and/or anxiety, while a decrease of 2 levels decreased risk (Knox et al., under review).

Table 7. DSM-IV depression and/or anxiety disorders at Wave 2 by Wave 1 WHO risk drinking level (drinks per day) and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in risk level	% with Wave 2 depressive and/or anxiety disorder	aOR (95% CI)
Very High (n=512)			
	No change	29	Reference
	Decreased 1 level	15	**0.42 (0.29-0.63)**
	Decreased 2 levels	13	**0.37 (0.22-0.62)**
	Decreased 3 levels	22	**0.67 (0.49-0.93)**
	Became abstainer	19	**0.56 (0.37-0.83)**
High (n=546)			
	No change	16	Reference
	Decreased 1 level	20	**1.32 (1.03-1.69)**
	Decreased 2 levels	12	**0.70 (0.51-0.96)**
	Became abstainer	7	**0.40 (0.28-0.57)**

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

Other substance use disorders. As shown in Table 8, among Wave 1 very high-risk drinkers, each non-abstinent decrease in WHO risk level predicted significantly lower odds of Wave 2 SUD compared to respondents with no change in drinking level (the reference group). Results were less consistent for those who were high-risk drinkers at Wave 1. Among this group, a 1-level decrease in WHO risk level predicted significantly higher odds of Wave 2 SUD compared to those whose WHO risk level remained unchanged, while a 2-level decrease predicted a significantly lower risk of SUD at Wave 2. Sensitivity analyses (not shown) indicated that controlling for alcohol dependence at Wave 1 did not affect the findings, nor did limiting the sample to the subset of participants who had alcohol dependence at Wave 1.

Table 8. DSM-IV substance use disorders at Wave 2 by Wave 1 WHO risk drinking level (drinks per day) and change in level by Wave 2

Wave 1 WHO risk level	Wave 2 change in risk level	% with Wave 2 substance use disorder	aOR (95% CI)
Very High (n=512)			
	No change	3	Reference
	Decreased 1 level	0.4	**0.15 (0.09-0.23)**
	Decreased 2 levels	0	**0.01 (0.01-0.02)**
	Decreased 3 levels	0.7	**0.24 (0.12-0.50)**
	Became abstainer	0	---
High (n=546)			
	No change	0.5	Reference
	Decreased 1 level	0.7	**1.37 (1.10-1.69)**
	Decreased 2 levels	0.6	1.04 (0.81-1.33)
	Became abstainer	0	---

Note: An aOR = 1.0 would indicate that the decrease in WHO risk drinking level did not affect the risk for the outcome at Wave 2. An aOR < 1.0 indicates that the decrease in WHO risk drinking level decreased the risk for the outcome. An aOR > 1.0 would indicate that the decrease in WHO risk drinking level increased the risk for the outcome. All aORs with CI ranges that do not include 1.00 are statistically significant. All significant aORs are in **bold**.

Summary and Conclusions

As summarized in Table 9, results for the "WHO very high-risk drinkers" there were significant improvement of function in all 6 categories examined in those who had either a WHO risk drinking reduction of 1 or 2 levels. For "WHO high-risk drinkers" a benefit in function is evident for 3 out of our 6 outcomes is observed (depending on level of reduction), including the prevalence and persistence of DSM-IV alcohol dependence, the prevalence and persistence of impaired functioning, and a positive score on the AUDIT-C a screening measure that is widely used in primary care settings. Decreased prevalence of liver disease was also predicted by 1- or 2-level reduction in WHO risk drinking levels by Wave 2 – however, since liver disease was a very rare outcome in this general population sample, and the reductions in odds were only statistically significant in the WHO very high-risk drinking group.

Table 9. Summary of results for non-abstinent reductions in WHO risk drinking levels by 3-year follow-up

	Wave 1 Very high risk drinkers (n=512)			**Wave 1 High risk drinkers (n=546)**	
Outcomes	**Decreased 1 level aOR (95% CI)**	**Decreased 2 levels aOR (95% CI)**	**Decreased 3 levels aOR (95% CI)**	**Decreased 1 level aOR (95% CI)**	**Decreased 2 levels aOR (95% CI)**
DSM-IV alcohol dependence	**0.28 (0.19-0.42)**	**0.17 (0.11-0.28)**	**0.07 (0.05-0.10)**	**0.67 (0.56-0.80)**	**0.12 (0.09-0.15)**
SF-12v2 functioning	**0.34 (0.25-0.45)**	**0.51 (0.38-0.67)**	**0.51 (0.35-0.75)**	**0.56 (0.43-0.74)**	1.09 (0.84-1.41)
AUDIT-C screening scale	**0.27 (0.20-0.36)**	**0.09 (0.07-0.12)**	**0.03 (0.02-0.04)**	**0.61 (0.54-0.69)**	**0.25 (0.23-0.29)**
Liver disease	**0.34 (0.21-0.54)**	**0.23 (0.15-0.36)**	**0.17 (0.10-0.29)**	0.71 (0.36-1.41)	0.63 (0.34-1.14)
DSM-IV depressive/ anxiety disorders	**0.42 (0.29-0.63)**	**0.37 (0.22-0.62)**	**0.67 (0.49-0.93)**	1.32 (1.03-1.69)	**0.70 (0.51-0.96)**
DSM-IV substance use disorders	**0.15 (0.09-0.23)**	**0.01 (0.01-0.02)**	**0.24 (0.12-0.50)**	1.37 (1.10-1.69)	1.04 (0.81-1.33)

Note: aOR = 1 indicates no risk; aOR < 1 indicates decreased risk; aOR > 1 indicates increased risk. aORs with CIs that include 1.0 are not statistically significant. Significant reductions in outcome are in **bold**.

It appears that any reduction in the WHO risk drinking level by Wave 2 was associated with a significant decrease in the risk for a wide variety of outcomes that reflect several aspects of how individuals feel and function. Thus, the implications of the findings particularly for this highest risk group, arguably the group most in need of treatment, are that reductions in WHO risk levels can be used as valid indicators of improvement of how a patient might feel and function under natural conditions. As such this suggests that those same reductions observed in AUD clinical trials do have meaning and can be translatable to real world settings.

WHO Risk Drinking Level Reductions in an AUD Clinical Trial Sample - Relationship to Behavioral and Physical Function

As shown above, a reduction of the WHO risk drinking level over time appears to be indicative of improvement in how people feel and function in real world/natural settings. A next step would be to evaluate how this same metric reflects improvements in individual functioning in AUD pharmacotherapy trials. ACTIVE examined the validity of reductions in WHO risk drinking levels in a very large alcohol clinical trial.

Methods

To examine the clinical relevance of reductions in WHO risk drinking levels in an alcohol clinical trial we used the COMBINE study data (Anton et al., 2006), a multi-site, randomized, double-blind placebo-controlled clinical trial conducted in the United States that evaluated the efficacy of combinations of medications and behavioral interventions in the treatment of alcohol dependence. To be included in COMBINE, all participants met the criteria for alcohol dependence based on the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (American Psychiatric Association, 1994) and reported at least 2 heavy drinking days in a consecutive 30-day period within the 90 days prior to the baseline assessment. Exclusion criteria included the presence of another substance use disorder (other than nicotine or cannabis), a psychiatric disorder requiring medication, or unstable medical conditions, including serum liver enzyme levels that were more than 3 times the upper limit of normal.

Participants (n=1383) were randomized using a 2 x 2 x 2 design in which they received: 1) active naltrexone (100 mg/day) or placebo naltrexone, 2) active acamprosate (3000 mg/day) or placebo acamprosate, 3) medication management with a combined behavioral intervention (CBI) or medication management (MM) alone, and an additional group received the CBI intervention without MM or pills. Participants completed follow-up assessments at the end of treatment (week 16) and at 3 post-treatment follow-ups: 10 weeks (week 26 after baseline), 36 weeks (week 52 after baseline), and 1 year following treatment (week 68 after baseline).

COMBINE measures of WHO risk drinking levels. Daily alcohol consumption was measured using the calendar-based Form-90 (Miller, 1996) and Timeline Follow-Back interview (Sobell & Sobell, 1992). We calculated WHO risk drinking levels (see Table 1) based on participants' reports of the number of standard drinks (defined as 0.6 ounces of absolute alcohol) consumed, which we converted to grams of pure alcohol (0.6 ounces=14 grams). Then, WHO risk drinking levels were calculated based on the average grams of alcohol consumed per day (i.e., drinks per day). Importantly, the WHO risk drinking level based on average grams per day incorporates both drinking days and abstinent days to define average alcohol consumption across days over a specific time period (in the current study we averaged over 1-month time periods). For the baseline period, we calculated the WHO risk drinking level using data from the month prior to screening. The final endpoint WHO risk drinking level during treatment was defined during the last month of treatment, assessed at week 16.

COMBINE outcome measures. The COMBINE study also included self-report measures of several conditions and biological markers/laboratory results. These measures assess aspects of how participants were feeling and/or functioning at the post-treatment (week 16), and/or 36 weeks (week 52 post-baseline) or 52 weeks (week 68 post-baseline) follow-ups. Outcome measures included the following:

1. Drinking Consequences. Drinking consequences were assessed with the Drinker Inventory of Consequences (DrInC; Miller, Tonigan, & Longabaugh, 1995)MD","title":"The Drinker Inventory of Consequences (DrInC), a 50-item measure that uses a 4-level response scale (0=never, 3=daily or almost daily). We used the DrInC total score (based on 45 drinking consequences, excluding the 5 control items)

to assess alcohol-related consequences over the prior 3 months at baseline, the prior 2 months at the end of treatment (week 16), and the prior 4 months at the 1-year follow-up (week 68, 1-year post-treatment). DrInC internal consistency and reliability exceeded Cronbach's α = 0.93 at all time periods.

2. Mental Health Functioning. Mental health was assessed using the 12-item Short Form Health Survey (SF-12; Ware, Kosinski, & Keller, 1996) mental health subscale, which included 6 items assessed on a Likert-type response scale (1=all of the time, 5=none of the time). We used T-scores (with average functioning of 50 and standard deviation of 10 in the general population) from the SF-12 mental health subscale, with higher scores indicating better mental health functioning over the past month at baseline, end of treatment (week 16), and the last follow-up at which the SF-12 was administered (week 52, 9 months following treatment). The reliability of the SF-12 items exceeded Cronbach's α = 0.80 at all time periods.

3. Physical/Biological Markers. Systolic blood pressure (SBP) was assessed at each clinic visit by clinical staff. Blood samples for liver enzyme levels and liver cell pathology, including percent carbohydrate-deficient transferrin (%CDT), aspartate aminotransferase (AST), alanine aminotransferase (ALT), and γ-glutamyltransferase (GGT), were collected at specified clinic visits. AST, ALT, and GGT are liver enzyme markers which typically reflect liver cell damage while %CDT is a sensitive and alcohol specific biochemical marker of alcohol's effects on the protein transferrin. SBP and liver enzyme/pathology levels were included as biological markers of physical health that have been shown to be associated with heavy drinking and AUD (Baros et al., 2008; LoCastro et al., 2009; Stewart et al., 2008). SBP was examined rather than diastolic blood pressure (DBP) because prior research has shown SBP is more impacted by reductions in alcohol use than DBP (Stewart et al., 2008). Lower SBP and liver enzyme/pathology levels are associated with more positive health outcomes.

4. Quality of Life. Quality of life was assessed with the World Health Organization Quality of Life Brief version (WHOQOL-BREF; World Health Organization, 1998), a 25-item measure with response options ranging from 1 ("not at all") to 5 ("an extreme amount") of quality of life in four domains: physical health, psychological health, social relationships, and environmental quality of life. Higher scores on each domain indicate better quality of life with maximum possible domain scores of 35 (physical health), 30 (psychological health), 15 (social health) and 40 (environmental domain). Scores in the COMBINE sample represented the full range of scores on each domain. Internal consistency reliabilities for each domain exceeded Cronbach's α=.70 at all time points.

Analysis. The method of analysis was the same for all outcomes. Data from the post-treatment assessments were used to examine the associations between WHO risk drinking level reduction and biopsychosocial indicators of clinical benefit. Regression analyses with WHO risk level reduction as an independent variable were conducted with the following covariates: baseline levels of the clinical outcome, sex (male=1), age, body mass index, smoking status (smoker=1), and WHO risk drinking level at baseline. Results are reported as adjusted, unstandardized regression coefficients betas (B) (with standard errors), which can be interpreted as the decrease in outcomes based on achieving at least the 1- and 2-level reductions, holding all other covariates constant. B >0.00 indicates an increase in the outcome versus the reference group, and B <0.00 indicates a decrease in the outcome relative to the reference group. Statistical significance of a B is indicated by a 95% confidence interval that does not include 0.00 between its lower and upper limit. The Cohen's *d* effect sizes are also reported, which represent the standardized mean differences between groups, where *d*=.2 is a small effect, *d*=.5 is a medium effect, and *d*= .8 is a large effect (Cohen, 1992).

For all analyses, we computed 2 binary variables that reflected at least 1- or 2-level reductions in the WHO risk drinking levels based on the reduction from baseline to the last month of treatment. The reference group for the 1-level reduction was no change or an increase in the WHO risk drinking level from baseline to the last month of treatment and the reference group for the 2-level reduction was the 1-level reduction, no change, or increase in the WHO risk level from baseline to the last month of treatment. This reference

group was chosen to be consistent with the EMA guidelines which define a responder as achieving at least a 2-level reduction (and thus a non-responder would someone who did not achieve those reductions). Individuals with missing data on the WHO risk drinking level were coded as nonresponders (e.g., no change or increase in WHO risk drinking level). Missing data in the outcomes were estimated using maximum likelihood estimation, thus all individuals were included in the regression analyses.

Results

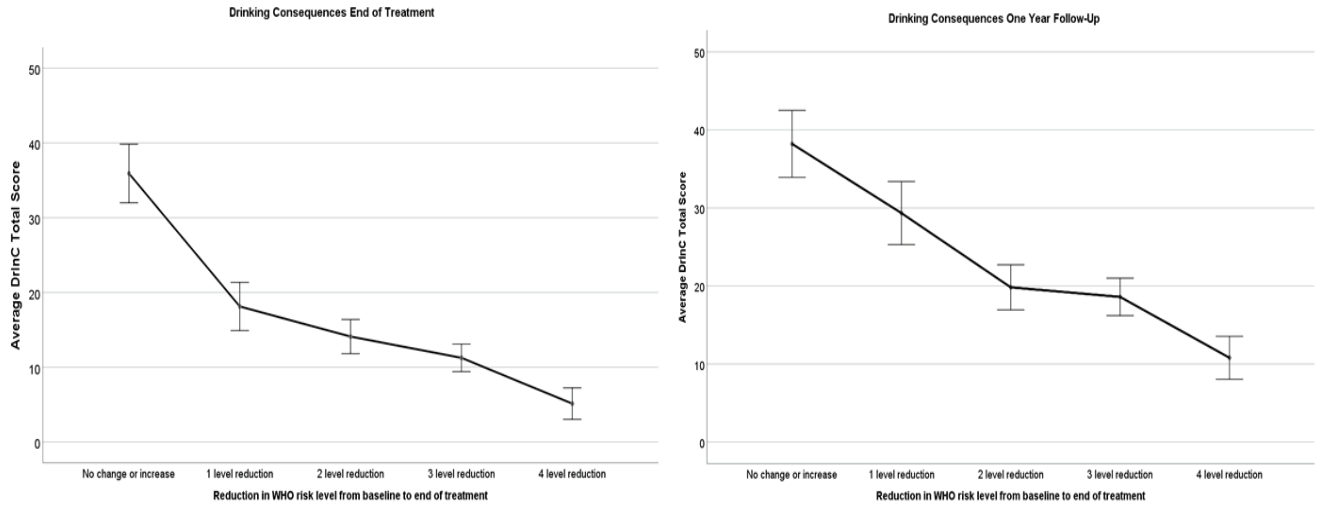
The number of participants categorized at each WHO risk drinking level based on drinking during the 28 days prior to screening and the last month of treatment were examined. At baseline, most individuals (69.1%) were in the "very high-risk" category and there were no abstainers. During the last month of treatment (month 4), 36.3% of the sample was abstinent, 39.4% categorized as "low risk" drinkers, 9.9% categorized as "medium risk" drinkers, 6.3% categorized as "high risk drinkers", and 8.1% categorized as "very high-risk" drinkers. The binary WHO drinking risk level reduction variables were created by calculating the reduction in risk drinking level from baseline to the last month of treatment. The majority of the sample reduced WHO risk drinking by at least 1 level (n=1011, 82.5%) or at least 2 levels (n=881, 71.9%) from baseline to the last month of treatment.

Drinking Consequences. As shown in Table 10, individuals with at least 1-level or at least 2-level reductions in WHO risk level from baseline to the last month of treatment reported significantly fewer drinking consequences at the end of treatment and up to 1 year following treatment (all $p<0.001$) with large effect sizes reductions in drinking consequences in those with at least 1-level or at least 2-level reductions (all Cohen's $d>.80$). As reported by Witkiewitz et al (2017), and shown in Figure 1, even a 1-level reduction in WHO risk drinking was associated with large and significant reductions in drinking consequences at the end of treatment and up to 1 year following treatment.

Table 10. Drinking consequences by WHO risk drinking level reduction from baseline to last month of treatment

End of Treatment (week 16) Drinking Consequences			
WHO Risk Drinking: 1 Level Reduction	n	Mean drinking consequences (SD)	Adjusted B (95% CI)
No change or increase	214	35.91 (26.16)	Reference
At least 1-level reduction	1011	11.05 (16.64)	-26.29 (-31.66, -20.92)
WHO Risk Drinking: 2 Level Reduction	n	Mean drinking consequences (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	25.30 (23.49)	Reference
At least 2-level reduction	881	10.09 (16.16)	-20.76 (-23.00, -18.52)
One Year Follow-Up Drinking Consequences			
WHO Risk Drinking: 1 Level Reduction	n	Mean drinking consequences (SD)	Adjusted B (95% CI)
No change or increase	214	38.21 (27.22)	Reference
At least 1-level reduction	1011	18.13 (20.25)	-19.44 (-26.12, -12.77)
WHO Risk Drinking: 2 Level Reduction	n	Mean drinking consequences (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	33.51 (25.43)	Reference
At least 2-level reduction	881	16.52 (19.30)	-20.28 (-23.84, -16.72)

Figure 1. Average Drinker Inventory of Consequences (DrInC) total scores by reduction in WHO Risk level. Vertical bars indicate 95% Confidence Intervals (CIs).



Mental Health Functioning. As shown in Table 11, individuals with at least 1-level or at least 2-level reductions in WHO risk drinking level from baseline to the last month of treatment reported significantly greater mental health functioning, as measured by SF-12 mental health composite T-scores (with average functioning of 50 and standard deviation of 10 in the general population), at the end of treatment and up to 9 months following treatment (all $p<0.001$) with medium effect sizes improvements in mental health functioning in those with at least 1-level or at least 2-level reductions (all Cohen's $d>.60$).

Table 11. Mental health functioning by WHO risk drinking level reduction from baseline to the last month of treatment (16 weeks)

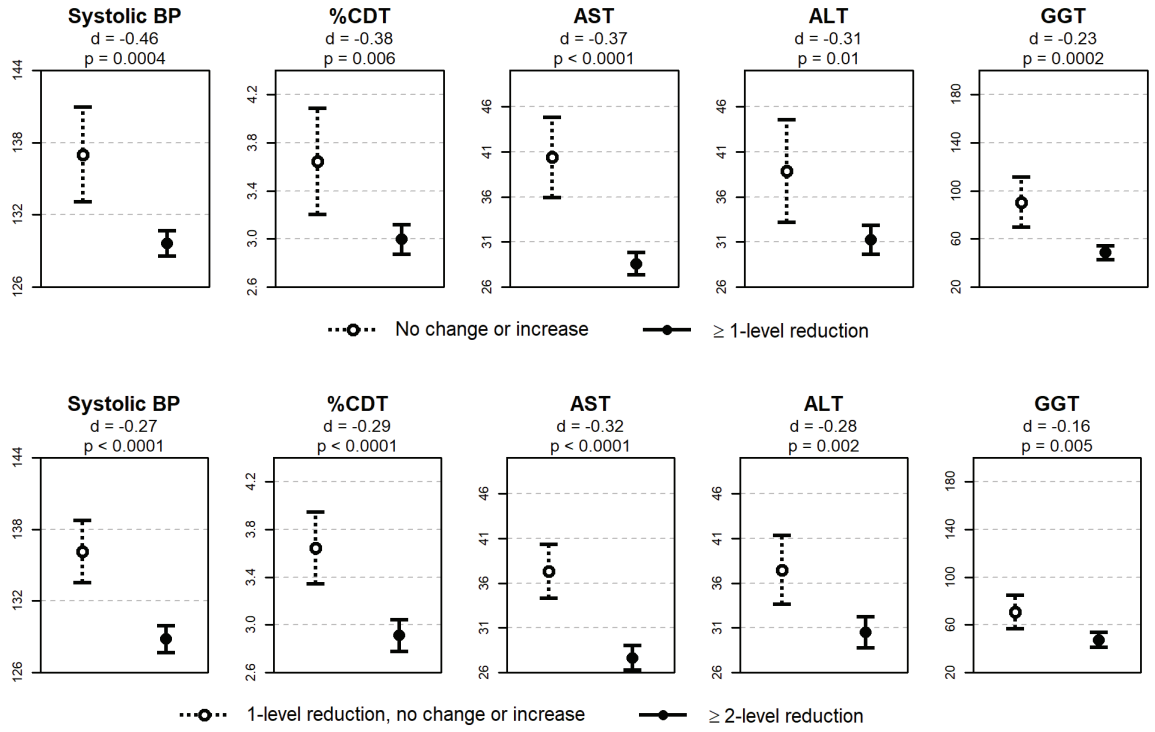
End of Treatment (week 16) Mental Health Functioning			
WHO Risk Drinking: 1 Level Reduction	n	Mean mental health (SD)	Adjusted B (95% CI)
No change or increase	214	42.67 (11.37)	Reference
At least 1-level reduction	1011	50.06 (9.18)	7.75 (5.28, 9.82)
WHO Risk Drinking: 2 Level Reduction	n	Mean mental health (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	44.97 (10.99)	Reference
At least 2-level reduction	881	50.55 (8.89)	6.86 (5.51, 8.22)
9-Month Follow-Up Mental Health Functioning			
WHO Risk Drinking: 1 Level Reduction	n	Mean mental health (SD)	Adjusted B (95% CI)
No change or increase	214	43.65 (10.00)	Reference
At least 1-level reduction	1011	48.34 (10.26)	4.54 (3.25, 5.83)
WHO Risk Drinking: 2 Level Reduction	n	Mean mental health (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	45.42 (9.79)	Reference
At least 2-level reduction	881	48.54 (10.36)	4.25 (2.72, 5.78)

Note: Higher scores indicate improvement in mental health functioning.

Biological Markers. As reported in Witkiewitz et al (in press) and summarized in Table 12, individuals with at least 1-level or at least 2-level reductions in WHO risk drinking level, from baseline to the last month of treatment, reported significantly lower SBP, lower %CDT, lower AST, lower ALT, and lower GGT (all $p<0.01$). As summarized in Figure 2, there were small-to-medium effect sizes reductions in biological markers in those with at least 1-level or at least 2-level reductions (all Cohen's $d>.20$).

Table 12. Biological markers by WHO risk drinking level reduction from baseline to the last month of treatment			
Systolic Blood Pressure (SBP)			
WHO Risk Drinking : 1-Level Reduction	n	Mean SBP (SD)	Adjusted B (95% CI)
No change or increase	214	137.77 (17.53)	Reference
At least 1-level reduction	1011	129.82 (16.56)	-7.87 (-11.77, -3.97)
WHO Risk Drinking: 2-Level Reduction	N	Mean SBP (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	134.16 (17.03)	Reference
At least 2-level reduction	881	129.54 (16.58)	-7.31 (-11.02, -3.61)
% Carbohydrate-deficient Transferrin (%CDT)			
WHO Risk Drinking: 1-Level Reduction	n	Mean %CDT (SD)	Adjusted B (95% CI)
No change or increase	214	3.68 (2.12)	Reference
At least 1-level reduction	1011	2.96 (1.81)	-0.65 (-0.95, -0.34)
WHO Risk Drinking: 2-Level Reduction	n	Mean %CDT (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	3.52 (2.89)	Reference
At least 2-level reduction	881	2.91 (1.72)	-0.73 (-1.11, -0.36)
Aspartate Aminotransferase (AST)			
WHO Risk Drinking: 1-Level Reduction	n	Mean AST (SD)	Adjusted B (95% CI)
No change or increase	214	40.88 (35.38)	Reference
At least 1-level reduction	1011	28.27 (15.99)	-11.56 (-20.96, -2.15)
WHO Risk Drinking: 2-Level Reduction	n	Mean AST (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	36.25 (28.99)	Reference
At least 2-level reduction	881	27.65 (14.71)	-9.86 (-14.17, -5.56)
Alanine Aminotransferase (ALT)			
WHO Risk Drinking: 1-Level Reduction	n	Mean ALT (SD)	Adjusted B (95% CI)
No change or increase	214	40.79 (32.34)	Reference
At least 1-level reduction	1011	30.77 (23.48)	-8.13 (-15.23, -1.04)
WHO Risk Drinking: 2Level Reduction	n	Mean ALT (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	38.00 (28.82)	Reference
At least 2-level reduction	881	30.08 (23.06)	-7.74 (-11.25, -4.22)
γ-Glutamyltransferase (GGT)			
WHO Risk Drinking: 1-Level Reduction	n	Mean GGT (SD)	Adjusted B (95% CI)
No change or increase	214	142.78 (450.92)	Reference
At least 1-level reduction	1011	43.47 (61.97)	-41.09 (-72.15, -10.03)
WHO Risk Drinking: 2-Level Reduction	n	Mean GGT (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	86.93 (301.86)	Reference
At least 2-level reduction	881	42.96 (63.34)	-23.66 (-38.28, -9.04)

Figure 2. WHO Risk Drinking Level Reductions from Baseline to End of Treatment Predicting End of Treatment Biomarkers (N=1142) for at least 1- and 2-Level WHO Risk Drinking Reduction with 95% Confidence Intervals. Cohen's d indicates effect size of reduction vs. no change.



Quality of Life. As shown in Table 13, individuals with at least 1-level or at least 2-level reductions in WHO risk drinking level from baseline to the last month of treatment reported significantly greater quality of life across all domains at the end of treatment (all p<0.01) with small-to-medium effect sizes improvements in mental health functioning in those with at least 1-level or at least 2-level reductions (all Cohen's d>.20, see Witkiewitz et al., in press).

Table 13. Quality of Life by WHO risk drinking level reduction from baseline to the last month of treatment			
End of Treatment Quality of Life (QoL)			
WHO Risk Drinking: 1- Level Reduction	n	Mean Physical QoL (SD)	Adjusted B (95% CI)
No change or increase	214	27.11 (4.81)	Reference
At least 1-level reduction	1011	29.37 (4.19)	2.21 (1.88, 2.53)
WHO Risk Drinking: 2-Level Reduction	n	Mean Physical QoL (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	27.82 (4.31)	Reference
At least 2-level reduction	881	29.50 (4.24)	2.20 (1.66, 2.73)
WHO Risk Drinking: 1-Level Reduction	n	Mean Psychological QoL (SD)	Adjusted B (95% CI)
No change or increase	214	21.06 (3.81)	Reference
At least 1-level reduction	1011	22.97 (4.12)	2.21 (1.67, 2.75)
WHO Risk Drinking: 2-Level Reduction	n	Mean Psychological QoL (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	21.19 (4.45)	Reference
At least 2-level reduction	881	23.09 (4.11)	2.23 (1.75, 2.72)
WHO Risk Drinking: 1-Level Reduction	n	Mean Social QoL (SD)	Adjusted B (95% CI)
No change or increase	214	10.27 (2.63)	Reference
At least 1-level reduction	1011	10.90 (2.57)	0.77 (0.36, 1.18)
WHO Risk Drinking: 2-Level Reduction	n	Mean Social QoL (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	10.32 (2.55)	Reference
At least 2-level reduction	881	10.93 (2.57)	0.83 (0.49, 1.18)
WHO Risk Drinking: 1-Level Reduction	n	Mean Environmental QoL (SD)	Adjusted B (95% CI)
No change or increase	214	28.86 (6.46)	Reference
At least 1-level reduction	1011	31.57 (5.36)	2.14 (1.06, 3.21)
WHO Risk Drinking: 2-Level Reduction	n	Mean Environmental QoL (SD)	Adjusted B (95% CI)
1-level reduction, no change or increase	344	29.37 (6.01)	Reference
At least 2-level reduction	881	31.82 (5.28)	2.33 (1.45, 3.21)

Note: increase in score indicates improvement.

Stability of WHO Risk Level Reductions. In addition, data from the 1-year follow-up assessment was used to assess stability of the WHO risk level reductions over time. Specifically, we were interested in whether those who achieved at least 1- and/or 2-level reductions in WHO risk level were able to maintain the 1- and 2-level reductions at 1 year following treatment. Cross-tabulations of WHO risk levels across post-treatment and 1-year follow-up timepoints were conducted to assess stability of WHO risk level over time.

Results indicated that 85.5% of those who achieved at least a 1-level reduction by the end of treatment reported at least a 1-level reduction at 1 year following treatment (χ^2 (1) =147.23, p<0.001), and 77.8% of those who achieved at least a 2-level reduction by the end of treatment reported at least a 2-level reduction by 1 year following treatment (χ^2 (1) =204.50, p<0.001). Thus, there was significant stability of achieving the WHO risk reduction endpoints over time.

Summary and Conclusions

As summarized in Table 14, at least 1- and 2-level reductions in the WHO risk drinking level were associated with significant improvements at the end of treatment for a variety of outcomes that reflect different aspects of how individuals feel and function.

Table 14. Summary of results unstandardized regression results (B (95% CI)) on reductions in WHO risk drinking levels in the COMBINE Study

	Risk Drinking level reduction	
Outcomes	At least 1-level reduction vs. no change or increase	At least 2-level reduction vs. 1-level reduction, no change or increase
Drinking Consequences	-26.29 (-31.66, -20.92)	-20.76 (-23.00, -18.52)
SF-12 Mental Health Functioning	7.75 (5.28, 9.82)	6.86 (5.51, 8.22)
Systolic Blood Pressure	-7.87 (-11.77, -3.97)	-7.31 (-11.02, -3.61)
%CDT	-0.65 (-0.95, -0.34)	-0.73 (-1.11, -0.36)
AST	-11.56 (-20.96, -2.15)	-9.86 (-14.17, -5.56)
ALT	-8.13 (-15.23, -1.04)	-7.74 (-11.25, -4.22)
GGT	-41.09 (-72.15, -10.03)	-23.66 (-38.28, -9.04)
Physical Quality of Life	2.21 (1.88, 2.53)	2.20 (1.66, 2.73)
Psychological Quality of Life	2.21 (1.67, 2.75)	2.23 (1.75, 2.72)
Social Quality of Life	0.77 (0.36, 1.18)	0.83 (0.49, 1.18)
Environmental Quality of Life	2.14 (1.06, 3.21)	2.33 (1.45, 3.21)

The effects of reductions in WHO risk levels through the last month of treatment were associated with fewer drinking consequences and improvements in mental health through 9 months and 1 year following treatment (Witkiewitz et al., 2017). Reductions in WHO risk drinking levels were also stable through 1 year following treatment for the large majority of individuals who achieved the WHO risk level reductions.

Reductions in WHO Risk Drinking Levels and Health Care Costs in an Alcohol Use Disorder Clinical Trial Sample

In this section we examine another important outcome—health care costs —that are of interest to researchers and policymakers who want to assess improved alcohol-related functioning as it might relate to reductions in WHO risk drinking levels observed in clinical trials. Specifically, we first estimated the relationship between reductions in WHO risk drinking levels during the 4 months of the COMBINE clinical trial (same trial used for the previous analysis) and then we evaluated the long-term effect of these WHO risk drinking reductions on health care costs 1-year post treatment and 3-years post randomization.

Methods

We used data from the COMBINE trial (Anton et al., 2006), a multi-site randomized controlled trial designed to measure the effectiveness of combinations of 2 pharmacological treatments—Acamprosate and Naltrexone—and a behavioral intervention. The COMBINE trial recruited 1,383 patients across 11 sites. Patients received treatment for 16 weeks and provided data up through a year following treatment. A subset of patients (N=786 from 9 study sites) volunteered to participate in additional follow-up data collection for up to 3 years after randomization to support a cost and cost-effectiveness analysis called the COMBINE Economic Study (CES) (Zarkin et al., 2010).

We used data from the Economic Form 90 (Bray et al., 2007) and timeline follow-back (TLFB) instruments (Sobell & Sobell, 1992) collected at baseline, during treatment (at weeks 8 and 16 post-randomization), and after treatment (weeks 26, 52, and 68). CES study subjects participated in additional interviews at weeks 88, 105, 122, 139, and 156. The Economic Form 90 collected demographic, socioeconomic (e.g.., employment, income, and criminal justice), and health care utilization data. We used unit costs for health care events from the peer-reviewed literature (French & Martin, 1996; Roebuck, French, & McLellan, 2003; Zarkin et al., 2010) and inflated them to 2016 U.S. prices using the medical care consumer price index (Bureau of Labor Statistics, 2018).

COMBINE measures of WHO risk drinking levels. Daily alcohol consumption was measured using the calendar-based Form-90 (Miller, 1996) and Timeline Follow-Back interview (Sobell & Sobell, 1992). Mutually-exclusive risk drinking groups were created based on sustained reductions in drinking across the entire 16-week treatment period: 1) no sustained risk drinking reduction, 2) sustained risk drinking reduction of 1 level, but not 2 or more levels, 3) sustained risk drinking reduction of at least 2 levels. A sensitivity analysis was conducted evaluating the WHO risk drinking reduction over different time periods during the trial (e.g., the last month, the last 2 months, the last 3 months of the trial).

COMBINE Health Care Outcomes. The dependent variable in our analysis is total healthcare costs calculated by multiplying the number of healthcare events recorded on the Economic Form 90 by the unit cost in 2016 U.S. dollars. The Economic Form 90 collected data on inpatient hospital stays (for behavioral and physical health), outpatient visits (for behavioral and physical health), and emergency department visits. The dependent variable for the 1-year cost analysis is the sum of total healthcare costs from the end of treatment (week 16) through 1-year post treatment completion. The dependent variable for the 3-year analysis is the sum over the entire data collection period from the end of treatment (week 16) through 3-years post treatment completion.

Covariates. Covariates in our analysis include demographic controls (gender, race/ethnicity, and age), socioeconomic controls (indicators of unemployment, marital status, and years of educational attainment), and health-related controls (indicators of whether participants had ever used marijuana or other illicit drugs and physical and psychological health domain scores from the WHO Quality of Life Instrument (The WHOQOL Group, 1998)). We also controlled for pre-randomization health care costs and the study site

where the patient received COMBINE treatment. Finally, because there was some attrition over the 3 years of data collection, we included in the 3-year cost models a covariate for the number of days for which each patient reported their costs (Aldridge et al., 2016).

Analysis. We estimated generalized gamma regression models with a log-link function to account for the positive skew commonly encountered in health care cost data (Manning & Mullahy, 2001). Tests of the distribution and heteroskedasticity of log-scale residuals and the modified Park test as suggested by Manning and Mullahy (2001) confirmed that this was the proper estimation technique. The dependent variables in our analysis were health care costs over 1-year post-treatment and 3-years post randomization.

For each dependent variable, we estimated models comparing health care costs for people with a reduction of exactly 1 level and a reduction of at least 2 levels to those with no shift or an increase in risk. We estimated the models with heteroskedasticity-consistent standard errors (White, 1980). We also present model-predicted average costs per reduction group and percent differences in cost relative to the reference group. We estimated this model for 4 time periods: the entire 4-month treatment period, the last 3 months of treatment (i.e., excluding the first month), the last 2 months of treatment, and the last month of treatment only. We considered different time periods to explore the extent to which our findings are sensitive to the period over which post-treatment initiation risk drinking is measured. One motivation for dropping the first months of treatment is to provide a "grace period" under the assumption that treatment might not be immediately effective (see also Aldridge et al., 2016; Witkiewitz et al., 2017 who look at drinking only in the last month of treatment.)

Results

For 1-year post-treatment, there is no statistically significant association between health care costs and a sustained 1-level reduction of WHO risk drinking; however, we find that sustaining a risk drinking reduction of at least 2 levels over the course of either the entire 4-month treatment period, the last 3 months or the last 2 months is associated with a significant reduction in health care costs relative to subjects who do not sustain a risk drinking shift (Table 15, column 1 – 3). While the magnitude and significance of the results are barely affected by dropping the first 1 or 2 months of treatment, we find that the effect is diminished when only the last month of treatment is considered – compare the 52.3% lower health care costs for a sustained at least 2-level shift over the 4-month treatment period ($p < 0.001$) to a 31.7% reduction for those who sustained the same reduction for the last month of treatment only ($p = 0.10$).

For 3-years post-randomization, the magnitude and significance of the associations are slightly diminished, consistent with the attenuation demonstrated in previous analyses (e.g., Aldridge et al., 2016). One notable difference between the 1- and 3-year analyses is that the association between sustaining only a 1-level risk reduction and health care costs is significant (Table 14, column 5; $p <0.05$). This may be explained by the fact that the group that sustains a 1-level shift but not a 2-level shift is the smallest (N=132 in the 1-year analysis; N=87 in the 3-year analysis) and likely sensitive to small fluctuations brought on by the exclusion of some subjects in the 3-year analysis. Note that the magnitudes of the coefficients are similar.

Table 16 presents the predicted costs for the WHO risk drinking groups. The difference in 1-year post-treatment costs among those who sustain at least a 2-level shift over the 4-month treatment period compared to those who do not sustain a shift is $2,188, 52.3% of the estimated costs for those who do not sustain a shift (Table 15, column 1, $4,182 vs $1,994). The difference in costs over 3-years post-randomization for the same drinking reduction group is $4,966, 44.0% of the estimated costs for those who do not sustain a shift (Table 15, column 5, $11,287 vs $6,321). The predicted reductions in costs for those who sustain a 2-level reduction in the final 3, 2, and 1 months of the trial are similar in magnitude. The estimated reductions at 1-year for those who sustain a 2-level reduction in the final 3, 2, and 1 months of the trial are $1,959 (49.3%), $1,624 (42.9%), and $1,095 (31.7%), respectively. At 3 years, the estimated reductions in costs are $5,077 (44.7%), $5,351 (45.6%), and $2,957 (30.5%).

Table 15. Estimated reductions in healthcare costs by alternative periods of sustained reductions in WHO risk levels

	1 Year Post-Treatment				3 Years Post-Randomization			
	All 4 Months	Last 3 Months	Last 2 Months	Last 1 Month	All 4 Months	Last 3 Months	Last 2 Months	Last 1 Month
Sustained WHO 1-level but not 2-level reduction throughout part of treatment period (ref: did not sustain a risk drinking reduction)	-0.3682	-0.1971	-0.1837	-0.2997	-0.4283*	-0.3341	-0.3041	-0.1767
	(0.2320)	(0.2553)	(0.2625)	(0.2968)	(0.2185)	(0.2342)	(0.2492)	(0.2859)
	[-30.8%]	[-17.9%]	[-16.8%]	[-25.9%]	[-34.8%]	[-28.4%]	[-26.2%]	[-16.2%]
Sustained at least WHO 2-level reduction throughout part of treatment period (ref: did not sustain a risk drinking reduction)	-0.7406***	-0.6785***	-0.5602**	-0.3819	-0.5798**	-0.5925**	-0.6097**	-0.3632
	(0.1901)	(0.1987)	(0.2123)	(0.2329)	(0.1915)	(0.2046)	(0.2190)	(0.2472)
	[-52.3%]	[-49.3%]	[-42.9%]	[-31.7%]	[-44.0%]	[-44.7%]	[-45.6%]	[-30.5%]
Observations	964	964	964	964	651	651	651	651

Coefficients and robust standard errors in parentheses from Gamma GLM with log link. Percentage difference from reference category in brackets, calculated as exp(coefficient)-1. Each column reflects an individual model in which the length of period of a sustained shift is reduced. Models control for demographic and socioeconomic characteristics, health status, baseline risk drinking, and study center.

WHO risk levels based on average drinks per day.
* p<0.05, ** p<0.01, *** p<0.001

Table 16. Predicted healthcare costs based on the estimated model								
	1 Year Post-Treatment				3 Years Post-Randomization			
	All 4 Months	Last 3 Months	Last 2 Months	Last 1 Month	All 4 Months	Last 3 Months	Last 2 Months	Last 1 Month
Did not sustain a risk drinking reduction	$4,182	$3,977	$3,787	$3,450	$11,287	$11,358	$11,723	$9,710
Sustained WHO 1-level but not 2-level reduction throughout part of treatment period	$2,893	$3,265	$3,151	$2,557	$7,355	$8,132	$8,649	$8,137
	(-30.8%)	(-17.9%)	(-16.8%)	(-25.9%)	(-34.8%)	(-28.4%)	(-26.2%)	(-16.2%)
Sustained at least WHO 2-level reduction throughout part of treatment period	$1,994	$2,018	$2,163	$2,355	$6,321	$6,281	$6,372	$6,753
	(-52.3%)	(-49.3%)	(-42.9%)	(-31.7%)	(-44.0%)	(-44.7%)	(-45.6%)	(-30.5%)

Predicted costs by WHO level reduction group based on models with percentage differences in parentheses from the reference group (no sustained drinking reduction).

Summary and Conclusions

There is strong evidence that sustained reductions in WHO risk drinking levels through the 4 months of COMBINE treatment are associated with statistically significant and meaningfully lower health care costs post-treatment. One-year post-treatment we find that sustaining at least a 2-shift reduction is associated with 52.3% lower health care costs ($p < 0.001$) and that sustaining a 1-shift but not a 2-shift reduction in risk levels is associated with 30.8% lower health care costs (not statistically significant). The estimated relationship between health care costs and reductions in drinking risk levels was generally similar for the last 2 and 3 months of treatment but was smaller in magnitude and not statistically significant when only the last month of treatment is considered.

We note that the COMBINE study required participants to be abstinent from alcohol for at least 4 days before randomization. In such a study, sustained reductions in drinking from the outset of treatment may be meaningful; however, for other studies which do not require pretreatment abstinence and for medications where the effects may not be immediate, it might take some time before a stable and sustained reduction (WHO risk drinking level reduction) to occur and translate into stable reductions in health expenditures. Higher success rates based on sustained reductions in risk drinking during the later months of the treatment could encourage sustained participation in treatment.

We found that our estimated associations generally persist when costs incurred up to 3 years post-randomization are included with some attenuation observed in the magnitude and significance of the estimated effects.

Reductions in WHO Risk Drinking Levels as Outcomes in Alcohol Pharmacotherapy Trials

The above data has supported the idea that a reduction in WHO risk drinking level over time, in the general population or during the course of treatment, are associated with improvements in how individuals with AUD "feel and function" and reductions in health costs. A crucial additional issue was to evaluate how sensitive the WHO 1- and 2-level risk drinking reductions are to medication treatment effects in clinical trials, i.e. can these reductions capture medication efficacy over placebo in a statistically reliable manner? This is an important question from the perspective of clinical trial sponsors and patients alike. Ideally, one would like to determine if the treatment effect sizes (i.e., the magnitude of difference between medication and placebo on the outcome of interest) obtained using the WHO outcomes are at least comparable to those obtained using the current FDA-recommended drinking outcomes of "abstinence" and "no heavy drinking days". In one published study, a WHO 2-level risk drinking reduction differentiated the drug nalmefene from placebo in Phase III trials of alcohol dependence among individuals with baseline WHO "high or very high-risk" levels (Aubin 2015). However, these outcomes were not contrasted with the FDA-recommended endpoints.

To address these issues, the ACTIVE conducted secondary data analyses of the WHO 1- and 2-level risk drinking reduction outcomes to evaluate their sensitivity in distinguishing between active medication and placebo in 3 multisite alcohol pharmacotherapy trials. Response rates and medication effect sizes for WHO risk drinking reduction outcomes were compared with those for the current FDA-recommended outcomes.

Methods

Data were from 3 multisite, randomized, double-blind placebo-controlled alcohol dependence pharmacotherapy trials in which the active medications were superior to placebo in reducing alcohol consumption:

- COMBINE Study: (N=1383; Phase IV, 11 sites; 2001-2004) focusing on the naltrexone arms without a comprehensive behavioral therapy (n=607)
- Varenicline: (N=200; Phase II, 5 sites; 20011-2012) testing varenicline tartrate (2 mg/d)
- Topiramate: (N=371; Phase III, 17 sites; 2004-2006) testing oral topiramate (300 mg/d)

In the COMBINE trial, participants were required to abstain from alcohol for at least 4 days prior to randomization, whereas in the other trials, participants could drink up to randomization. Participants in all 3 trials were mainly middle-aged, white, men. The majority consumed alcohol at the Very-High or High-risk levels. These demographic and alcohol consumption profiles are typical in AUD clinical trials, suggesting that the results could be expected to generalize to other AUD trials with similar profiles.

Measures of WHO Risk Drinking Levels. All trials measured daily alcohol consumption using the calendar-based Form-90 (Miller, 1996) and Timeline Follow-Back interview (Sobell & Sobell, 1992). WHO risk drinking levels (Table 1) were calculated based on average grams of pure ethanol consumed per day (which includes zero for abstinent days) during 28-day periods before and during treatment. The WHO 1- and 2-level risk drinking reduction outcomes were defined as the percentage of participants who decreased by at least 1 or 2 risk drinking levels, respectively, from baseline to the last 4 weeks of treatment (used in Table 16) or each month of treatment (used in Figure 3). For the 1-level reduction outcome, this included participants who reduced from Very High to High risk (or below); High to Medium risk (or below); and Medium to Low risk (or below). For the 2-level¬ reduction outcome, this included participants who reduced from Very High to Medium risk (or below); High to Low risk (or below); and Medium risk to Abstinence.

Measures of Other Endpoints. Abstinence and no heavy drinking days were computed for the same time periods. Abstinence was defined as no drinking during the month. No heavy drinking days was defined as never consuming 4 or more standard drinks (women) or 5 or more standard drinks (men) on any day during the month, with a standard drink containing 14 grams of ethanol.

Analysis. To facilitate the comparison of treatment effects (i.e., difference in prevalence rates obtained using the various drinking outcomes), a metric called Cohen's h was used. The size of the treatment effect using this metric can be interpreted as: small=0.20, medium=0.50, and large=0.80. The number needed to treat (NNT), another measure of size of the treatment effect, was also used. NNT is defined as the number of patients needed to be treated with the medication for 1 patient to be a responder on the given outcome compared with the control; lower NNT is indicative of a larger treatment effect size. Participants with any missing drinking data within a given 28-day period were considered "non-responders." We also performed a sensitivity analysis in which missing data were not imputed.

Results

Response Rates and Treatment Effect Sizes (last 4 weeks of treatment; Table 16). Across trials, the percentage of participants in the active condition achieving each measure of success during the last 4 weeks of treatment was lowest for abstinence (7%-35%), followed by no heavy drinking days (21%-51%), the WHO 2-level reduction (45%-75%), and WHO 1-level (55%-83%) reduction outcomes.

Across the trials, the size of the treatment effects and NNT for the WHO 1- and 2-level risk drinking reduction outcomes were as large or larger than outcomes currently accepted by the FDA (abstinence or no heavy drinking days). The naltrexone treatment effect for the WHO 2-level reduction outcome was greater than for abstinence, no heavy drinking days, and the WHO 1-level reduction. The varenicline treatment effect for the WHO 1- and 2-level reduction outcomes was larger than for either abstinence or no heavy drinking days. There was a slightly larger treatment effect for the WHO 1-level reduction than the 2-level reduction outcome. The topiramate treatment effect for the 2-level reduction outcome was comparable to that for no heavy drinking days, but smaller than for abstinence. The treatment effect for the 1-level reduction was the smallest of the outcomes.

Table 16. Treatment Effects for Traditional and WHO Risk Drinking Level Reduction Definitions of Response during the Last 4 Weeks of Treatment					
Trial (N)	**Responder Outcome**	**Placebo %**	**Medication %**	**h (95% CI)**	**NNT (95% CI)**
Naltrexone (n=288); Placebo (n=302)	Abstinent	28%	35%	0.142 (-0.020 – 0.303)	16 (111.9 [H] to ∞ to 7.1 [B])
	No heavy drinking days	44%	51%	0.140 (-0.021 – 0.302)	15 (96.2 [H] to ∞ to 6.6 [B])
	WHO 2-level reduction	65%	75%	0.214 (0.053 – 0.375)	11 (5.8 [B] to 41.4 [B])
	WHO 1-level reduction	79%	83%	0.116 (-0.046 – 0.277)	23 (55.7 [H] to ∞ to 9.2 [B])
Varenicline (n=96); Placebo (n=101)	Abstinent	4%	7%	0.146 (-0.133—0.426)	30 (32.6 [H] to ∞ to 10.3 [B])
	No heavy drinking days	15%	24%	0.232 (-0.048—0.511)	11 (54.3 [H] to ∞ to 5.0 [B])
	WHO 2-level reduction	42%	55%	0.273 (-0.006—0.553)	8 (486.2 [H] to ∞ to 3.6 [B])
	WHO 1-level reduction	54%	70%	0.338 (0.058—0.617)	7 (3.4 [B] to 34.3 [B])
Topiramate (n=179); Placebo (n=185)	Abstinent	3%	12%	0.369[a] (0.163—0.574)	12 (7.0 [B] to 26.2 [B])
	No heavy drinking days	13%	21%	0.207 (0.002—0.413)	13 (6.5 [B] to 1581.1 [B])
	WHO 2-level reduction	34%	45%	0.230 (0.024—0.435)	9 (4.7 [B] to 81.8 [B])
	WHO 1-level reduction	54%	55%	0.014 (-0.192—0.219)	144 (10.5 [H] to ∞ to 9.2 [B])

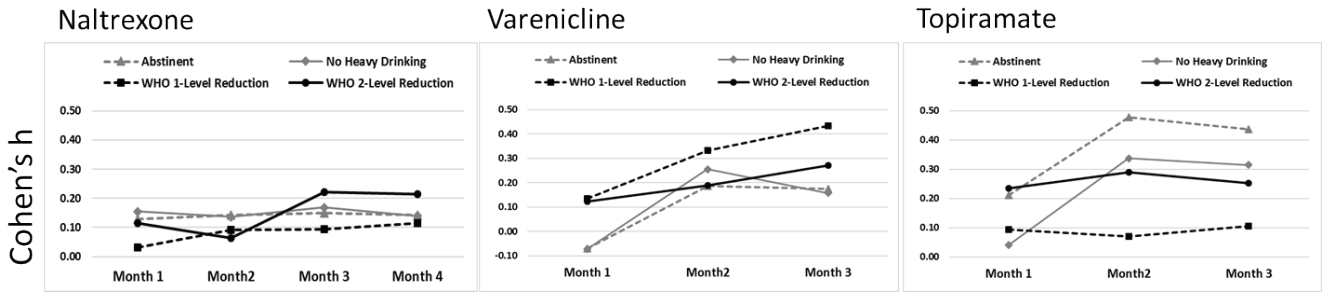
Note: Missing drinking data were imputed as non-responders for all outcomes. H=Number Needed to Harm, B=Number Needed to Benefit, WHO = World Health Organization. The WHO 1- and 2-level reduction outcomes represent the percentage of participants that reduce by 1 or more and 2 or more WHO risk drinking levels, respectively.

For Cohen's h, 95% confidence intervals (CI) that do not include zero are statistically significant, $p<0.05$. For NNT, statistically significant estimates are indicated by CIs with lower and upper bounds both expressed as Number Needed to Benefit (B). Nonsignificant estimates are indicated by CIs with a lower bound expressed as Number Needed to Harm (H) - indicating less efficacy in the medication than placebo group - and an upper bound expressed as B; infinity is included in the confidence interval because NNT is infinity when the absolute risk reduction is 0.

At first glance, the Cohen's h for the abstinence outcome in the topiramate trial is larger than might be expected given there is only a 9% difference between topiramate and placebo. Indeed, one artifact of using Cohen's h is that it gets disproportionately bigger for a given prevalence difference when one of the prevalence rates is very small. Here, it was the rate of abstinence in the placebo group that was very small (3%).

Treatment Effects by Month. As shown in Figure 3, treatment effects for all outcomes were generally smallest in the first month and increased in later months. Because participants in these trials (and individuals with AUD generally) typically have a long history of habitual heavy drinking, treatment providers and patients may reasonably expect that maximal drinking reductions may take 1 or more months to emerge, even with efficacious treatments.

Figure 3. Treatment effects (Cohen's h) for Responder Outcomes by Treatment Month



Note: Missing drinking data were imputed as non-responders for all outcomes.

Summary and Conclusions

Overall, outcomes based on reductions in WHO risk drinking levels, measured in grams of ethanol per day, performed well in 3 large multisite alcohol pharmacotherapy trials—reductions of 1 and 2 levels differentiated medication effects as well, or better than, outcomes of abstinence and no heavy drinking days. However, the 2-level reduction may have some advantages over the 1-level reduction. Specifically, the 2-level reduction produced larger effect sizes in 2 of 3 trials and had less variable and lower NNTs across studies (8-11 vs. 7-144).

In addition, the higher percentage of patients classified as responders using the WHO risk drinking reduction outcomes could lead to greater optimism about the value of treatment by patients and providers. For example, in the varenicline trial, 55% of participants experienced a clinically meaningful reduction in drinking (WHO-2 shift) versus only 7% who achieved abstinence, data that clinicians could use to encourage individuals to engage pharmacotherapy to assist in successful treatment.

Stability of the WHO Risk Drinking Reduction Outcomes in a 6-Month Clinical Trial

The FDA currently recommends a 6-month duration for phase 3 alcohol pharmacotherapy trials, however, most phase 2 clinical trials conducted to date were 3-4 months in duration. One question involves how stable the WHO 1- and 2-level risk drinking reduction outcomes are over the course of a longer (6-month) clinical trial and how this might inform trial length considerations required by the FDA for Phase III registration. To answer this question, we conducted a secondary data analysis of a recent NIAAA-sponsored 6-month multisite clinical trial evaluating the safety and efficacy of the medication, Horizant (gabapentin enacarbil), for the treatment of AUD (n=338) (primary manuscript under review in *Alcoholism: Clinical and Experimental Research*).

Methods

We computed WHO 1- and 2-level risk drinking reduction outcomes for each month during the 6-month treatment period. For this analysis the dependent variable was the WHO 1- or 2-level risk drinking reduction outcome from baseline to month 6 of the trial. We then examined how well the same WHO outcomes calculated in earlier Months 1-5 agree with the outcome at Month 6 . To formally quantify the level of agreement between each previous month and Month 6, we use a metric called kappa (k). A high k indicates substantial agreement between months, i.e. participants who were *successful* in attaining the WHO outcome in an early month (e.g., month 3) were also successful in attaining the WHO outcome in month 6, and participants who were *unsuccessful* in attaining the WHO outcome in an early month (e.g., month 3) were also unsuccessful in attaining the WHO outcome in month 6.

Results

In Figures 4 & 5 (below), we can see that by Month 3 there is already substantial agreement with Month 6 for both the WHO 1- and 2-level risk drinking reduction outcomes. Thus, it could be indicated that these outcomes stabilize by Month 3 and were sustainable through Month 6 (end of the trial).

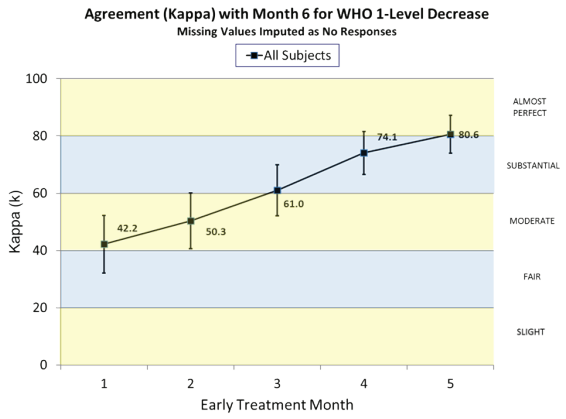


Figure 4:

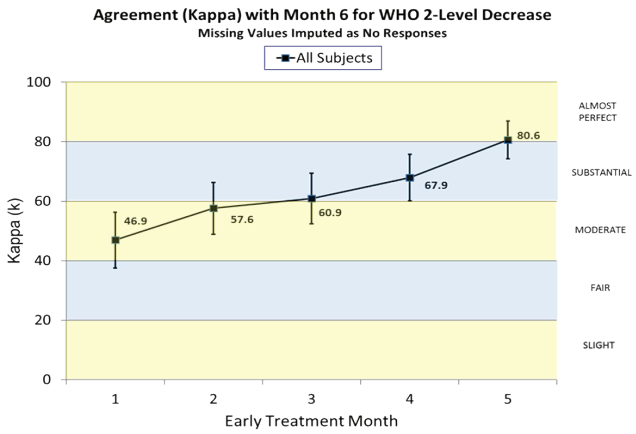


Figure 5:

Summary and Conclusions

This analysis shows that the WHO risk drinking reduction outcome is sufficiently stable by month 3-4 to suggest that trials of this length, using a WHO risk drinking outcome, might be considered for regulatory guidance. It is our opinion that a 3-4 month trial, compared to a 6-month trial for instance, would be more acceptable to clinical-trial participants thereby leading to higher adherence rates. Trials of 3-4 months would also be more cost-effective and less burdensome for medication development plans, leading to more drug development for AUD.

Review of Important Points and Observations:

- Alcohol Use Disorder is highly prevalent and costly to our society, leading to great suffering and disability.
- There have been only 3 drugs and 1 variant administration approved by the FDA in the last 60 years with the last one being over 12 years ago.
- It is becoming more apparent that many individuals do not seek treatment because of conservative treatment goals such as “complete abstinence”.
- There is growing data to suggest that a reduction in alcohol consumption is both more acceptable to individuals seeking treatment, can improve function, and reduce disease burden.
- Recently a new metric for measuring drinking reduction, the WHO risk drinking level change has been accepted in other countries, such as the EMA in Europe, as one potential endpoint of improvement in AUD trials.
- Both epidemiological and clinical trial samples show that a 1 or 2 level change in WHO risk drinking level is associated with improvements across an array of diagnostic, mental health, and physical functions, as well a reduction in health care costs..
- A reduction in WHO risk drinking level endpoint in clinical trials is able to distinguish efficacy of active medication from placebo in a reliable and statistically appropriate manner, and at a level equal to, or comparable with, currently FDA guided endpoints of “abstinence” and “no heavy drinking days”.
- Initial data suggests that the WHO risk drinking level reduction stabilizes around 3-4 months of a clinical trial, providing support for a shorter AUD clinical trial length than the currently FDA guided 6-month duration.
- Many more subjects would meet “success criteria” using the WHO risk drinking reduction metric thereby encouraging treatment seeking, monitoring, and long term treatment for this relapsing and costly illness. This has large public health consequences including likely health care cost savings.

References

Aldridge AP, Zarkin GA, Dowd WN, Bray JW. (2016). The Relationship Between End-of-Treatment Alcohol Use and Subsequent Healthcare Costs: Do Heavy Drinking Days Predict Higher Healthcare Costs? *Alcohol Clin Exp Res.* 40(5):1122-8.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.; DSM-IV)* (4th ed.). Washington, DC: Author.

Anton RF, et al. (2012). The Alcohol Clinical Trials Initiative (ACTIVE): purpose and goals for assessing important and salient issues for medications development in alcohol use disorders. *Neuropsychopharmacology.* 37(2): 402-411.

Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D., Donovan DM, … Zweben A. (2006). Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. *JAMA : The Journal of the American Medical Association,* 295(17):2003–17. http://doi.org/10.1001/jama.295.17.2003

Aubin HJ, Reimer J, Nutt DJ, et al. (2015). Clinical relevance of as-needed treatment with nalmefene in alcohol dependent patients. *Eur Addict Res,* 21(3):160-168

Baros AM, Wright TM, Latham PK, Miller PM, & Anton RF. (2008). Alcohol consumption, %CDT, GGT and blood pressure change during alcohol treatment. *Alcohol and Alcoholism*, 43(2):192–197.

Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR. (2007). AUDIT-C as a brief screen for alcohol misuse in primary care. *Alcohol Clin Exp Res,* 31(7):1208-17.

Bray JW, Zarkin GA, Miller WR, Mitra D, Kivlahan DR, Martin DJ, … Cisler RA. (2007). Measuring economic outcomes of alcohol treatment using the Economic Form 90. *Journal of Studies on Alcohol and Drugs,* 68(2):248–255.

Carey KB. (1997). Reliability and validity of the time-line follow-back interview among psychiatric outpatients: a preliminary report. *Psychology of Addictive Behaviors,* 11:26–33.

Cohen J (1992). A power primer. *Pscyho Bull,* 112(1):155-9.

Cottler LB, Grant BF, Blaine J, Mavreas V, Pull C, Hasin D, Compton WM, Rubio-Stipec M, Mager D. (1997). Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. *Drug Alcohol Depend*, 47(3):195-205.

Dawson DA, Grant BF, Stinson FS, Chou PS, Huang B, Ruan WJ. (2005). Recovery from DSM-IV alcohol dependence: United States, 2001-2002. *Addiction,* 100(3):281-92.

Dawson DA, Smith SM, Saha TD, Rubinsky AD, Grant BF. (2012). Comparative performance of the AUDIT-C in screening for DSM-IV and DSM-5 alcohol use disorders. *Drug Alcohol Depend,* 126(3):394-8

Falk DE, Ryan ML, Fertig JB, et al. (2018). A double-blind, placebo-controlled, multisite trial assessing the efficacy of gabapentin enacarbil extended-release for alcohol use disorder. To be submitted.

Falk DE, Wang X-Q, Liu L, Fertig F, Mattson M, Ryan M, Johnson B, Stout R, Litten RZ. (2010). Percentage of subjects with no heavy drinking days: Evaluation as an efficacy endpoint for alcohol clinical trials. *Alcohol Clin Exp Res*, 34:2022-2034.

French MT, Martin RF. (1996). The costs of drug abuse consequences: a summary of research findings. *J Subst Abuse Treat,* 13(6):453–466.

Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. (2003). The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend,* 71(1):7-16.

Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, Pickering RP, Ruan WJ, Smith SM, Huang B, Hasin DS. (2015). Epidemiology of DSM-5 alcohol use disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry,* 72(8):757–766.

Grant BF, Harford TC, Dawson DA, Chou PS, Pickering RP. (1995). The Alcohol Use Disorder and Associated Disabilities Interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. *Drug and Alcohol Dependence,* 39:37–44.

Hasin D, Carpenter KM, McCloud S, Smith M, Grant BF. (1997). The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. *Drug Alcohol Depend,* 44(2-3):133-41.

Hasin D, Grant BF, Cottler L, Blaine J, Towle L, Ustün B, Sartorius N. (1997). Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study. *Drug Alcohol Depend,* 47(3):217-26.

Hasin DS, Grant BF. (2004). The co-occurrence of DSM-IV alcohol abuse in DSM-IV alcohol dependence: results of the National Epidemiologic Survey on Alcohol and Related Conditions on heterogeneity that differ by population subgroup. *Arch Gen Psychiatry,* 61(9)891-6.

Hasin DS, Wall M, Witkiewitz K, Kranzler HR, Falk D, Litten R, Mann K, O'Malley SS, Scodes J, Robinson RL, Anton R .(2017). Change in non-abstinent WHO drinking risk levels and alcohol dependence: A 3 year follow-up study in the US general population. *Lancet Psychiatry.* 4:469-476.

Kline-Simon AH, Falk DE, Litten RZ, Mertens JR, Fertig J, Ryan M, Weisner CM. (2013). Posttreatment low-risk drinking as a predictor of future drinking and problem outcomes among individuals with alcohol use disorders. *Alcohol Clin Exp Res,* 37(Suppl 1):E373-E380.

Kline-Simon AH, Weisner CM, Parthasrathy S, Falk DE, Litten RZ, Mertens JR. (2014). Five-year healthcare utilization and costs among lower-risk drinkers following alcohol treatment. *Alcohol Clin Exp Res,* 38:579-586.

Laramee P, Leonard S, Buchanan-Hughes AB, Warnakula S, Daeppen JB, Rehm J. (2015). Risk of all-cause mortality in alcohol-dependent individuals: A systematic literature review and meta-analysis. *EBioMedicine,* 2:1394-1404.

Litten RZ, Ryan ML, Fertig JB, Falk DE, Johnson B, Dunn KE, Green AI, Pettinati HM, Ciraulo DA, Sarid-Segal O, Kampman K, Brunette MF, Strain EC, Tiouririne NA, Ransom J, Scott C, Stout R. (2013). A double-blind, placebo-controlled trial assessing the efficacy of varenicline tartrate for alcohol dependence. *J Addict Med,* 7:277–286.

Litten RZ, et al. (2010). Alcohol biomarkers in applied settings: recent advances and future research opportunities. *Alcoholism Clinical & Experimental Research,* 34(6): 955-967.

Litten RZ, Wilford BB, Falk DE, Ryan ML, Fertig JB. (2016). Potential medications for the treatment of alcohol use disorder: An evaluation of clinical efficacy and safety. *Subs Abuse,* 37:286–298.

LoCastro JS, Youngblood M, Cisler RA, Mattson ME, Zweben A, Anton RF, Donovan DM. (2009). Alcohol treatment effects on secondary nondrinking outcomes and quality of life: the COMBINE study. *J Stud Alcohol Drugs,* 70(2):186-96.

Manning WG, Mullahy J. (2001). Estimating log models: to transform or not to transform? *J Health Econ,* 20(4):461-94.

Maremmani I Cibin M, Pani PP, Rossi A, Turchetti G. (2015). Harm reduction as "continuum care: in alcohol abuse disorder. *Int J Environ Res Public Health,* 12:14828-14841.

Miller WR. (1996). *Form 90: A structured assessment interview for drinking and related behaviors*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

Miller WR, Tonigan JS, & Longabaugh R. (1995). *The Drinker Inventory of Consequences (DrInC)* . Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

Knox J, Wall M, Witkiewitz K, Kranzler HR, Falk DE, Litten RZ, Mann KF, O'Malley SS, Scodes J, Anton RF, & Hasin DS. Reduction in Non-Abstinent WHO Drinking Risk Levels and Change in Risk for Liver Disease and Positive AUDIT-C Scores: Prospective 3-Year Follow-Up Results in the US General Population. *Alcoholism: Clinical and Experimental Research.* (In-Press)

Probst C, Manthey J, Martinez A. Rehm J. (2015). Alcohol use disorder severity and reported reasons not to seek treatment: A cross-sectional study in European primary care practices. S*ubs Abuse Treat Prev Policy on line.*

Rehm J, Gmel GE, Gmel G, Hasan O, et al. (2017). The relationship between different dimensions of alcohol use and the burden of disease-an update. *Addiction,* 112:968–1001.

Rehm J, Roerecke M. (2013). Reduction of drinking in problem drinkers and all-cause mortality. *Alcohol Alcohol,* 48: 509-13.

Roebuck MC, French MT, McLellan AT. (2003). DATStats: results from 85 studies using the Drug Abuse Treatment Cost Analysis Program (DATCAP). *J Subst Abuse Treat,* 25(1):51–57.

Ruan WJ, Goldstein RB, Chou SP, Smith SM, Saha TD, Pickering RP, Dawson DA, Huang B, Stinson FS, Grant BF. (2008). The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. *Drug Alcohol Depend,* 92:(1-3):27-36.

Ryan ML, Falk DE, Fertig JB, Rendenbach-Mueller B, Katz DA, Tracy KA, Strain EC, Dunn KE, Kampman K, Mahoney E, Ciraulo DA, Sickles-Colaneri L, Ait-Daoud N, Johnson BA, Ransom J, Scott C, Koob GF, Litten RZ. (2017). A phase2, double-blind, placebo-controlled randomized trial assessing the efficacy of ABT-436, a novel V1b receptor antagonist, for alcohol dependence. *Neuropsychopharmacology,* 42:1012–1023.

Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction,* 88(6):791-804.

Shield KD, Parry C, Rehm J. (2013). Chronic diseases and conditions related to alcohol use. *Alcohol Res,* 35:155–171.

Sobell LC, & Sobell MB. (1992). Timeline Follow-Back. In Measuring alcohol consumption: Psychosocial and biochemical methods (pp. 41–72). *Humana Press,* http://doi.org/10.1007/978-1-4612-0357-5_3

Sobell LC, Brown J, Leo GI, Sobell MB. (1996). The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug Alcohol Depend,* 42:49–54.

Sobell LC, Sobell MB, Leo GI, Cancilla A. (1988). Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *Br J Addict,* 83:393–402.

Sobell LC, Agrawal S, Sobell MB, Leo GI, Young LJ, Cunningham JA, Simco ER. (2003). Comparison of a quick drinking screen with the timeline followback for individuals with alcohol problems. *J Stud Alcohol,* 64(6):858–861.

Stewart SH, Latham PK, Miller PM, Randall P, & Anton RF. (2008). Blood pressure reduction during treatment for alcohol dependence: results from the Combining Medications and Behavioral Interventions for Alcoholism (COMBINE) study. *Addiction*, 103(10):1622–1628.

Tramacere I, Negri E, Bagnardi V, Garavello W, Rota M, Scotti L, Islami F, Corrao G, Boffetta P, La Vecchia C. (2010). A meta-analysis of alcohol drinking and oral and pharyngeal cancers. Par 1: Overall results and dose-risk relation. *Oral Oncology,* 46:497–503.

van Amsterdam, van den Brink W. (2013). Reduced-risk drinking as a viable treatment goal in problematic alcohol use and alcohol dependence. *J Psychopharmacol,* 27:987-997.

Ware JE, Jr., Kosinski M, Turner-Bowker DM, Gandek B. (2002). How to Score Version 2 of the SF-12® Health Survey (With a Supplement Documenting Version 1) Lincoln, RI: Quality Metric Incorporated.

Ware JE, Kosinski M, & Keller SD. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care*, 34(3), 220–33.

Witkiewitz K, Hallgren KA, Kranzler HR, Mann KF, Hasin DS, Falk DE, Litten RZ, O'Malley SS, Anton RF. (2017). Clinical validation of reduced alcohol consumption after treatment for alcohol dependence using the World Health Organization risk drinking levels. *Alcohol Clin Exp Res,* 41:179-186.

Witkiewitz K, Kranzler HR, Hallgren KA, O'Malley SS, Falk DE, Litten RZ, Hasin DS, Mann KF, & Anton RF. (in press). Drinking risk level reductions associated with improvements in physical health and quality of life among individuals with alcohol use disorder. *Alcoholism: Clinical and Experimental Research.*

Witkiewitz K. (2013). "Success" following alcohol treatment moving beyond abstinence. *Alcohol Clin Exp Res,* 37(Suppl 1):E9-E13.

Wood AM, Kaptage S, Butterworth AS, Willeit P, Warnakula S, Bolton T, Paige E, Paul DS, et al. (2018). Risk thresholds for alcohol consumption: Combined analysis of individual-participant data for 599,912 current drinkers in 83 prospective studies. *Lancet,* 391:1513–1523.

World Health Organization. (1998). *World Health Organization Quality of Life*. Geneva, Switzerland: World Health Organization.

Zarkin GA, Bray JW, Aldridge A, Mills M, Cisler RA, Couper D, McKay JR, O'Malley S. (2010). The effect of alcohol treatment on social costs of alcohol dependence: results from the COMBINE study. *Med Care,* 48(5)396-401.