Testing different thresholds for risky episodic drinking – what's so special about five drinks?

Author/Contributor:

Livingston, Michael

Publication details:

Drug and Alcohol Dependence

v. 133

Chapter No. 1

pp. 248-253

Publication Date:

2013

Publisher DOI:

<http://dx.doi.org/10.1016/j.drugalcdep.2013.04.027>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/53415> in <https://unsworks.unsw.edu.au> on 2023-04-26

Testing different thresholds for risky episodic drinking – what’s so special about five drinks?

Michael Livingston^{1,2}

1. Drug Policy Modelling Program, National Drug and Alcohol Research Centre, University of New South Wales
2. Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre

Word count: 3536

Abstract word count: 232

Figures: 3

Tables: 4

Pages: 18

Competing interests: None.

Abstract

Background: Studies of episodic drinking typically use a measure based on the frequency of drinking five or more standard drinks (a definition which itself varies based on the standard units being used). While this threshold clearly defines drinking behaviour with a range of risks and negative consequences, there has been limited research outside of US college-based studies to determine the appropriateness of this definition. This study examines fifteen different risky-drinking thresholds to assess which definitions of risky drinking best predict negative outcomes. **Methods:** This paper presents an analysis of a national survey sample of 19,757 drinkers. The appropriateness of each threshold is assessed using basic risk-curves, specificity and sensitivity analyses and the performance of each threshold definition in multivariate logistic regression models. Risky drinking was defined in fifteen ways (based on frequency and volume) and tested against a series of self-reported negative outcomes and risky behaviours. **Results:** The study finds that the most appropriate risky drinking threshold for these data varies based on the mode of analysis and on the type of outcome being considered. Across all approaches used, risky drinking thresholds of seven or fewer drinks performed better than higher thresholds. **Conclusions:** While individual level risks peak at higher levels of consumption, these findings support the continuing use of relatively low thresholds for defining risky-drinking, as risk across the total population is highest at these levels.

Key words: Alcohol, drinking guidelines, risky drinking

Acknowledgments

This work was supported by the Foundation for Alcohol Research and Education, an independent, charitable organisation working to prevent the harmful use of alcohol in Australia www.fare.org.au. Dr Livingston is supported by an NHMRC Early Career Fellowship. The funders had no role in study design, analysis or interpretation or in the decision to submit this paper for publication. The Australian Institute of Health and Welfare manage the data collection and dissemination of the National Drug Strategy Household Survey and I am grateful to them for facilitating access to the data.

Introduction

Alcohol consumption contributes to a substantial public health burden in Australia, with the most recent estimates available attributing 3.2% of the burden of disease to risky drinking (Begg et al., 2007). Given this burden, there is significant research interest in examining the characteristics, behaviours and negative consequences experienced by people who drink at 'risky' levels. Typically, the definition of 'risky drinking' in Australian research has been based on guidelines produced by the National Health and Medical Research Council (NHMRC), which specify thresholds for both episodic consumption and longer-term drinking. In other words, the NHMRC guidelines conceptualise risky drinking of two types, based on drinking patterns and total volume of consumption. There is increasing evidence that drinking pattern (i.e., risky episodes of drinking) is a key driver of negative health and social consequences from alcohol consumption (Bobak et al., 2004; Gmel et al., 2007; Rehm et al., 2001). Similarly, the focus of much media and policy attention in the alcohol field is on 'binge' or episodic heavy drinking (Gartrell and Veness, 2008; Stark, 2007).

Given this focus, developing a robust definition of episodic risky drinking is a key requirement. Under the Australian 2009 NHMRC guidelines, short-term risky drinking is defined as an episode of drinking where five or more standard drinks¹ are consumed. This threshold was derived from meta-analyses of data from emergency department studies examining the role of alcohol consumption in injury morbidity and mortality. The authors of the guidelines picked an absolute risk threshold of 1 in 100, which was met by drinkers who consumed 5 or more drinks twice-weekly. While there was some evidence that mortality risk increased more rapidly above the 5+ threshold, the use of the 1 in 100 risk level was arbitrary, and based on acceptable risks in other settings (National Health and Medical Research Council, 2009). This definition is widely used in alcohol research in Australia to define 'risky drinkers' (AIHW, 2011). There has been some criticism of this threshold as being overly inclusive. For example, using this threshold, 45% of Australians aged 18 and over are classified as short-term risky drinkers (AIHW, 2011). Further, this threshold doesn't capture extremely heavy drinking, which may be increasing in Australia. For example, a recent Victorian study found that more than 40% of 16-24 year olds reported drinking episodes of more than 20 standard drinks, up from 26% in 2002 (Victorian Drug and Alcohol Prevention Council, 2010).

In US studies of risky drinking a threshold equivalent to seven Australian standard drinks (5 US drinks) is more widely used (e.g. Wechsler et al., 1995; Weitzman et al., 2003). Again, this threshold has been criticised as being too low, particularly for college populations (White et al., 2006). In response to these criticisms, a series of studies have been undertaken to assess the validity of the US threshold. Using survey data, both Wechsler et al. and Weitzman et al. find no evidence that a higher threshold for defining episodic risky drinking is more valid, with increases in the threshold resulting in more accurate prediction of harm rates amongst the 'risky drinking' group, but also higher rates of false negatives (i.e. non-risky drinkers who report harm) (Wechsler and Nelson, 2006; Weitzman and Nelson, 2004). Framing their results in terms of the prevention paradox, Weitzman et al. argue that the use of a lower risky-drinking threshold is the most likely to reduce alcohol-related harms as, while prevalence of harm is higher among the heaviest drinking young people, the number of lower-level drinkers is so much higher that they account for the majority of harms experienced (Weitzman

¹ Note that an Australian standard drink is 10g of alcohol, while US standard drinks are ~14g.

and Nelson, 2004). Other studies have broadly supported the use of risky-drinking thresholds at or below 7 Australian drinks. Using event-based data, Jackson et al. found little evidence that higher thresholds had better predictive value (Jackson, 2008), while Dawson et al., examined a range of potential drinking guidelines, finding that a risky-drinking threshold of 4 US drinks (~6 Australian standard drinks) performed the best (Dawson et al., 2012). Outside of the US, analyses of Finnish drinkers examined the contribution of drinking occasions across a range of quantities and frequencies, finding that individual-level risk peaked at high levels of consumption (above thirteen drinks²) but that the highest levels of risk at the population level occurred at much lower levels (between five and seven drinks) (Mäkelä and Mustonen, 2007).

Thus, there is a reasonable evidence base from US studies of college populations that using episodic risky drinking thresholds higher than 5-7 drinks provides little additional benefit. In addition, there is evidence from Finland that, while individual level risk is highest for people who drink at very high levels, population level risk is highest at much lower levels. Aside from this Finnish study, there has been little attempt to validate these findings either on broader populations (i.e. non-college samples) or in a non-US context. This study uses Australian survey data to examine the predictive utility of different definitions of short-term risky drinking on a range of alcohol-related harms.

It is worth highlighting the broad and disparate approaches that make up the research literature in this field: researchers have used a wide variety of harm measures, have studied differing sets of potential thresholds (both in terms of volumes and frequencies) and have not used consistent timeframes for analysis. Further, there has been little theoretical discussion of the balancing between sensitivity and specificity required for risky drinking measures. This will depend on the purpose of the measure, and with broad-based population approaches like those discussed here, there is no obvious argument to favour sensitivity over specificity (or vice versa). This study takes a broad approach, aiming to assess the validity of as broad a range of potential thresholds as possible, over a range of different harm measures, while balancing sensitivity and specificity.

Methods

Data

This study uses data from the 2010 National Drug Strategy Household Survey (NDSHS). The NDSHS is a national survey of the Australian population aged 12 and over. The data were collected using a drop and collect approach and a clustered stratified random sample with regions sampled randomly within 15 strata (capital city and rest of state for each state and territory of Australia, except the Australian Capital Territory, which was a single strata), households selected randomly within regions and respondents randomly selected within households. Data were collected from a final sample of 26,648 respondents (based on a participation rate of 50.6%). The survey collects a wide range of data on alcohol, tobacco and other drug use, along with items on consequences of and attitudes to alcohol and drug use, health-related items and a suite of socio-demographic measures. Full details of the methods, the questionnaire and the broad findings of the 2010 NDSHS are available in the main survey report (AIHW, 2011).

² Finnish standard drinks, which include 12g of pure alcohol.

The analyses presented here were based only on respondents who had consumed at least one alcohol drink in the last 12 months and who had provided complete answers to the detailed graduated frequency items relating to alcohol consumption at varying levels. This led to the exclusion of 5,099 (19.1%) respondents who had not consumed alcohol in the last 12 months and 1,752 respondents (6.7%) who provided insufficiently detailed or inconsistent consumption data, leaving a final sample of 19,757 (74.1% of the original sample). All analyses are based on unweighted data for this group of respondents.

The main aim of this study was to compare the utility of a range of different threshold measures for episodic risky drinking. Thus, a series of thresholds were set based on frequency and quantities of consumption, derived from the standard graduated quantity-frequency questions used in the survey (Greenfield, 2000). The quantity thresholds (linked to the specific quantities asked about in the survey) were 3, 5, 7, 11 and 20 Australian standard drinks (10g alcohol), with three frequencies examined for each (at least once in the last year, at least 12 times in the last year and at least 52 times in the last year). Table 1 provides the prevalence of drinking at these various thresholds in this sample (i.e. among drinkers).

TABLE 1 ABOUT HERE

These thresholds will be presented in the rest of this paper as 'quantity/frequency'. For example, the threshold based on drinking 7 or more drinks on 12 or more occasions will be presented as '7+ /monthly', while 5 or more on 52 or more occasions will be presented as 5+ /weekly.

Using each of these thresholds, respondents were classified as either risky-drinkers or non-risky drinkers based on their answers to the graduated quantity-frequency questions (Gmel et al., 2006) in the NDSHS. Thus, for example, a respondent who reported drinking fifteen drinking occasions in the last year and consumed six drinks on each of these occasions would be classified as a risky-drinker using the 3+ /yearly, 3+ /monthly, 5+ /yearly and 5+ /monthly thresholds, but as a non-risky-drinker using all the other thresholds. The utility of these fifteen thresholds was tested by examining how useful they were for predicting the experience of alcohol-related harm. Three different measures of harm were used, again based on the questions available in the NDSHS.

The first measure of harm was based on a single item, "Have you, or someone else, been injured because of your drinking?", with respondents whose drinking had resulted in an injury in the last 12 months coded as 'Yes' and respondents who reported no injuries or injuries more than 12 months ago coded as 'No'. There were 205 respondents who refused to answer this question who were excluded from the analyses based on it.

The remaining harm measures were based on a ten-item scale, relating to behaviours and consequences related to alcohol in the 12 months prior to the survey. Previous analyses have demonstrated that these items cluster into two groups, described as hazardous behaviours and delinquent behaviours (Livingston and Room, 2009). The ten items and their classification into the two broad groups are presented in Table 2. Respondents were classified as having engaged in hazardous behaviours if they had engaged in any of the five relevant items (prevalence = 19.2%). Similarly, any positive response to the delinquent behaviour question meant that a respondent was classified as having engaged in delinquent behaviour (prevalence = 6.3%). Respondents who did not

respond to any of the particular items (n=830 for delinquent behaviour, n=668 for hazardous behaviour) were excluded.

TABLE 2 ABOUT HERE

Analysis

Analyses for this study were descriptive, with basic measures of sensitivity, specificity and model performance compared across thresholds. These measures cannot strictly be compared statistically, raising the possibility that the thresholds identified are artefacts of sampling rather than concrete findings. For this reason, analyses proceeded in three stages, to provide a broad picture of the relationship between drinking at various threshold levels and negative outcomes.

Firstly, simple descriptive tables were run to calculate the risk functions of each of the three harms being examined across each of the thresholds. These risk functions were then plotted to provide a rudimentary means of assessing whether any particular threshold represented a key point beyond which risk increased rapidly. Secondly, these simple descriptive tables were used to calculate straightforward sensitivity and specificity values for each of the three harms (across all thresholds). An ideal threshold would have 100% sensitivity and 100% specificity – that is, it would perfectly discriminate between drinkers who experience harm and those that do not. To assess which threshold provided the best balance of sensitivity and specificity, a Receiver Operator Characteristic (ROC) curve was developed for each outcome, summarising the sensitivity and specificity of each of the fifteen potential thresholds. For each outcome, the best performing threshold was the point at which the ROC curve was closest to a perfectly discriminating test. This was done by calculating the distance between each threshold and the point (0,1) on the ROC plot, with smaller distances indicating better performing thresholds (Beck and Shultz, 1986).

In the second stage of analyses, logistic regression models were developed to determine the predictive utility of each of the fifteen thresholds being examined. Separate models were run for each threshold and harm with age, sex and marital status included as control variables. Three goodness-of-fit measures were calculated for each of these fifteen models: the Bayesian Information Criteria (BIC), the McFadden pseudo R-squared and the Pearson Chi-Squared. The BIC is a model selection metric based on the likelihood function and is lower for a better-fitting model (Schwarz, 1978). The McFadden pseudo R-squared for logistic regression is a rough measure of the proportion of variation in the data explained by the model, and is higher for a better-fitting model (Veall and Zimmermann, 1996). Finally, the Pearson Chi-Squared statistic is based on the sum of the difference between observed and expected outcome and is lower for a better performing model (Bluman, 1992). Values of each of these measures were compared across the fifteen models to assess which threshold values most appropriately fitted the data in this sample when predicting each type of harm.

Results

The relationships between risky-drinking threshold and the risk of experiencing each of the three harms are presented in Figures 1-3. Across each harm-type, the risk functions increase steadily as the risky-drinking threshold is raised, and are higher when frequency is increased. There is some indication, particularly for delinquent and hazardous behaviours (Figures 2 and 3) that the increase

in risk is not straightforwardly linear, with the increase in risk slowing as the risky-drinking threshold is raised. For a risk curve to provide an obvious threshold the reverse relationship would need to be true – that is, there would need to be an ideal risk threshold above which risk increases more quickly than below. Thus, these exploratory charts do not point to a clearly defined risky-drinking threshold.

FIGURES 1-3 ABOUT HERE

The sensitivity and specificity of each of the proposed thresholds are presented in Table 3. The sensitivity of each threshold is defined as the proportion of respondents who experienced each outcome who were classified as ‘risky-drinkers’ based on each threshold (true positives). The sensitivity of each threshold is defined as the proportion of respondents who did not experience each harm who were not classified as ‘risky-drinkers’ by the threshold (true negatives). As discussed earlier, ROC curves were produced to compare thresholds, and the distances between each threshold and the perfect test are provided in Table 3 (smaller values imply a better balance of specificity and sensitivity).

TABLE 3 ABOUT HERE

This analysis suggests that the thresholds that best balance sensitivity and specificity vary by harm type. A risky-drinking threshold of 7 or more drinks at least 12 times per year performs the best on this sample for alcohol-related injuries, while 5+/monthly was observed as most appropriate for hazardous behaviour. For delinquent behaviour, two different thresholds performed equally well: 7+/yearly and 7+/monthly. Separate analyses were run for males and females, with higher thresholds observed for males than those for females. For injury, the male threshold that produced the best balance of sensitivity and specificity was 11+/yearly, while for females it was 5+/monthly. The differences were smaller for delinquent (males 7+/monthly, females 5+/monthly) and hazardous (males 7+/yearly, females 5+/yearly) behaviour. These analyses make no attempt to deal with the variation in drinking patterns and underlying harm rates that occur across sub-groups of the population. Thus, the subsequent analyses use multivariate methods to assess risky drinking thresholds while controlling for age, gender and marital status.

The regression outputs are not reproduced here; instead Table 4 presents the goodness-of-fit measures for each of the 45 (3 outcomes x 15 thresholds) models to compare the appropriateness of the risky-drinking thresholds for each outcome. The three measures (described earlier) used were: the BIC (lower = better performing model), the pseudo R-squared (higher = better performing model) and the Pearson’s Chi-squared (lower = better performing model).

The results of the multivariate analyses have a similar pattern to the more straightforward results presented earlier – thresholds that best fit these data for the injury and delinquent behaviour harms are slightly higher than those for hazardous behaviour. The threshold that produces the best goodness-of-fit measures for the injury outcome in the multivariate models is 7+/monthly for two of the measures and 7+/yearly for the other. The thresholds identified for hazardous behaviour are much lower – 3+/monthly using two measures and 5+/yearly using the other. Delinquent behaviour is somewhere in between, with a threshold of 5+/monthly using two of the measures and 3+/weekly using the other. Again, gender specific models were generated and compared, with the full set of results not reproduced here. As in the previous analyses, the higher thresholds performed better for

males than for females, although the magnitude of the differences between males and females varied substantially across outcome type and goodness-of-fit measure.

TABLE 4 ABOUT HERE

Discussion

The overall picture provided by both the multivariate and bivariate analyses is that the best performing risky drinking threshold will vary considerably depending on the outcome of interest. For the most serious outcome examined here (injury), the thresholds identified were generally higher than those for the less serious outcomes. Even within a single outcome measure, the threshold identified varied based on the analytical approach used, with differing results between bivariate and multivariate approaches and even within the same multivariate approach, depending on the measure used to assess model fit.

Overall, it is clear that, for these particular outcome measures in this population survey, risky drinking thresholds set at around 5-7 Australian standard drinks (3.5 – 5 U.S. standard drinks) are likely to provide the optimal balance between sensitivity and specificity. While people who drink at higher levels clearly put themselves at greater risk of alcohol-related harms, setting thresholds at 11 or 20 standard drinks results in large numbers of people not classified as risky drinkers who experience harm from alcohol. These findings fit with the previous Finnish analyses, which found individual level risk peaking at high thresholds, but overall population risk peaking at levels similar to those found here (Mäkelä and Mustonen, 2007). These findings are also broadly similar to the previous studies in the US, which were based largely on college student populations and typically found that a risky drinking threshold of 5 US drinks (7 Australian drinks) provides the appropriate balance between sensitivity and specificity (Wechsler and Nelson, 2006; Weitzman and Nelson, 2004).

Given the results of this study, there is no obvious need to move away from using the current NHMRC guidelines (5 or more drinks in a session = risky-drinking) to assess episodic risky drinking in Australia. The use of this threshold has the advantage of being consistent with national health advice and with much of the existing research in the field, and is clearly a reasonable level use based on the findings presented here. However, it's worth noting that the ideal threshold for 'risky drinking' will vary substantially depending on the particular 'risk' being examined. As in this study, where the most serious outcome (injury) was best modelled using a higher threshold, previous studies have found that higher thresholds are more appropriate for more severe outcomes (Jackson, 2008; Wechsler and Nelson, 2006).

Gender-specific analyses point to some differences in identified thresholds between males and females, with higher thresholds generally performing producing better goodness-of-fit statistics for males. These findings contrast with earlier analyses of National Drug Strategy Household Survey data (Livingston and Room, 2009) and provide some support to the setting of gender-specific drinking guidelines with higher thresholds for male drinkers, as seen in many other jurisdictions (National Health and Medical Research Council, 2009, p19). Further analyses could explore whether or not best-fitting thresholds vary by age or other demographic characteristics.

This study was limited by the range of specific negative outcomes examined, which were restricted to those that were included in the NDSHS and thus incorporate a fairly narrow range of negative outcomes and behaviours. Future work using a broader range of alcohol-related consequences would provide a better indication of how risky-drinking thresholds vary across types of harm.

The study was also limited by its use of survey data, based on a 50% cooperation rate and covering only 50% of the alcohol actually consumed in Australia. These limitations are common to most studies of these issues and it is worth considering how these findings compare to risky-drinking thresholds based on other methodologies. For example, a meta-analysis of emergency room studies found some evidence of non-linear risk functions, with risk increasing more rapidly above 60g-70g of alcohol (6-7 standard drinks) (Taylor et al., 2010), although these studies are based on drinking immediately prior to the injury and may over-estimate risk levels for methodological reasons (Gmel and Daepfen, 2007). The analyses underlying the recent Australian drinking guidelines based on previous modelling (Rehm et al., 2008) attempted to model lifetime risk of injury mortality based on given drinking patterns, and found that non-linear risk functions best fitted the data, with inflection points at around 5-8 Australian standard drinks, suggesting that 5 standard drinks is a reasonable threshold for defining risky drinking (National Health and Medical Research Council, 2009, p46).

Thus, the findings of this survey-based study fit broadly with the findings of previous studies using both survey data and objective outcome measures. This suggests that, even considering the weaknesses identified above, the findings presented here are relatively robust. In spite of the increased levels of risk experienced by drinkers who engage in extremely heavy drinking episodes, there is no evidence that the use of risky drinking thresholds at these very high levels provides a better means of assessing risky drinking in the population.

References

- AIHW, 2011. 2010 National Drug Strategy Household Survey Report., Drug Statistics Series, No 25. AIHW, Canberra.
- Beck, J.R., Shultz, E.K., 1986. The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of pathology & laboratory medicine* 110, 13.
- Begg, S., Vos, T., Barker, B., Stevenson, C., Stanley, L., Lopez, A.D., 2007. The burden of disease and injury in Australia 2003. Australian Institute of Health and Welfare, Canberra.
- Bluman, A.G., 1992. *Elementary Statistics*. Wm. C. Brown, Dubuque.
- Bobak, M., Room, R., Pikhart, H., Kubinova, R., Malyutina, S., Pajak, A., Kurilovitch, S., Topor, R., Nikitin, Y., Marmot, M., 2004. Contribution of drinking patterns to differences in rates of alcohol related problems between three urban populations. *J Epidemiol Community Health* 58, 238-242.
- Dawson, D.A., Smith, S.M., Pickering, R.P., Grant, B.F., 2012. An empirical approach to evaluating the validity of alternative low-risk drinking guidelines. *Drug and Alcohol Review* 31, 141-150.
- Gartrell, A., Veness, P., 2008. Rudd to tackle binge drinking 'epidemic'. *The Herald Sun*, Melbourne.
- Gmel, G., Daeppen, J.B., 2007. Recall bias for seven-day recall measurement of alcohol consumption among emergency department patients: implications for case-crossover designs. *Journal of Studies on Alcohol and Drugs* 68, 303-310.
- Gmel, G., Graham, K., Kuendig, H., Kuntsche, S., 2006. Measuring alcohol consumption-should the 'graduated frequency' approach become the norm in survey research? *Addiction* 101, 16-30.
- Gmel, G., Room, R., Kuendig, H., Kuntsche, S., 2007. Detrimental drinking patterns: Empirical validation of the pattern values score of the Global Burden of Disease 2000 study in 13 countries. *Journal of Substance Use* 12, 337-358.
- Greenfield, T.K., 2000. Ways of measuring drinking patterns and the difference they make: experience with graduated frequencies. *Journal of Substance Abuse* 12, 33-49.
- Jackson, K.M., 2008. Heavy episodic drinking: Determining the predictive utility of five or more drinks. *Psychology of Addictive Behaviours* 22, 68-77.
- Livingston, M., Room, R., 2009. Variations by age and sex in alcohol-related problematic behaviour per drinking volume and heavier drinking occasion. *Drug and Alcohol Dependence* 101, 169-175.
- Mäkelä, P., Mustonen, H., 2007. How Do Quantities Drunk Per Drinking Day And The Frequencies Of Drinking Those Quantities Contribute To Self-Reported Harm And Positive Consequences? *Alcohol and Alcoholism* 42, 610-617.
- National Health and Medical Research Council, 2009. *Australian Guidelines to Reduce Health Risks from Drinking Alcohol*. NHMRC, Canberra.
- Rehm, J., Gmel, G., Room, R., Frick, U., 2001. Average volume of alcohol consumption, drinking patterns and related burden of mortality in young people in established market economies of Europe. *Eur Addict Res* 7, 148-151.
- Rehm, J., Room, R., Taylor, B., 2008. Method for moderation: measuring lifetime risk of alcohol-attributable mortality as a basis for drinking guidelines. *Int J Methods Psychiatr Res* 17, 141-151.
- Schwarz, G., 1978. Estimating the dimension of a model. *The annals of statistics* 6, 461-464.
- Stark, J., 2007. Alarm over binge-drinking epidemic. *The Age*. Fairfax, Melbourne.
- Taylor, B., Irving, H.M., Kanteres, F., Room, R., Borges, G., Cherpitel, C., Greenfield, T., Rehm, J., 2010. The more you drink, the harder you fall: A systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. *Drug and Alcohol Dependence* 110, 108-116.
- Veall, M.R., Zimmermann, K.F., 1996. Pseudo-R² Measures For Some Common Limited Dependent Variable Models. *Journal of Economic Surveys* 10, 241-259.
- Victorian Drug and Alcohol Prevention Council, 2010. 2009 Victorian Youth Alcohol and Drug Survey - Final Report. Victorian Department of Health, Melbourne.
- Wechsler, H., Moeykens, B., Davenport, A., Castillo, S., Hansen, J., 1995. The adverse impact of heavy episodic drinkers on other college students. *Journal of Studies on Alcohol* 56, 628-634.

Wechsler, H., Nelson, T.F., 2006. Relationship Between Level of Consumption and Harms in Assessing Drink Cut-Points for Alcohol Research: Commentary on "Many College Freshmen Drink at Levels Far Beyond the Binge Threshold" by White et al. *Alcoholism: Clinical and Experimental Research* 30, 922-927.

Weitzman, E.R., Folkman, A., Folkman, K.L., Wechsler, H., 2003. The relationship of alcohol outlet density to heavy and frequent drinking and drinking-related problems among college students at eight universities. *Health & Place* 9, 1-6.

Weitzman, E.R., Nelson, T.F., 2004. College Student Binge Drinking and The "Prevention Paradox": Implications For Prevention And Harm Reduction. *Journal of drug education* 34, 247-265.

White, A.M., Kraus, C.L., Swartzwelder, H.S., 2006. Many college freshmen drink at levels far beyond the binge threshold. *Alcoholism: Clinical and Experimental Research* 30, 1006-1010.

*Table 1 – Proportion of current drinkers drinking at various quantity and frequency thresholds, 2010
National Drug and Alcohol Strategy Household Survey.*

Quantity	Frequency		
	1+ occasions (yearly)	12+ occasions (monthly)	52+ occasions (weekly)
3 or more standard drinks	69.4%	60.4%	35.5%
5 or more standard drinks	51.8%	42.4%	19.2%
7 or more standard drinks	37.2%	27.3%	9.4%
11 or more standard drinks	22.5%	15.8%	3.9%
20 or more standard drinks	13.9%	5.5%	1.5%

Table 2 – Classification of survey items into broad harm indicators

Specific survey item	Broad harm group	Sample prevalence
Went to work under the influence of alcohol	Hazardous behaviour	4.5%
Went swimming under the influence of alcohol	Hazardous behaviour	6.4%
Operated a boat under the influence of alcohol	Hazardous behaviour	0.9%
Drove a motor vehicle under the influence of alcohol	Hazardous behaviour	13.7%
Operated hazardous machinery under the influence of alcohol	Hazardous behaviour	0.6%
Created a public disturbance or nuisance while under the influence of alcohol	Delinquent behaviour	2.3%
Caused damage to property while under the influence of alcohol	Delinquent behaviour	1.3%
Stole money, goods or property while under the influence of alcohol	Delinquent behaviour	0.4%
Verbally abused someone while under the influence of alcohol	Delinquent behaviour	4.9%
Physically abused someone while under the influence of alcohol	Delinquent behaviour	0.8%

Table 3 – Sensitivity and specificity of each risky-drinking threshold for three types of alcohol-related harm, best-fitting thresholds highlighted.

Threshold	Injury (n=19552)			Hazardous behaviour (n=19089)			Delinquent behaviour (n=18927)		
	Sensitivity	Specificity	Distance from perfect test	Sensitivity	Specificity		Sensitivity	Specificity	
3+/yearly	95.1%	34.1%	0.661	92.2%	38.6%	0.619	97.0%	34.8%	0.653
3+/monthly	90.9%	43.3%	0.574	88.0%	48.8%	0.526	94.9%	44.4%	0.558
3+/weekly	70.7%	66.9%	0.442	63.6%	72.6%	0.456	74.3%	68.5%	0.407
5+/yearly	89.0%	53.1%	0.482	80.1%	58.9%	0.457	92.0%	55.5%	0.452
5+/monthly	84.4%	62.6%	0.405	71.6%	68.5%	0.424	87.8%	64.3%	0.377
5+/weekly	54.9%	83.5%	0.480	41.0%	87.8%	0.602	57.0%	85.0%	0.455
7+/yearly	79.9%	68.0%	0.378	64.7%	73.7%	0.440	82.3%	69.7%	0.351
7+/monthly	71.9%	77.4%	0.361	51.7%	82.3%	0.514	71.7%	79.2%	0.351
7+/weekly	38.6%	92.6%	0.618	21.3%	94.8%	0.789	36.1%	93.6%	0.642
11+/yearly	62.2%	82.0%	0.419	44.5%	86.5%	0.571	62.7%	83.6%	0.407
11+/monthly	53.1%	87.9%	0.484	33.1%	91.4%	0.675	52.5%	89.4%	0.487
11+/weekly	21.6%	97.2%	0.784	9.5%	98.2%	0.905	18.8%	97.8%	0.812
20+/yearly	45.2%	89.4%	0.558	28.9%	92.4%	0.715	45.7%	90.7%	0.551
20+/monthly	26.5%	96.2%	0.736	12.6%	97.5%	0.874	23.5%	96.9%	0.766
20+/weekly	11.0%	98.9%	0.890	3.2%	99.2%	0.968	6.8%	99.2%	0.932

Table 4 –Model fit statistics for logistic regression models testing each risky-drinking threshold across three types of alcohol-related harm, best-fitting thresholds highlighted.

Threshold	Injury			Hazardous behaviour			Delinquent behaviour		
	Pseudo R ²	BIC	Chi-Squared	Pseudo R ²	BIC	Chi-Squared	Pseudo R ²	BIC	Chi-Squared
3+/yearly	0.192	4,440.4	93.4	0.120	16,301.9	122.0	0.180	7,278.6	97.1
3+/monthly	0.197	4,414.4	81.4	0.133	16,061.6	125.3	0.197	7,130.1	103.6
3+/weekly	0.215	4,316.6	86.0	0.132	16,072.7	132.1	0.215	6,977.0	79.8
5+/yearly	0.205	4,368.7	87.3	0.071	16,190.3	110.8	0.203	7,083.1	100.5
5+/monthly	0.216	4,312.6	72.7	0.128	16,141.1	114.9	0.221	6,922.2	93.7
5+/weekly	0.220	4,289.7	92.1	0.118	16,337.1	133.2	0.217	6,958.5	101.0
7+/yearly	0.213	4,328.7	71.7	0.122	16,264.5	143.6	0.212	7,004.1	172.2
7+/monthly	0.225	4,263.3	74.8	0.116	16,360.2	150.7	0.217	6,957.6	136.4
7+/weekly	0.216	4,310.9	103.8	0.095	16,750.6	143.9	0.196	7,144.6	105.1
11+/yearly	0.214	4,320.4	79.1	0.112	16,445.2	158.8	0.205	7,065.9	131.6
11+/monthly	0.219	4,284.7	78.0	0.104	16,589.1	165.4	0.208	7,035.9	121.6
11+/weekly	0.199	4,402.6	115.4	0.083	16,983.1	165.1	0.174	7,335.4	136.2
20+/yearly	0.207	4,362.7	81.2	0.098	16,700.6	180.1	0.194	7,161.2	132.2
20+/monthly	0.201	4,393.7	118.0	0.085	16,931.5	159.8	0.176	7,317.0	161.6
20+/weekly	0.191	4,460.4	79.3	0.074	17,143.5	157.2	0.151	7,530.9	139.8

Models control for age, sex and marital status. Note that regression analyses controlling for income, neighbourhood socio-economic status and region along with sex, age and marital status were developed, but made no difference to the thresholds identified.

Figure 1 – Proportion of risky-drinkers who report that either they or somebody else have been injured in the past 12 months due to their drinking, by risky-drinking threshold (amount and frequency)

Figure 2 – Proportion of risky-drinkers who report hazardous behaviour in the past 12 months due to their drinking, by risky-drinking threshold (amount and frequency)

Figure 3 – Proportion of risky-drinkers who report delinquent behaviour in the past 12 months due to their drinking, by risky-drinking threshold (amount and frequency)