Harmful Drinking Phenotype in a Large Dutch Community Sample

Koen Hogenelst 1,*, Almar Snippe², and Marjana Martinic³

¹Department of Human Performance, The Netherlands Organisation for Applied Scientific Research (TNO), Soesterberg 3769 DE, The Netherlands

²Department of Weapon Systems, The Netherlands Organisation for Applied Scientific Research (TNO), The Hague 2597 AK, The Netherlands ³MM Science & Policy Advisors, LLC, Washington, DC 20007, USA

*Corresponding author: Human Performance, The Netherlands Organisation for Applied Scientific Research (TNO), Kampweg 55, 3769 DE Soesterberg, The Netherlands. Tel: +31-6 316 676 33; E-mail: koen.hogenelst@tno.nl

Abstract

Aims: Harmful drinking patterns are shaped by a broad complex interaction of factors, societal and individual, psychological and behavioral. Although previous studies have focused on a few variables at a time, the current study simultaneously examines a large number of variables in order to create a comprehensive view (i.e. phenotype) of harmful drinking, and to rank the main predictors of harmful and non-harmful drinking by order of importance.

Methods: We surveyed a large sample of Dutch adults about their habitual drinking characteristics and attitudes, perceptions and motives for drinking. We fed 45 variables into a random forest machine learning model to identify predictors for (1) drinking within and in excess of Dutch guideline recommendations and (2) harmful and non-harmful drinking.

Results: In both models, respondents' subjective perceptions of 'responsible drinking', both per occasion and per week, showed the strongest predictive potential for different drinking phenotypes. The next strongest factors were respondents' reason for drinking, motives for drinking and age. Other variables, such as drinking location, knowledge about alcohol-related health risks and consumption of different beverage types, were not strong predictors of drinking phenotypes.

Conclusions: Although the direction of the relationship is unclear from the findings, they suggest that interventions and policy measures aimed at individuals and social norms around drinking may offer promise for reducing harmful drinking. Messaging and promotion of drinking guidelines should be tailored with this in mind.

INTRODUCTION

Harmful drinking is associated with a range of short- and long-term health risks (Ronksley *et al.*, 2011; Piano, 2017; WHO, 2018; Carvalho *et al.*, 2019; Roerecke *et al.*, 2019; Shield *et al.*, 2020) and with negative psychosocial consequences (Appleton *et al.*, 2018). As a result, it is the focus of high-profile initiatives like the World Health Organization's (WHO) Global Strategy to reduce the harmful use of alcohol (WHO, 2010) and many national-level efforts.

In the Netherlands, the net social cost of alcohol use is estimated at Euro 2.5 billion annually [de Wit *et al.*, 2019; Central Bureau of Statistic (CBS), 2021] and includes the burden from lost productivity, law enforcement and legal costs, healthcare costs, and life-years lost to accidental injuries (e.g. road traffic), chronic (non-communicable) diseases and alcohol abuse. Given the health risks associated with harmful drinking, the Health Council of the Netherlands has developed drinking guidelines for the general adult population, which recommends 'Do not drink alcohol or drink no more than one glass daily'. Effectively, the guideline is an average of at most one drink per day. A 'glass' (or drink) in the Netherlands is officially defined as containing ~10 g of pure alcohol (Health Council of the Netherlands, 2015).

A 2020 population survey in the Netherlands found that 44% of the Dutch population over the age of 18 years met

the guideline of the Dutch Health Council, whereas 6.9% were considered excessive drinkers and 7.8% heavy drinkers (CBS, 2021). In line with definitions used by the two collaborating entities, Trimbos and RIVM, we considered 'excessive' and 'heavy' drinkers to be two distinct groups though not mutually exclusive. We defined 'Excessive' drinkers as consuming on average more than two drinks per day for women and more than three drinks per day for men. We defined 'Heavy' drinkers as engaging in heavy episodic drinking (HED) (four or more drinks on an occasion for women and six or more drinks on an occasion for men, at least once a week). Yet, heavy drinkers may also drink excessively (i.e. HED at least once a week and consuming on average more than two drinks per day for women and more than three drinks per day for men). In the present study, excessive drinking and HED are jointly referred to as 'harmful' drinking. Harmful drinking patterns can be modified through prevention and policy interventions, which rely on a better understanding of their prevalence and of the factors that drive them.

There is a wealth of studies that have assessed the associations between behavioral, psychological and social variables and harmful drinking. These variables include, for example, drinking settings and locations (Callinan *et al.*, 2016), social context (Cullum *et al.*, 2012; Monk *et al.*, 2020), beverage

Received: February 22, 2022. Revised: July 18, 2022. Accepted: July 29, 2022

© The Author(s) 2022. Medical Council on Alcohol and Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution NonCommercial-NoDerivs licence (http://creativecommo ns.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

type (Ally et al., 2016; Kilian et al., 2021), awareness and knowledge of alcohol-related health risks (Bowden et al., 2014), individual reasons or motives for consuming alcohol (Crutzen and Kuntsche, 2013), and perceptions and attitudes towards alcohol and drinking (Livingstone and McCafferty, 2015). Also included are perceptions of what is considered 'responsible' drinking, shaped in part by the general drinking culture and in part by individual-level factors. The term 'responsible drinking' is poorly defined in terms of consumption volume and frequency and has been criticized for its use in campaigns supported by the alcohol industry (Maani Hessari and Petticrew, 2018; Gray et al., 2021). Despite criticism, 'responsible drinking' can be a useful concept for capturing individual-level views and definitions of appropriate or acceptable drinking patterns and is used within that context in this study. It may be expected that perceptions of what is 'responsible drinking' in terms of frequency and amount may vary among individuals and reflect other factors and motivations that underlie their drinking behavior. Insights into the relationship between people's subjective perception of responsible drinking and their actual intake may guide future interventions to combat harmful drinking.

In this study, we examine the combination of characteristics of harmful drinking among a Dutch sample of adults. Previous studies generally focused on single variables and their associations with harmful drinking behavior. In contrast, in this study, we use a random forest model (Breiman, 2001), an innovative approach that allowed us to consider multiple factors simultaneously and create a more comprehensive picture of what, when, where, with whom and why people drink. This method was previously used to predict harmful drinking (Bonnell *et al.*, 2020). Here, we used it to rank both individuallevel variables and behavioral aspects, motivations, attitudes and perceptions in order to distinguish harmful drinkers from non-harmful drinkers.

Improving our ability to understand the interactions among behavioral, social and psychological factors that characterize harmful drinking supports the development of better public education and awareness around drinking and its potential risks. This knowledge can help to create more effective health promotion and interventions to reduce harmful drinking.

METHODS

Ethical procedures

We performed this study in accordance with the Declaration of Helsinki (World Medical Association, 2018). The study protocol received a favorable review of the TNO internal review board (reg. nr. 2020-106). The data reported here are part of a larger study conducted in October and November 2020. Participants started with retrospective questionnaires in which they reported on their drinking behavior prior to the Coronavirus Disease 2019 (COVID-19) outbreak in the Netherlands (February 2020 as the retrospective period) and during the outbreak (September 2020 as the retrospective period). Subsequently, participants completed a 4-week diary reporting their daily drinking behavior. The data used in the present study include only the retrospective data on drinking behavior prior to the COVID-19 outbreak and implementation of social measures in the Netherlands.

Participants

We recruited participants through existing survey panels throughout the Netherlands to include a nationally representative sample in terms of age, sex and educational distribution. We sent panel members an email announcement including a brief study description. Through a link in the email, we directed participants to a webpage where they could view the participant information and provide their consent to participate by continuing to the screening questions. Inclusion criteria were age (between 18 and 65 years) and currently drinking alcohol at least once a month.

Individuals reported their age and biological sex and subsequently answered three alcohol consumption questions from the short-form Alcohol Use Disorders Identification Test-Concise to screen for heavy drinking and/or alcohol abuse or dependence (Bush *et al.*, 1998). We screened women who scored three or more points on these questions and men who scored four or more using diagnostic criteria for Alcohol Use Disorders (AUD) as defined in DSM-5 (American Psychiatric Association, 2013). We excluded those meeting four or more criteria for AUD from the study, whereas the others continued to the questionnaire. We financially reimbursed participants for their participation.

Questionnaire

The variables and questions are briefly described in the following sections.

Consumption

For consumption of alcoholic beverages we asked participants about their drinking behavior in 'the past month'. When estimating their consumption, we instructed participants to consider the Dutch definition of a standard drink expressed as alcohol by volume (ABV), which contains ~ 10 g of pure ethanol (Health Council of the Netherlands, 2015) and is equivalent to one 250 ml serving of beer (5% ABV), 100 ml of wine (12% ABV) or a 35 ml serving of distilled spirits (35% ABV).

To determine frequency of drinking and amount consumed, we first asked participants to respond to the question: 'In February 2020 how often did you drink alcohol-containing beverages?'. Options included 'never', 'monthly', 'weekly' and 'daily'. For 'monthly' and 'weekly' responses, we assessed the number of days on which alcohol was consumed. Subsequently, we assessed the number of standard drinks per drinking occasion. Participants then reported the three types of alcoholic beverages they drank most commonly, up to three of the usual locations where they drank, up to three of their most common reasons to drink and their usual number of drinking companions. We then asked participants how often they drank six or more drinks (for men) or four or more drinks (for women) on a single occasion. Response options included 5-6 times per week, 3-4 times per week, 1-2 times per week, 2–3 times per month and once per month.

We considered participants who indicated they consumed no more than one drink on an occasion to have met the Dutch drinking norm of drinking within the recommended guideline (Health Council of the Netherlands, 2015). In addition, we calculated the average number of drinks consumed per day using frequency and amount of consumption. We defined light drinkers as consuming on average up to one glass of alcohol a day and moderate drinking as an average of one to two (for women) or one to three (for men) glasses of alcohol a day. In line with definitions used in national alcohol consumption surveys in the Netherlands (Statistics Netherlands, CBS, 2021), we defined 'excessive drinking' as consuming on average more than two (for women) or three (for men) glasses a day and 'HED' as drinking four or more (for women) or six or more (for men) glasses on one occasion, at least once a week. For the purpose of this study, we defined harmful drinking as including both excessive drinking and HED. The other two groups (i.e. light and moderate drinkers) together formed the category of non-harmful drinking.

Attitudes towards alcohol consumption

We assessed general attitudes towards alcohol consumption by three items: 'Having a drink is one of the pleasures of life', 'Having a drink with someone is a way of being friendly' and 'There is nothing good to be said about drinking'. We asked participants to rate items on a 5-point Likert scale ranging from 'Strongly agree' (coded as 1) to 'Strongly disagree' (coded as 5) (adapted from Alcohol Research Group, 2009).

Furthermore, we assessed attitudes towards alcohol consumption in specific situations, which included at party at someone's home; a mother spending time with a young child; a father spending time with a young child; a man dining alone in a restaurant; a woman dining alone in a restaurant; a man at a bar with friends; a woman at a bar with friends; a woman out with colleagues; a man out with colleagues; at home with friends; when having to drive; a man having dinner at home with a partner; a woman having dinner at home with a partner [questions adapted from Gender, Alcohol, and Culture: An International Survey (GENACIS); GENACIS, 2018]. For each situation, we asked participants to indicate how much a person in that situation should be able to drink. Response options were (1) zero drinks, (2) some (one or two) drinks but not enough to feel the effects, (3) enough to feel the effects but not become drunk, (4) getting drunk is sometimes alright and (5) getting drunk is always alright.

To determine perceptions of acceptable drinking we asked participants to respond to the question: 'What do you consider to be responsible drinking?' in 2-fold: (1) by the number of days per week (from 0 to 7) on which drinking usually takes place and (2) by the number of drinks consumed on those days (from 0 to 10). By multiplying these two answers we computed a third variable reflecting perceptions of the 'responsible number of drinks per week'.

Knowledge and perceptions related to the consumption of alcohol

We assessed knowledge of health risks related to the consumption of alcohol as follows: 'How much do you agree or disagree that the consumption of alcoholic beverages can increase the risk of the following health conditions?'. We provided participants a list of health risks including liver disease, heart disease, cancer, asthma, depression and birth defects. We asked participants to rate each health risk on a 5-point Likert scale ranging from 'Strongly agree' (coded as 1) to 'Strongly disagree' (coded as 5) (we derived items from European Commission, 2010). We assessed perceptions about the health risks associated with specific consumption patterns by asking: 'How much do you think people risk harming themselves if they do each of the following?'. We provided a list of behaviors including: try alcohol once or twice; have one or two drinks nearly every day; have four or five drinks nearly every day; have five or more drinks each weekend. We asked participants to rate risk on a 4-point Likert scale ranging from 'No risk' (coded as 1) to 'Great risk' (coded as 4) (we derived items from Hibell et al., 2012). In addition, we assessed attitudes towards the acceptability of drinking a glass of beer versus a glass of wine versus a glass of spirits.

Motivations for alcohol consumption

We used the Drinking Motive Questionnaire revised short form (DMQ–R SF) (Kuntsche and Kuntsche, 2009) to assess motivations for alcohol consumption. It consists of 12 items that are rated on a 3-point Likert scale from 'Never' (coded as 1) to 'Sometimes' (coded as 2) to 'Almost always' (coded as 3), with three items for each of the four subscales: drinking for (a) enhancement (e.g. drinking to have fun and to get drunk), (b) social (e.g. to better enjoy social gatherings), (c) coping (e.g. to alleviate personal problems and worries) and (d) conformity motives (e.g. not to feel left out of the group). We averaged the three items within each subscale for scoring.

Knowledge regarding alcoholic beverage types

We assessed knowledge of alcohol content in beverages by three items: 'Does a regular glass of beer have [more/less/equal amount of] alcohol than a regular glass of wine?', 'Does a regular glass of beer have [more/less/equal amount of] alcohol than a regular glass of distilled spirits?' and 'Does a regular glass of wine have [more/less/equal amount of] alcohol than a regular glass of distilled spirits?'.

Analysis

To study the classification of drinking types using numerous behavioral and psychological variables, we used a random forest model. Random forest models can handle mixtures of continuous and categorical data, and are therefore among the most powerful, fully automated, machine learning techniques (Breiman, 2001). The random forest method is a generalization of the decision tree method. For the present study, we split all included participants into two groups for each variable in the dataset (e.g. meeting drinking norms versus not meeting drinking norms). After a certain number of splits, the end groups (i.e. leaves) are homogeneous subsets (i.e. subsets with drinkers who meet the norm and subsets of drinkers who do not). The random forest model builds many decision trees using a random selection of variables and random subsets of the dataset for each decision tree. Consequently, the number of variables and participants used varies from decision tree to decision tree. All decision trees, with their error rates, are compared to determine the joint set of variables that produce the strongest classification model (meet norm/exceed norm). The random forest algorithm estimates the (mean Gini) importance of a variable (from 0 to 100) by looking at how much the impurity of the leaves (subsets) increases when one variable is removed from the model while all others are left unchanged; little increase of impurity implies low importance (Breiman, 2001). In the present study, we used 45 behavioral and psychological variables as input. For the final model, we considered variables relevant if their mean Gini importance (i.e. the increase of impurity of the leaves in the model when the variable was excluded) was above the 'elbow' when looking at the importance of all variables in decreasing order (Genuer et al., 2010). In short, only a few variables have a (distinctively) high importance followed by a bunch of variables with mutually comparable importance; this introduces an elbow (or plateau) when plotting the importance in decreasing order. We built two random forest models to classify (1) individuals meeting the Dutch drinking

Age and sex distribution in the two classification models



Fig. 1. Sample distribution of age groups and sex (%) for drinking within versus exceeding the Dutch drinking norm (left graph) and harmful drinkers versus non-harmful drinkers (right graph).

norm versus drinkers who exceed this norm, and (2) harmful drinkers versus non-harmful drinkers.

Above all, we used random forest analyses as an exploratory data mining technique to find the best subset of variables for constructing a model to describe a specific drinking category. The outcome is a model (a forest with multiple decision trees) that can be used on data of new persons to 'classify' a type of drinker, though this is not the focus of the present study. We performed all random forest analyses using the randomForest R package (Liaw and Wiener, 2002) in R (R Development Core Team, 2019).

RESULTS

Baseline characteristics

The final sample consisted of 1184 included individuals. Young males (aged 18-29 years) were underrepresented (by \sim 50%), but otherwise the sample was representative of the Dutch population. Figure 1 illustrates the distribution per model group by age and sex. Among the final sample, 30% (360 respondents) met the Dutch drinking norm. The remaining 70% (824 respondents) exceeded the norm (Model 1). Those meeting the norm were older (mean ages of 47 and 44 years, respectively) and more likely to be female (69 versus 55%) than those who did not. Out of the 1184 individuals, 10.5% (124 respondents) were 'harmful drinkers' (Model 2). The remaining 89.5% (1060 respondents) were not (Model 2). The mean age among harmful drinkers was 46 years, compared with 45 among non-harmful drinkers, and harmful drinkers were also less likely to be female (52 versus 60%), but not significantly so. The P-value on the two-sample proportion test was greater than 0.05.

Random forest model: ranking predictive variables Model 1: drinking within versus exceeding the Dutch drinking norm

A total of six variables emerged in the classification model as significant predictors of drinking within and in excess of the Dutch drinking norm (Fig. 2). In order of importance, these included perception about how many drinks constitute 'responsible drinking' on a given occasion (rank 1) and in a week (rank 2), the most common reasons for drinking (rank 3), age (rank 4), drinking for enhancement (rank 5) and social motives (rank 6). For this model, an accuracy of 89% was found in the training set that included 884 subjects and an accuracy of 71% in the validation set that included 300 subjects. Table 1 shows the group characteristics and the means and variations for each group of the six most important model variables. Figure 3 illustrates the distribution within each group for each of these six variables.

Compared with individuals drinking within the Dutch norm, those drinking in excess of the norm considered 'responsible drinking' to involve more drinks both per occasion (2.46 versus 1.44 standard drinks) and per week (6.06 versus 3.65 standard drinks). The most common reason for drinking reported by both groups was 'At home, because they enjoy it'. As the second most common reason to drink, individuals drinking in excess of the Dutch norm reported 'Drinking to party or have fun', whereas those drinking within the norm reported 'To accompany dinner at home or home of others'. Individuals drinking more than the norm were younger (44.0 versus 47.4 years) and had a stronger motive to drink for enhancement (1.80 versus 1.52) and for social reasons (1.95 versus 1.62) compared with drinkers who stayed within the Dutch norm.

Model 2: harmful drinkers versus non-harmful drinkers

In this classification model, six variables also emerged as significant predictors of harmful versus non-harmful drinking (Fig. 4). Ranked in order of importance, these include perceptions of 'responsible drinking' in terms of the number of drinks per week (rank 1), drinking for coping (rank 2), enhancement motives (rank 3), perceptions of 'responsible drinking' in terms of the number of drink per occasion (rank 4), the most common reasons to drink (rank 5) and age (rank 6). For this model, we found an accuracy of 80% in the training set of 884 subjects and an accuracy of 79% in the validation set of 300 subjects. Table 2 shows the group characteristics and the means and variation for each group of the six most important model variables. Figure 5 illustrates the distribution within each group for each of the six included variables. Compared with non-harmful drinkers, harmful drinkers on average perceived the 'responsible' number of drinks to be greater than non-harmful drinkers both per occasion (3.34 versus 2.01 standard drinks) and per week



Fig. 2. Random forest model with 45 variables ranked by importance (most important at the top) to classify norm drinkers versus more than norm drinkers. Variables above the black line have a (distinctively) high classification importance.

Table 1. Variable descriptives for Model 1

	Group	
	Meets Dutch drinking norm $(n=360)$	Exceeds Dutch drinking norm (<i>n</i> = 824)
Age in years, mean (SD)	47.4 (12.6)	44.0 (13.4)
Most common reason to drink (rank)	'At home because I enjoy it'	'At home because I enjoy it'
1.	'To accompany dinner at home or home of	'To party or have fun'
2.	others'	'To accompany dinner at home or home of
3.	'To celebrate something'	others'
Perception of responsible drinks per occasion,	1.44 (0.65)	2.46 (1.31)
mean (SD)	3.65 (2.69)	6.06 (5.23)
Perception of responsible number of		
drinks/week, mean (SD)		
DMQ—Social, mean (SD)	1.62 (0.53)	1.95 (0.60)
DMQ—Enhance, mean (SD)	1.52 (0.41)	1.80 (0.45)

(10.60 versus 4.71 standard drinks). Both groups reported the most common reason to drink as 'At home, because they enjoy it'. As the second most common reason to drink, harmful drinkers reported 'Drinking to party or have fun', whereas non-harmful drinkers reported 'To accompany dinner at home or home of others'. Compared with non-harmful drinkers, harmful drinkers had a stronger motive to drink for enhancement (2.10 versus 1.67) and coping reasons (1.80 versus 1.33) and were older (46.1 versus 44.9 years).

DISCUSSION

This study investigated drinking patterns in a large Dutch community sample using a machine learning approach to understand the relative importance of 45 behavioral and psychological variables in determining drinking patterns. We built random forest models to distinguish between (I) individuals who drank within and those who exceeded Dutch drinking guidelines and (II) harmful versus non-harmful



Drinking within or more than the Dutch Drinking norm

Fig. 3. Distribution within each group for each of the six highest ranked variables in Model 1. Categories for 'most common reason to drink' are as follows: 1 = party, have fun; 2 = celebrate something; 3 = dinner in restaurant; 4 = dinner at home or home of others; 5 = drinks at or after work; 6 = at home, because its available; 7 = at home, because I enjoy it; 8 = at home, because it's a habit; 9 = other.

Harmful drinkers



Fig. 4. Random forest model with 45 variables ranked by importance (most important at the top) to classify harmful drinkers versus non-harmful drinkers. Variables above the black line have a (distinctively) high classification importance.

Table 2. Variable descriptives for Model 2

	Group	
	Non-harmful drinkers (n=1060)	Harmful drinkers (<i>n</i> = 124)
Age in years, mean (SD)	44.9 (13.3)	46.1 (13.2)
Most common reason to drink (rank)	'At home because I enjoy it'	'At home because I enjoy it'
1	'To accompany dinner at home or home of	'To party, have fun'
2	others'	'At home, because it's a habit'
3	'To party, have fun'	
Perception of responsible drinks per occasion,	2.01 (1.04)	3.34 (1.96)
mean (SD)	4.71 (3.61)	10.60 (8.53)
Perception of responsible number of		
drinks/week, mean (SD)		
DMQ—Coping, mean (SD)	1.33 (0.43)	1.80 (0.61)
DMQ—Enhance, mean (SD)	1.67 (0.43)	2.10 (0.49)



Fig. 5. Distribution within each group for each of the six highest ranked variables in Model 2. Categories for 'most common reason to drink' are as follows: 1 = party, have fun; 2 = celebrate something; 3 = dinner in restaurant; 4 = dinner at home or home of others; 5 = drinks at or after work; 6 = at home, because its available; 7 = at home, because I enjoy it; 8 = at home, because it's a habit; 9 = other.

drinkers. The group of harmful drinkers included both 'excessive' drinkers and those who engaged in HED, as defined for the Netherlands. Of the 45 variables included those with substantial differentiating potential, in both models, included respondents' subjective perceptions of what constituted a 'responsible' number of drinks both per occasion and per week, the most commonly cited reason to drink, drinking motives, including enhancement, coping and social reasons, and respondents' age.

Discussion of the main findings

The findings show that respondents' own definitions of 'responsible drinking' were the strongest predictors of drinking in excess of the Dutch drinking norm and of harmful drinking patterns. The term 'responsible drinking' has been identified by some as a charged term, often used in industry-led campaigns. A criticism of the use of the term has been a lack of a coherent definition with specificity in quantity and frequency (Moss and Albery, 2018). However, in the present

study, we asked respondents to quantify, in frequency and amount, their own definition of acceptable drinking behavior. As such, it may be seen as useful concept for exploring how consumers view acceptable drinking behavior as well as how this predicts their own drinking behavior. Compared with non-harmful drinkers, harmful drinkers perceive a higher number of alcoholic drinks per occasion and per week as responsible drinking. Moreover, this variable was the strongest predictor of all studied variables in distinguishing harmful from non-harmful drinkers as well as drinking within or in excess of the Dutch drinking norm. These findings are consistent with other studies showing that perceptions of normative drinking, as well as perceptions of risk correlate with drinking behavior (Foster et al., 2015; Bräker and Soellner, 2016; Kuntsche et al., 2017; Lau-Barraco et al., 2017; Dillard et al., 2018). Differences in how definitions of responsible drinking vary between those who engage in harmful drinking and those who do not merit further study.

The second most salient variable influencing drinking behavior was motivation. Specifically, enhancement, coping and social motives to drink were important variables in distinguishing drinking groups. For each type of motivation, the scores were higher among those that exceed the Dutch drinking norm and those that engage in harmful drinking compared with the other groups. These findings are in line with prior research (e.g. Kuntsche et al., 2014), including a study in a large Dutch sample (Crutzen and Kuntsche, 2013). In that study, social motives were positively associated with the number of drinks on the heaviest drinking day and enhancement motives with the number of drinking days. Coping motives were positively associated with both drinking measures (Crutzen and Kuntsche, 2013). Together, these findings underscore the relevance to target motivation when targeting harmful drinking patterns.

Another important variable in both random forest models was 'most common reason to drink'. Both harmful and nonharmful drinkers reported drinking 'at home, because I enjoy it' as the most common reason to drink. As the second most important reason, harmful drinkers reported 'party, have fun' (regardless of location), whereas non-harmful drinkers reported 'dinner at home or home of others'. As the third most important reason, harmful drinkers reported 'at home, because it's a habit', whereas non-harmful drinkers reported 'party, have fun'.

These findings reflect subtle differences between groups, specifically, that harmful drinkers are more likely to (a) drink to party and have fun and (b) habitually drink at home. However, since we did not assess the variable 'most common reason to drink' in terms of amount consumed, it is not possible to conclude whether harmful drinking was more likely to occur at home than elsewhere. The findings do imply that most drinking occurs at home, for both non-harmful and harmful drinkers. Other studies across countries have shown that most drinking, including harmful drinking, occurs within the home, and not in licensed premises (Callinan *et al.*, 2016, 2021).

Lastly, age was an important variable in both models, albeit in opposite ways. Individuals drinking within the norm were older than those drinking in excess of the norm (Model 1), whereas non-harmful drinkers were younger than harmful drinkers (Model 2). It should be noted that the group of moderate drinkers, which in the present study was a relatively younger group, exceeds the Dutch drinking norm (Model 1), but in Model 2 is part of the non-harmful drinking group. The outcome of Model 1 is in line with outcomes of another large survey in the Netherlands showing that with increasing age, the percentage of people meeting the Dutch drinking norm also increases (Schouten et al., 2021). The study also reported that harmful drinking decreased with age (at least until the age of 65), and that young adults aged 18-29 years were particularly likely to engage in harmful drinking. This finding is at odds with the outcome of Model 2. This discrepancy with prior research may be in part related to the underrepresentation of young males, who are on average more likely to engage in harmful drinking (Schouten et al., 2021, US National Survey on Drug Use and Health, 2019).

Other correlates

Variables such as biological sex, perception of alcohol-related health risks, social context, drinking location and beverage type were not strong predictors in either model of whether respondents drank within or in excess of the Dutch norm, or whether they engaged in non-harmful versus harmful drinking.

In the present study, there was a higher proportion of females in the group drinking within the norm (69%) than in the group exceeding it (55%), and also among non-harmful drinkers (60%) compared with harmful drinkers (52%). Previous national surveys have consistently shown that harmful drinking is more likely among males than females (Schouten *et al.*, 2021, US National Survey on Drug Use and Health, 2019). Apparently, however, in the present study, the difference between males and females was not sufficient to render 'sex' an important distinguishing variable in either model. The underrepresentation of young males may have influenced this limited importance of 'sex' in the model outcomes.

Awareness and knowledge about alcohol-related health risks are important elements in prevention (Bowden et al., 2014). Drinking guidelines and health literacy around alcohol are predicated upon the notion that knowledge and understanding of risks associated with different drinking patterns will act to moderate drinking behavior and prevent harmful drinking. The fact that perception of alcohol-related health risks was not a strong differentiator between drinking groups might reflect poor awareness of alcohol-related harms in general or at least reflect few differences in knowledge or perceptions of alcohol-related health risks between these groups. In a recent national survey in the Netherlands, most participants were unaware that excessive drinking (part of harmful drinking in the present study) increases the risk of different types of cancer and dementia (Schouten et al., 2021). However, a lack of awareness of alcohol-related harms was most widespread among those engaging in harmful drinking. Taken together, these and other similar findings call for more intensive efforts to share information and increase awareness and knowledge of alcohol-related health risks.

Drinking location also did not strongly distinguish drinking phenotypes from each other. For both harmful and nonharmful drinkers, the most important 'reason' to drink was 'at home, because I enjoy it'. These results are in line with findings from other large nationally representative surveys reporting that people consume the majority of alcohol they drink in their own homes (Callinan et al., 2016, 2021). However, the present study shows drinking location was not strongly distinguishing non-harmful from harmful drinkers. The importance of the social context within which drinking takes place has been well described, including drinking in (or due to) social isolation, drinking as a social activity, the role of 'drinking confederates' and group pressure in drinking situations (Collins and Marlatt, 1981; de Wit and Sayette, 2018; Kelly et al., 2018; Monk et al., 2020). However, in the present study, social context was also not a strong drinking phenotype characteristic.

Finally, the study showed no difference across the two models in the importance of the type of alcohol beverage in predicting whether drinking was harmful or whether or not it exceeded the norm. Other studies have suggested that the type of beverage most associated with harmful drinking was the one that is most consumed in a country (Rossow, 2001; Schmidt *et al.*, 2003), and not any particular type of drink. In the Netherlands, beer is the most consumed beverage type, followed by wine (WHO, 2018), and it is reasonable to assume that it is also more likely to be consumed by harmful drinkers and those exceeding the drinking norm. Taken together, these results suggest that individual characteristics such as perceptions, reasons and motivation to drink may be better targets for combatting harmful drinking than drinking location or beverage type.

Strengths and weaknesses of the present study

The present study has several strengths. It uses a large sample with a wide range of drinking levels and a broad age group (18–65 years), allowing disaggregation of drinking patterns within the survey population. Its main strength is the simultaneous inclusion of 45 behavioural and psychological variables that could be ranked according to importance and predictive value. These include habitual drinking characteristics as well as attitudes towards, motives for, and perceptions about alcohol consumption. Furthermore, the application of a machine learning technique such as the random forest model offers an opportunity to simultaneously assess the relative strength of different variables related to alcohol consumption in a novel way.

At the same time, the study also has its limitations. First, younger males (between 18 and 29 years) were underrepresented in the sample, which is problematic given their usually strong representation in the excessive and heavy drinker categories (Schouten et al., 2021). Second, whereas the addition of perceptions and subjective definitions of 'responsible drinking' was novel and relevant for the model, we did not assess knowledge among drinkers about current Dutch drinking guidelines. Previous research, including a recent government-funded alcohol survey, found that knowledge about drinking guidelines is lower among those who drink excessive or heavily than among lighter drinkers (Schouten et al., 2021). The inclusion of this dimension would have shed light on the extent to which perceptions of responsible drinking are related to knowledge of drinking guidelines. Another limitation is that we defined harmful drinkers as those engaging in excessive drinking and HED. In reality, for many people, harmful drinking is likely to involve levels below the threshold for either of these but higher than levels defined by the Dutch drinking norm. Furthermore, statistical power was too low to examine excessive drinking and HED as separate categories. The two groups of drinkers might be different and have different predictive variables.

Implications/future perspectives

In the present study, self-rated 'responsible drinking' is the most powerful variable that separates those drinking within the Dutch norm from those exceeding the norm, and also harmful drinkers from non-harmful drinkers. Evidently, from our (correlational) data, the directionality of the relationship remains to be seen; in other words, whether a person's idea of 'responsible drinking' determines their drinking behavior or whether actual drinking behavior determines the subjective definition of 'responsible drinking', or perhaps both. Clearly, individuals' perceptions of 'responsible drinking', in terms of both the number of drinks per occasion and per week, deviate from the Dutch Health Council's drinking norm that advises not drinking at all or at least no more than one drink a day. This was particularly the case for harmful drinkers. Whether this deviation between perception of 'responsible drinking' and the Dutch drinking norm is related to a lack of awareness about the norm or whether it reflects a general acceptance of drinking levels that exceed the norm and the associated risk cannot be concluded from this study. Previous research performed by the Trimbos Institute in October 2020 shows that 55.6% of the total Dutch population were unaware of the official drinking guidelines, and 43.6% of people who drink excessively (Schouten *et al.*, 2021). Together with the outcome of the present study, these findings underscore the need for strengthening and continuing to provide information and education about alcohol consumption, particular attention to specific target groups like excessive and heavy drinkers, and guidelines that are realistic within the context of prevailing drinking patterns.

CONCLUSION

In the present study, we employed a random forest model to classify and rank a wide range of behavioral and psychological variables related to alcohol consumption in a large Dutch community sample. Though the direction of the relationship is unclear from the findings, they suggest that interventions and policy measures aimed at individuals' perceptions, motivation and social norms around drinking may offer promise for reducing harmful drinking. Messaging and promotion of drinking guidelines should be tailored with this in mind.

FUNDING

This research was funded by Diageo plc. Diageo had no role in study design, data collection, analyses or reporting.

CONFLICT OF INTEREST STATEMENT

KH and AS are employed by TNO, a not for profit independent research organization. KH and AS were previously involved in a public private partnership that was 54% government funded and 46% funded by an alliance of Dutch Brewers.

MM was previously employed by the alcohol beverage industry-funded International Center for Alcohol Policies (ICAP) and International Alliance for Responsible Drinking (IARD). As an independent consultant, MM now supports the work of both industry and non-governmental organizations.

References

- Alcohol Research Group (2009). National Alcohol Survey (NAS12) 2009-2010. https://arg.org/wp-content/uploads/2015/04/N12_Ite mMatrix.pdf (6 June 2022, date last accessed).
- Ally AK, Lovatt M, Meier PS *et al.* (2016) Developing a social practicebased typology of British drinking culture in 2009-2011: implications for alcohol policy analysis. *Addiction* **111**:1568–79.
- American Psychiatric Association (APA). (2013) Diagnostic and Statistical Manual of Mental Disorders (DSM-5), 5th edn. Washington, DC: APA.
- Appleton A, James R, Larsen J. (2018) The association between mental wellbeing, levels of harmful drinking, and drinking motivations: a cross-sectional study of the UK adult population. *Int J Environ Res Public Health* 15:1333.
- Bonnell LN, Littenberg B, Wshah SR et al. (2020) A machine learning approach to identification of unhealthy drinking. J Am Board Fam Med 33:397–406.
- Bowden J, Delfabbro P, Room R *et al.* (2014) Alcohol consumption and NHMRC guidelines: has the message got out, are people conforming and are they aware that alcohol causes cancer? *Aust N Z J Public Health* 38:66–72.

- Bräker AB, Soellner R. (2016) Alcohol drinking cultures of European adolescents. *Eur J Public Health* 26:581-6.
- Breiman L. (2001) Random forests. Mach Learn 45:5-32.
- Bush K, Kivlahan DR, McDonell MB et al. (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 158:1789–95.
- Callinan S, Livingston M, Room R *et al.* (2016) Drinking contexts and alcohol consumption: how much alcohol is consumed in different Australian locations? *J Stud Alcohol Drugs* 77: 612–9.
- Callinan S, Mojica-Perez Y, Wright CJ et al. (2021) Purchasing, consumption, demographic and socioeconomic variables associated with shifts in alcohol consumption during the COVID-19 pandemic. Drug Alcohol Rev 40:183–91.
- Carvalho AF, Heilig M, Perez A et al. (2019) Alcohol use disorders. Lancet 394:781–92.
- Central Bureau of Statistics (CBS) (2021). Health Survey/Lifestyle Monitor 2014-2020. RIVM (Rijksinstituut voor Volksgezondheid en Milieu) and Trimbos-Instituut. https://www.rivm.nl/leefstijlmo nitor/alcoholgebruik-onder-volwassenen (6 June 2022, date last accessed).
- Collins RL, Marlatt GA. (1981) Social modeling as a determinant of drinking behavior: implications for prevention and treatment. *Addict Behav* 6:233–9.
- Crutzen R, Kuntsche E. (2013) Validation of the four-dimensional structure of drinking motives among adults. *Eur Addict Res* 19: 222-6.
- Cullum J, O'Grady M, Armeli S *et al.* (2012) The role of context-specific norms and group size in alcohol consumption and compliance drinking during natural drinking events. *Basic Appl Soc Psychol* **34**: 304–12.
- de Wit GA, van Gils PF, Over EAB et al. (2019) Social cost-benefit analysis of regulatory policies to reduce alcohol use in The Netherlands. Eur J Public Health 29:ckz185.794.
- de Wit H, Sayette M. (2018) Considering the context: social factors in responses to drugs in humans. *Psychopharmacology (Berl)* 235: 935–45.
- Dillard AJ, Ferrer RA, Bulthuis KRK *et al.* (2018) Positive excessive drinker prototypes predict greater drinking and alcohol problems. *Br J Health Psychol* 23:1000–20.
- European Commission. (2010) EU Citizens' Attitudes Towards Alcohol. Special Eurobarometer 331. Brussels: Directorate-General for Communications.
- Foster DW, Neighbors C, Krieger H. (2015) Alcohol evaluations and acceptability: examining descriptive and injunctive norms among heavy drinkers. Addict Behav 42:101–7.
- GENACIS (Gender, Alcohol, and Culture: An International Survey) (2018). Questionnaires. https://www.kettilbruun.org/projects/gena cis/11.html (6 June 2022, date last accessed).
- Genuer R, Poggi JM, Tuleau-Malot C. (2010) Variable selection using random forests. *Pattern Recogn lett* 31:2225–36.
- Gray HM, Wiley RC, Williams PM *et al.* (2021) A scoping review of "responsible drinking" interventions. *Health Commun* 36: 236–56.
- Health Council of the Netherlands. (2015) Dutch Dietary Guidelines 2015. The Hague, Health Council of the Netherlands, publication no. 2015/24E.
- Hibell B, Guttormsson U, Ahlström S et al. (2012) The 2011 ESPAD Report. Substance Use Among Students in 36 European Countries. Stockholm: Swedish Council for Information on Alcohol and Other Drugs (CAN).
- Kelly S, Olanrewaju O, Cowan A *et al.* (2018) Alcohol and older people: a systematic review of barriers, facilitators and context of drinking in older people and implications for intervention design. *PLoS One* 13:e0191189.

- Kilian C, Manthey J, Kraus L *et al.* (2021) A new perspective on European drinking cultures: a model-based approach to determine variations in drinking practices among 19 European countries. *Addiction* **116**:2016–25.
- Kuntsche E, Gabhainn SN, Roberts C et al. (2014) Drinking motives and links to alcohol use in 13 European countries. J Stud Alcohol Drugs 75:428–37.
- Kuntsche E, Kuntsche S. (2009) Development and validation of the Drinking Motive Questionnaire Revised short form (DMQ-R SF). J Clin Child Adolesc Psychol 38:899–908.
- Kuntsche E, Kuntsche S, Thrul J et al. (2017) Binge drinking: health impact, prevalence, correlates and interventions. Psychol Health 32: 976–1017.
- Lau-Barraco C, Linden-Carmichael AN, Hequembourg A et al. (2017) Motivations and consequences of alcohol use among heavy drinking nonstudent emerging adults. J Adolesc Res 32:667–95.
- Liaw A, Wiener M. (2002) Classification and regression by random Forest. *R News* 2:18–22.
- Livingstone AG, McCafferty S. (2015) Explaining reactions to normative information about alcohol consumption: a test of an extended social identity model. *Int J Drug Policy* **26**:388–95.
- Maani Hessari N, Petticrew M. (2018) What does the alcohol industry mean by 'Responsible drinking'? A comparative analysis. J Public Health 40:90–7.
- Monk RL, Qureshi A, Heim D. (2020) An examination of the extent to which mood and context are associated with real-time alcohol consumption. *Drug Alcohol Depend* **208**:107880.
- Moss AC, Albery IP. (2018) The science of absent evidence: is there such a thing as an effective responsible drinking message? *Alcohol* 53:26–30.
- National Survey on Drug Use and Health 2019. U.S. Department of Health and Human Services (2019). (NSDUH-2019-DS0001). https://datafiles.samhsa.gov/ (6 June 2022, date last accessed).
- Piano MR. (2017) Alcohol's effects on the cardiovascular system. Alcohol Res 38:219–41.
- R Development Core Team. (2019) R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.
- Roerecke M, Vafaei A, Hasan OSM *et al.* (2019) Alcohol consumption and risk of liver cirrhosis: a systematic review and meta-analysis. *Am J Gastroenterol* 114:1574–86.
- Ronksley PE, Brien SE, Turner BJ et al. (2011) Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. BMJ 342:d671.
- Rossow I. (2001) Alcohol and homicide: a cross-cultural comparison of the relationship in 14 European countries. Addiction 96:77–92.
- Schmidt H, Ter Bogt T, Godeau E et al. (2003) Drunkenness among young people: a cross-national comparison. J Stud Alcohol Drugs 64:650–61.
- Schouten F, Tuithof M, van Hasselt N et al. (2021) Kennismonitor Alcohol 2020. [Alcohol Knowledge Monitor 2020]. Utrecht: Trimbos Instituut.
- Shield K, Manthey J, Rylett M *et al.* (2020) National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. *Lancet Public Health* 5:e51–61.
- World Health Organization (WHO). (2010) Global Strategy to Reduce the Harmful Use of Alcohol. Geneva: World Health Organization.
- World Health Organization (WHO). (2018) Global Status Report on Alcohol and Health 2018. Geneva: World Health Organization.
- World Medical Association (WMA) (2018). WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects. https://www.wma.net/policies-post/wma-declaration-ofhelsinki-ethical-principles-for-medical-research-involving-humansubjects/ (6 June 2022, date last accessed).