HHS Public Access

Author manuscript

Drug Alcohol Depend. Author manuscript; available in PMC 2018 February 01.

Published in final edited form as:

Drug Alcohol Depend. 2017 February 01; 171: 91–96. doi:10.1016/j.drugalcdep.2016.11.015.

The Search for an Elusive Cutoff Remains: Problems of Binary Classification of Heavy Drinking as an Endpoint for Alcohol Clinical Trials

Matthew R. Pearson[*,1], **Adrian J. Bravo**[1], **Megan Kirouac**[1,2], and **Katie Witkiewitz**[1,2]

[1]Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, 2650 Yale Blvd SE MSC 11-6280, Albuquerque, NM 87106, United States

[2]Department of Psychology, University of New Mexico

Abstract

Background—To examine whether a clinically meaningful alcohol consumption cutoff can be created for clinical samples, we used receiver operator characteristic (ROC) curves to derive gender-specific consumption cutoffs that maximized sensitivity and specificity in the prediction of a wide range of negative consequences from drinking.

Methods—We conducted secondary data analyses using data from two large clinical trials targeting alcohol use disorders: Project MATCH (n = 1,726) and COMBINE (n = 1,383).

Results—In both studies, we found that the ideal cutoff for men and women that maximized sensitivity/specificity varied substantially both across different alcohol consumption variables and alcohol consequence outcomes. Further, the levels of sensitivity/specificity were poor across all consequences.

Conclusions—These results fail to provide support for a clinically meaningful alcohol consumption cutoff and suggest that binary classification of levels of alcohol consumption is a poor proxy for maximizing sensitivity/specificity in the prediction of negative consequences from drinking. Future research examining consumption-consequence associations should take advantage of continuous measures of alcohol consumption and alternative approaches for assessing the link between levels of consumption and consequences (e.g., ecological momentary assessment). Clinical researchers should consider focusing more directly on the consequences they aim to reduce instead of relying on consumption as a proxy for more clinically meaningful outcomes.

[*]Corresponding Author: Phone: 505 925 2322, mateo.pearson@gmail.com.

Contributors

MRP wrote the first draft and conducted the analyses. AJB prepared all tables, and assisted with data interpretation and helped write other aspects of the manuscript. MK and KW assisted with manuscript writing and data interpretation. All authors approve of this manuscript and have contributed significantly to the article, and will hold ourselves responsible for its content.

Conflict of Interest

The authors have no conflict of interest.

Keywords

heavy drinking; alcohol use disorders; alcohol treatment outcome; alcohol related consequences; receiver operating characteristics

1. INTRODUCTION

For the past few decades, the use of a cutoff for “binge drinking” (or “heavy episodic drinking”; Jackson, 2010) has proliferated in the college student literature, and has more recently gained serious traction in the broader public health and clinical fields. For example, the National Institute on Alcohol Abuse and Alcoholism (NIAAA, 2004) defines binge drinking as consuming 4+/5+ drinks for women/men within a two-hour period, and uses a 4+/5+ definition for public health messages regarding “low-risk” drinking. Similarly, the Food and Drug Administration (FDA; 2015) recently proposed to include percent of subjects with no heavy drinking days (PSNHDD) as a primary endpoint for alcohol clinical trials, with heavy drinking days defined using the 4+/5+ binge drinking cutoff. Despite its increased use in the field, one can argue there is little evidence supporting the use of this 4+/5+ binge drinking criterion in either college or clinical populations (Pearson et al., 2016a, 2016b).

Many critiques of the binge/heavy drinking criterion have been levied over the years ranging from its limited predictive validity and clinical utility, the perils of falsely dichotomizing a count or continuous variable (i.e., alcohol consumption), and the misguided tendency to use between-subject data to make within-subject inferences (Dejong, 2001; Edward et al., 1994; Pearson et al., 2016a, 2016b; Perkins et al., 2001). In its application to clinical populations, the use of this criterion seems particularly problematic as the cutoff was initially derived from college students and their increased risk of experiencing alcohol-related consequences of most relevance to college students (Wechsler et al., 1995). One area of serious concern is that this cutoff is being applied to clinical populations to identify treatment “failure” (Falk et al., 2010; FDA, 2006). Although Falk et al. (2010) found that individuals who reported any heavy drinking (defined by the 4+/5+ cutpoint) experienced more consequences than individuals with no heavy drinking, one would expect that any cutoff that separates lighter and heavier drinkers would find the heavier drinkers to experience more consequences. Thus, there is little empirical evidence that 4+/5+ drinks per drinking occasion is uniquely indicative or predictive of problem severity or clinical outcomes in alcohol use disorder (AUD) samples.

We argue there are three prerequisites for identifying a clinically meaningful alcohol consumption-based cutoff: 1) assessment of clinically meaningful outcomes, 2) use of clinical samples, and 3) testing of a wide range of potential cutoffs. To this end, we empirically evaluated alcohol consumption-based measures as predictors of drinking consequences with these three prerequisites guiding the present study. First, we considered a wide range of alcohol consequences as clinically meaningful outcomes. Second, we used data collected from clinical samples. Finally, we compared multiple operationalizations of

alcohol consumption to test whether a single cutoff could be identified that maximized sensitivity and specificity in predicting consequence outcomes.

2. METHOD

2.1. Participants and Procedures

We examined clinical populations of patients with alcohol use disorder by using data from Project MATCH (Project MATCH Research Group, 1993) and the COMBINE study (Anton et al., 2006). Project MATCH included 1,726 participants (952 outpatient, 774 aftercare) recruited for a multisite randomized trial examining treatment-matching hypotheses in response to one of three behavioral interventions (cognitive behavioral therapy, motivation enhancement therapy, or twelve-step facilitation) for alcohol dependence (for design details, see Project MATCH Research Group, 1993). In Project MATCH treatment was delivered over three months and assessments were conducted at baseline, 3-months (end of treatment), and 6-, 9-, 12-, and 15-months post-baseline. The COMBINE study included 1,383 participants recruited for a multisite randomized trial examining the effects of pharmacotherapy (naltrexone, acamprosate, or placebo) with or without a combined behavioral intervention for alcohol dependence (for design details, see COMBINE Study Research Group, 2003; Anton et al., 2006). In COMBINE treatment was delivered over four months and assessments were conducted at baseline, 16-weeks (end of treatment), and 6.5-, 12-, and 16-months post-baseline.

2.2. Measures

2.2.1. Alcohol consumption—In both studies, alcohol consumption was assessed using the Form 90 interview (Miller and Del Boca, 1994), which is a calendar-based measure in which participants report the number of standard drinks (defined as 0.5 ounces of pure alcohol) they consumed on each day during the assessment window (e.g., past 90 days). From this measure, we calculated three primary alcohol consumption variables: average drinks per drinking day (DDD), maximum number of drinks consumed during the assessment window (MXD), and average drinks per day (DPD; including non-drinking days in the average). To capture post-treatment drinking, each of these alcohol consumption variables were calculated for a 90-day period ending 6-months following baseline (i.e., 4–6 months post-baseline).

2.2.2. Alcohol Consequences—In both studies, alcohol consequences were assessed using the Drinker Inventory of Consequences (DrInC; Miller et al., 1995), which assesses 45 consequences including physical consequences (“I have had a hangover after drinking,” “While drinking or intoxicated, I have been physical hurt, injured, or burned”), intrapersonal consequences (“I have felt bad about myself because of my drinking,” “My spiritual or moral life has been harmed by my drinking”), social responsibility consequences (“I have failed to do what is expected of my because of my drinking,” “I have been suspended/fired from or left a job or school because of my drinking”), interpersonal consequences (“While drinking, I have said or done embarrassing things,” “I have lost a marriage or a close love relationship because of my drinking”), and impulse control consequences (“I have driven a motor vehicle after having three or more drinks,” “While drinking or intoxicated, I have

injured someone else"). For all items, see Supplemental Table 1[1]. In both Project MATCH and COMBINE, we examined consequences at 12-months post-treatment (15-month post-baseline follow-up in Project MATCH, 16-month post-baseline follow-up in COMBINE). This time point was selected based on its use in the field as a timepoint that is meaningful to clinicians (e.g., Miller and Manuel, 2008).

2.3. Statistical Analyses

In an attempt to examine whether a clinically meaningful cutoff can be created for clinical samples, we used receiver operator characteristic (ROC; Green and Swetz, 1966) curves. We used ROC curves to examine and compare the sensitivity/specificity of various gender-specific alcohol consumption-based cutoffs on DDD, MXD, and DPD. We used each of the 45 consequences assessed by the DrInC as outcomes variables, where each item was scored as absent (scored as 0) or present (scored as 1) to facilitate the requirement of dichotomous outcomes in ROC analyses. For each operationalization of alcohol consumption and for each alcohol consequence, we followed a two-step procedure to find the cutoff that maximized the sensitivity and specificity. In step one, we found the point with the smallest absolute value difference between sensitivity and specificity (i.e., equally weighting the importance of each). Then, we found if sensitivity or specificity continued to improve without worsening the other, and selected the point at which the sum of sensitivity and specificity were maximized. In all cases where there was a "tie" between two points, we selected the point that maximized the sum of sensitivity and specificity. In the present study, sensitivity reflected the degree to which individuals who experienced a consequence fell above a certain alcohol consumption cutoff (i.e., true positive rate), and specificity reflected the degree to which individuals who did not experience a consequence fell below a certain alcohol consumption cutoff (i.e., true negative rate). Values closer to 1 indicate greater sensitivity/specificity. All analyses were conducted prospectively (i.e., alcohol consumption predicting alcohol consequences during a later time period).

3. RESULTS

3.1. Bivariate Correlations and Descriptive Statistics: Project MATCH

The correlations among consumption and consequence variables in Project MATCH are shown in Supplemental Table 2[2]. The alcohol consumption variables were strongly correlated with each other ($.522 < rs < .859$) and modestly correlated with most alcohol consequences ($-.003 < rs < .283$, median $r = .134$). Correlations among the alcohol consequences were mostly in the moderate range (median $r = .294$), including one significant negative correlation ranging to strong positive inter-correlations ($-.182 < rs < .806$). There was a wide range in the proportions of individuals that experienced each alcohol consequence (8% to 86%, median = 57%).

[1]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...
[2]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...

3.2. Bivariate Correlations and Descriptive statistics: COMBINE

The correlations among consumption and consequence variables in COMBINE are shown in Supplemental Table 3[3]. The alcohol consumption variables were strongly correlated with each other (.660 < *r*s < .878) and modestly correlated with most alcohol consequences (.115 < *r*s < .345, median *r* = .258). The correlations among the alcohol consequences ranged from non-significant to strong (.027 < *r*s < .841), with most correlations in the moderate range (median *r* = .391). There was a wide range in the proportions of individuals that experienced each alcohol consequence (4% to 63%, median = 32.5%).

3.3. ROC Curves: Project MATCH

The results for all 45 consequences are available for DDD, MXD, and DPD in Supplemental Table 4[4]. Table 1 provides a summary of ranges and median values for ideal cutoffs as well as the ranges of sensitivity and specificity. For example, the average cutoff for women was 4.38 drinks/day based on average drinks per drinking day (DDD), 7.24 drinks/day based on maximum drinks during the assessment window (MXD), and 0.77 drinks/day based on average drinks per day (which includes non-drinking days as zero; DPD). The sensitivity/specificity for these ranged from .47–.71. For men, these average cutoffs were 5.52, 9.46, and 0.99 drinks/day based on DDD, MXD, and DPD, respectively, with sensitivity/specificity ranges from .50–.66. The drinking cutoff ranges across the different alcohol consequences were also wide, especially for MXD. Overall, these statistics reveal there is great variability in ideal cutoffs both across alcohol consumption variables and alcohol consequences for women and men. Further, these results indicate that none of the alcohol consumption cutoffs result in high sensitivity and high specificity (average sensitivity = .60, average specificity = .60).

3.4. ROC Curves: COMBINE

The results for all 45 consequences are available for DDD, MXD, and DPD in Supplemental Table 5[5]. The average cutoff for women was 3.83 drinks/day based on average drinks per drinking day (DDD), 6.15 drinks/day based on maximum drinks during the assessment window (MXD), and 1.18 drinks/day based on average drinks per day (which includes non-drinking days as zero; DPD). The sensitivity/specificity for these ranged from .42–.92. For men, these average cutoffs were 4.92, 7.74, and 1.06, respectively, based on DDD, MXD, and DPD with sensitivity/specificity ranges from .57–.72. As with Project MATCH data, there was great variability in ideal cutoffs across alcohol consumption variables, alcohol consequences, and genders (see Table 1), and none of the cutoffs demonstrated high sensitivity and specificity (average sensitivity = .64, average specificity = .65).

3.5. Supplemental Analyses

We conducted additional analyses using presence/absence of at least 1 problem on each of the five DrInC subscales and experiencing 0–2 vs. 3+ problems on the DrInC (Anton et al., 2006) as additional clinically meaningful outcomes in Project MATCH (Supplemental Table

[3]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...
[4]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...
[5]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...

4[6]) and COMBINE (Supplemental Table 5[7]). We found similar results as described above. In addition, for these outcomes, we examined both the concurrent associations as well as the prospective associations. As shown in Supplemental Table 6[8], we found higher sensitivity/specificity for concurrent associations (averages ranged from .70 – .80) compared to prospective associations (averages ranged from .61 – .71); however, in either case, great variability was found in ideal cutoffs across alcohol consumption variables, across DrInC subscales, and across gender.

4. DISCUSSION

The current 4+/5+ definition of binge/heavy drinking for women/men is pervasive and has gained support in its application to the clinical context (FDA, 2006). Using ROC curves, the present study sought to examine whether an optimal cutoff could be found for women and men that had high sensitivity and specificity in the prediction of a wide range of negative consequences from drinking. Two findings were clearly demonstrated: 1) the optimal cutoff for maximizing sensitivity/specificity greatly depended on the specific operationalization of alcohol consumption (e.g., DDD, MXD, or DPD) and the specific consequence being examined, and 2) no cutoff provided particularly strong levels of sensitivity and specificity.

Interestingly, across both samples, the average optimal cutoff for women was close to 4 (4.38 in MATCH and 3.83 in COMBINE) drinks and the average optimal cutoff for men was close to 5 (5.52 in MATCH and 4.92 in COMBINE) drinks when using average drinks per drinking day, thus providing some support for the 4+/5+ definition when using drinks per drinking day. This finding is curious considering that self-reported average drinks per drinking day was the measure originally used to support the 4/5+ binge/heavy drinking criterion among college students (Wechsler et al., 1995). However, the current focus on “no heavy drinking” (pg. 7; FDA, 2015) would suggest that the *maximum* number of drinks (i.e., the MXD variable in the current analyses) during an assessment window would be a more appropriate measure upon which to develop a cutoff for no heavy drinking. In the present study, the average and range of ideal cutoffs based on maximum drinks during an assessment period were substantially higher than the 4/5+ cutoff and again leads us to question whether the 4/5+ cutoff is a useful cutoff for clinical populations (Pearson et al., 2016a, 2016b; Wilson, Bravo, Pearson, and Witkiewitz, 2016).

We see the use of no heavy drinking based on the 4+/5+ definition of heavy drinking for evaluating clinical trials as a step in the right direction away from fully relying on abstinence-based outcomes. However, we did not find support for a single alcohol consumption cutoff that maximizes both sensitivity and specificity in the prediction of consequences among clinical populations of individuals who received treatment for alcohol use disorder. Therefore, if a treatment targets the reduction of alcohol consumption, we advocate that continuous or count measures of alcohol consumption be used as the outcome measures rather than falsely dichotomizing these measures. Dichotomization of continuous or count variables makes the untenable assumption that those in one group (i.e., non-heavy

[6]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...
[7]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...
[8]Supplementary material can be found by accessing the online version of this paper at http://dx.doi.org and by entering doi:...

drinkers) are qualitatively different from those in the other group (i.e., heavy drinkers). DeCoster and colleagues (2009) demonstrated that continuous variables in statistical models rarely perform worse than dichotomous variables, but dichotomous variables usually perform worse than continuous variables. Dichotomization of a continuous measure can result in a loss of statistical power, making statistical tests less sensitive to detecting existing effects (higher chance of a type II error), which has serious implications when these measures are used to quantify treatment effects.

Given the modest associations between post-treatment drinking and alcohol consequences at follow-up, we must also conclude that if a treatment targets the reduction of alcohol consequences, researchers ought to focus on alcohol consequences as their outcome. It seems that alcohol consumption is often used as a proxy for the risk of experiencing negative consequences from drinking; however, the present findings demonstrate that alcohol consumption may be a poor proxy for alcohol consequences (as indicated by small correlations between consumption and consequence items). Given the growing acceptance of harm reduction approaches (Davis and Rosenberg, 2012; Maremmani et al., 2015), it is quite conceivable that a reduction of alcohol consequences can occur in the absence of changing level of alcohol consumption through improving nutrition, improving hydration, reducing environmental risk (i.e., not drinking in risky contexts), slowing consumption rate, and using alcohol protective behavioral strategies (e.g., “Using a designated driver”; Pearson, 2013).

According to the Diagnostic and Statistical Manual for Psychiatric Disorders - fifth edition (DSM-5; American Psychiatric Association, 2013), alcohol use disorder (AUD) is diagnosed based on the endorsement of two or more of 11 criteria during the same 12-month period. Despite the fact that level of alcohol consumption is not one of the criteria for AUD, alcohol consumption outcomes are the primary outcomes used to test the efficacy of AUD treatments. We know of no other psychiatric disorders where the primary outcomes of its treatments are *not* the very symptoms that define the disorder. If we define AUD diagnostic criteria using alcohol-related consequences, why not define and evaluate AUD treatment success by alcohol-related consequences instead of consumption?

4.1. Study Limitations and Future Directions

It terms of identifying an alcohol consumption cutoff that is associated with experiencing negative consequences, there are limitations to the present study. First, we relied on retrospective self-report measures of alcohol consumption and consequences assessed over a relatively long period of time (e.g., past 90 days). Second, we set the optimal cutoff by maximizing sensitivity and specificity, but there are often tradeoffs between sensitivity and specificity and there are some situations when it may be better to maximize sensitivity over specificity, and *vice versa*. For example, if one wanted to define treatment success by identifying an optimal cutoff that was most likely to reflect the degree to which individuals who did not experience a consequence fell below a certain alcohol consumption cutoff (i.e., true negative rate), then one would want to maximize specificity. Conversely, if one wanted to define treatment failure by identifying an optimal cutoff that was most likely to reflect the degree to which individuals who experienced a consequence fell above a certain alcohol consumption cutoff (i.e., true positive rate), then one would want to maximize sensitivity.

Future research could examine cutoffs that maximize either sensitivity or specificity by setting a minimal allowable cutoff of either, or determining cutoffs using more sophisticated criteria like the Youdon index (Perkins and Schisterman, 2006).

The present study focused on finding an ideal alcohol consumption-based cutoff based on predictive validity. Alternatively, this could be attempted based on concurrent validity. However, given findings that suggest drinking patterns are potentially less stable and relapse rates are high early in the post-treatment window (e.g., Hunt et al., 1971), and given that clinicians are concerned with longitudinal outcomes (e.g., Miller and Manuel, 2008), we felt examination of predictive validity was the most appropriate way to address the present research question. Given the substantively null findings in the present study, future studies may consider using a concurrent approach to identify consumption-based cutoffs.

An additional limitation was the lack of information about the duration of alcohol consumption during a drinking episode; therefore, we are unable to distinguish between consuming 5 standard drinks in a 5-hour period from consuming 5 standard drinks in a 1-hour period. The use of ecological momentary assessment (EMA; Shiffman, 2009) has several advantages over the current methodology, including its ability to reduce recall biases, measure time of consumption in near real-time, and examine proximal associations between alcohol consumption and alcohol consequences (e.g., last night drinking predicting next day hangover symptoms). Use of EMA data holds promise for identifying alcohol consumption levels that are predictive of alcohol consequences at the individual level.

4.2. Conclusions

Despite the limitations of the present study, our results add to the growing evidence that questions the unique predictive validity of the 4+/5+ binge/heavy drinking criteria (Pearson et al., 2016a, 2016b; Wilson et al., 2016; Witkiewitz, 2013). Importantly, the present study results provide evidence that using percent subjects with no heavy drinking days (PSNHDD) relying upon 4/5+ as the maximum number of drinks per occasion in a 90-day assessment window is an excessively restrictive outcome for clinical populations. Rather than further reifying this definition based on tradition, the field should consider moving forward based on scientific evidence, which has largely failed to find support for 4+/5+ as an optimal predictor of clinically meaningful treatment outcomes over other cutoffs or predictive variables (e.g., temptation to drink; Witkiewitz, 2013). The myopic focus on alcohol consumption as a treatment target, rather than negative consequences from drinking, may have steered the field away from potentially impactful treatments and towards less impactful treatments. Although alcohol consumption may be a necessary condition for alcohol consequences, it is far from a sufficient condition and the alcohol treatment field may benefit from focusing more specifically on the actual outcomes of most interest rather than relying on alcohol consumption as a proxy for these clinically meaningful outcomes (Donovan et al., 2012).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of Funding

This research was supported by a grant (R01-AA022328; PI: Witkiewitz) from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). MRP is supported by a career development grant (K01-AA023233) from the NIAAA. AJB is supported by a training grant (T32-AA018108) by the NIAAA. MK is supported by a predoctoral training fellowship (F31-AA024559) from the NIAAA.

References

Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D, Donovan DM, Gastfriend DR, Hosking JD, Johnson BA, LoCastro JS, Longabaugh R, Mason BJ, Mattson ME, Miller WR, Pettinati HM, Randall CL, Swift R, Weiss RD, Williams LD, Zweben A. Combined pharmacotherapies and behavioral interventions for alcohol dependence. JAMA. 2006; 295:2003–2017. DOI: 10.1001/jama.295.17.2075 [PubMed: 16670409]

Davis AK, Rosenberg H. Acceptance of non-abstinence goals by addiction professionals in the United States. Psychol Addict Behav. 2013; 27:1102–1109. DOI: 10.1037/a0030563 [PubMed: 23088408]

Dejong W. Finding common ground for effective campus-based prevention. Psychol Addict Behav. 2001; 15:292–296. DOI: 10.1037/0893-164X.15.4.292 [PubMed: 11767259]

Donovan DM, Bigelow GE, Brigham GS, Carroll KM, Cohen AJ, Gardin JG, Hamilton JA, Huestis MA, Hughes JR, Lindblad R, Marlatt GA, Preston KL, Selzer JA, Somoza EC, Wakim PG, Wells EA. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. Addiction. 2012; 107:694–708. DOI: 10.1111/j.1360-0443.2011.03473.x [PubMed: 21781202]

Edwards, G., Anderson, P., Babor, TF., Casswell, S., Ferrence, R., Glesbrecht, N., Godfrey, C., Holder, HD., Lemmens, PHMM., Makela, K. Alcohol Policy And The Public Good. Oxford University Press; New York, NY: 1994.

Falk D, Wang XQ, Liu L, Fertig J, Mattson M, Ryan M, Johnson B, Stout R, Litten RZ. Percentage of subjects with no heavy drinking days: Evaluation as an efficacy endpoint for alcohol clinical trials. Alcohol Clin Exp Res. 2010; 34:2022–2034. DOI: 10.1111/j.1530-0277.2010.01290.x [PubMed: 20659066]

Food and Drug Administration. Medical Review Of Vivitrol: 21-897. Rockville, MD: 2006.

Food and Drug Administration. Alcoholism: Developing Drugs for Treatment. Rockville, MD: 2015. (No. PDA D-0152-001)

Green, DM., Swetz, JA. Signal Detection Theory And Psychophysics. John Wiley and Sons Inc; New York, NY: 1966.

Hunt WA, Barnett W, Branch LG. Relapse rates in addiction programs. J Clin Psychol. 1971; 27:455–456. [PubMed: 5115648]

Jackson KM. Heavy episodic drinking: Determining the predictive utility of five or more drinks Psychol. Addict Behav. 2010; 22:68–77. DOI: 10.1037/0893-164X.22.1.68.Heavy

Maremmani I, Cibin M, Pani PP, Rossi A, Turchetti G. Harm reduction as "continuum care" in alcohol abuse disorder. Int J Environ Res Public Health. 2015; 12:14828–14841. DOI: 10.3390/ijerph121114828 [PubMed: 26610535]

Miller WR, Del Boca FK. Measurement of drinking behavior using the Form 90 family of instruments. J Stud Alcohol Suppl. 1994; 12:112–118. [PubMed: 7722987]

Miller WR, Manuel JK. How large must a treatment effect be before it matters to practitioners? An estimation method and demonstration. Drug Alcohohl Rev. 2008; 27:524–528. DOI: 10.1080/09595230801956165

Miller, WR., Tonigan, JS., Longabaugh, R. The Drinker Inventory Of Consequences (DrInC), The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse. National Institute on Alcohol Abuse and Alcoholism; Bethesda: 1995.

National Institute on Alcohol Abuse and Alcoholism. NIAAA Council approves definition of binge drinking. NIAAA Newsletter. 2004:2–3.

Pearson MR. Use of protective behavioral strategies among college students: a critical review. Clin Psychol Rev. 2013; 33:1025–1040. [PubMed: 24036089]

Pearson MR, Kirouac M, Witkiewitz KA. Questioning the validity of the 4+/5+ binge or heavy drinking criterion in college and clinical populations. Addiction. 2016a epub ahead of print.

Pearson MR, Kirouac M, Witkiewitz KA. We still question the utility and validity of the binge/heavy drinking criterion. Addiction. 2016b epub ahead of rpint.

Perkins HW, Linkenbach J, Dejong W. Estimated blood alcohol levels reached by “binge” and “nonbinge” drinkers: a survey of young adults in Montana. Psychol Addict Behav. 2001; 15:317–320. DOI: 10.1037/0893-164X.15.4.317 [PubMed: 11767263]

Perkins NJ, Schisterman EF. The inconsistency of “optimal” cut-points using two ROC based criteria. Am J Epidemiol. 2006; 163:670–675. [PubMed: 16410346]

Project MATCH Research Group. Project MATCH: Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment. Alcohol Clin Exp Res. 1993; 17:1130–1145. [PubMed: 8116822]

Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess. 2009; 21:486–97. DOI: 10.1037/a0017074 [PubMed: 19947783]

The COMBINE Study Research Group. Testing Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (The COMBINE Study): a pilot feasibility study. Alcohol Clin Exp Res. 2003; 27:1123–1131. DOI: 10.1097/01.ALC.0000078020.92938.0B [PubMed: 12878918]

Wechsler H, Dowdall GW, Davenport A. A gender-specific measure of binge drinking among college students. Am J Public Health. 1995; 85:5–8.

Wilson AD, Bravo AJ, Pearson MR, Witkiewitz K. Finding success in failure: using latent profile analysis to examine heterogeneity in psychosocial functioning among heavy drinkers following treatment. Addiction. 2016; doi: 10.1111/add.13518

Witkiewitz K. “Success” following alcohol treatment: moving beyond abstinence. Alcohol Clin Exp Res. 2013; 37(Suppl 1):E9–13. DOI: 10.1111/acer.12001 [PubMed: 23075307]

Highlights

- We attempted to create a clinically meaningful alcohol use cutoff
- We conducted secondary data analyses using Project MATCH and COMBINE data
- The results fail to provide support for a clinically meaningful alcohol use cutoff
- Across a range of consequences, no cutoff resulted in high sensitivity/specificity
- No specific level of alcohol use is a good proxy for alcohol consequences

Table 1

Average, median and range of cutoffs based on maximizing sensitivity and specificity

Project MATCH															
	Cut-Offs					**Sensitivity**					**Specificity**				
Women	***M***	**Med**	***SD***	**Min**	**Max**	***M***	**Med**	***SD***	**Min**	**Max**	***M***	**Med**	***SD***	**Min**	**Max**
DDD	4.38	4.25	0.52	3.45	5.65	0.59	0.60	0.04	0.47	0.66	0.60	0.60	0.04	0.54	0.67
MXD	7.24	7.06	1.14	5.11	10.19	0.61	0.61	0.04	0.47	0.68	0.61	0.61	0.04	0.52	0.71
DPD	0.77	0.71	0.24	0.34	1.29	0.61	0.62	0.05	0.47	0.70	0.62	0.63	0.04	0.55	0.70
	Cut-Offs					Sensitivity					Specificity				
Men	*M*	Med	*SD*	*M*	Med	*M*	Med	*SD*	Min	Max	*M*	Med	*SD*	Min	Max
DDD	5.52	5.37	0.58	4.57	6.98	0.61	0.61	0.03	0.54	0.65	0.61	0.61	0.03	0.54	0.64
MXD	9.46	9.05	1.21	7.74	11.62	0.62	0.62	0.03	0.54	0.66	0.62	0.62	0.03	0.55	0.66
DPD	0.99	0.93	0.22	0.59	1.38	0.60	0.60	0.04	0.50	0.65	0.60	0.60	0.03	0.49	0.65
COMBINE															
	Cut-Offs					Sensitivity					Specificity				
Women	*M*	Med	*SD*	Min	Max	*M*	Med	*SD*	Min	Max	*M*	Med	*SD*	Min	Max
DDD	3.83	3.62	0.83	2.81	7.11	0.64	0.65	0.04	0.50	0.69	0.65	0.63	0.05	0.56	0.84
MXD	6.15	5.88	1.14	5.11	11.92	0.62	0.64	0.05	0.42	0.71	0.65	0.65	0.05	0.56	0.87
DPD	1.18	1.08	0.72	0.56	5.47	0.63	0.64	0.05	0.50	0.72	0.65	0.65	0.06	0.56	0.92
	Cut-Offs					Sensitivity					Specificity				
Men	*M*	Med	*SD*	Min	Max	*M*	Med	*SD*	Min	Max	*M*	Med	*SD*	Min	Max
DDD	4.92	4.79	0.63	3.48	6.32	0.65	0.66	0.03	0.60	0.71	0.66	0.66	0.03	0.59	0.71
MXD	7.74	7.51	1.31	5.45	10.52	0.67	0.67	0.03	0.57	0.71	0.67	0.68	0.03	0.57	0.71
DPD	1.06	0.98	0.34	0.53	2.33	0.66	0.66	0.03	0.59	0.72	0.66	0.66	0.03	0.59	0.72

Note. DDD = average drinks per drinking day, MXD = maximum number of drinks consumed during the assessment window, DPD = average drinks per day.