A systematic review of the efficacy of alcohol warning labels

Hassan, Louise; Shiu, Edward

Journal of Social Marketing

DOI:
10.1108/JSOCM-03-2017-0020

Published: 01/01/2018

Peer reviewed version

A systematic review of the efficacy of alcohol warning labels: Insights from qualitative and quantitative research in the new millennium

Abstract

Purpose – The placement of warning labels on alcoholic beverages is a policy area with renewed interest, yet a strong evidence base regarding the efficacy of text-based or pictorial warning labels has still to emerge. Increased interest by policy makers has spurred research into potential alcohol warning label designs and messages. However, research which draws together findings from these recent studies is lacking.

Design/methodology/approach – The current study seeks to review research that has sought to examine the effectiveness of alcohol warning labels. Searches for English language articles (since 2000) using the terms 'alcohol' and 'warning label*' were conducted in 2015 across four databases (Web of Science, PubMed, PsycInfo and Cochrane). Articles were included if they empirically assessed the effectiveness and/or design of alcohol warning labels. A narrative analysis approach was used for the 15 articles identified.

Findings – Findings are reported on five themes covering the design of the warning including the use of imagery or recommendations followed by a focus on the warning messages and whether or not they are specific, use signal words and are based on qualitative or quantitative information.

Research limitations/implications – Overall there was little consistency in approach and measures, with very limited research having explored the potential of pictorial warning labels. Numerous research gaps are identified thus much more research is needed in this area. The evidence base is weak and caution is needed by policy makers regarding the introduction and implementation of alcohol warning labels. Limitations are discussed.

Originality/value – The review provides a timely up-to-date evaluation of the alcohol warning labels literature that has seen a recent resurgence but has not been critically reviewed.

Keywords: Alcohol warning labels, health information, alcohol policy, persuasive communications

Excessive alcohol consumption is a significant problem within societies as alcohol consumption and problem drinking cause 4.6% of all ill-health and premature deaths (Rehm *et al*., 2009). Furthermore, over 60 diseases and other types of trauma (excluding social and other population-level problems) have a causal link to alcohol use (Rehm *et al.*, 2003). Resultantly, governments have issued guidelines on the consumption of alcohol (e.g. UK: 2-3 units of alcohol per day; Australia: 2 standard drinks per day; USA: 3 drinks per day). However, consumers do exceed the low-risk governmental guidelines (see World Health Organisation: WHO, 2014) with many countries in Europe having the highest proportions of consumers who engage in heavy episodic drinking, defined by the WHO (2014) as drinking 60 or more grams of pure alcohol (7.5 units) during a single occasion on at least a monthly basis. As a result, rather than rely on commercial marketing interventions, governments need to find more effective means to tackle excessive consumption of alcohol (Wymer, 2011). One arm of an alcohol misuse prevention strategy is the introduction of health warnings on alcohol products. Alcohol warning labels are currently used in 31 countries (WHO, 2014). Other countries such as the UK are proposing to adopt such measures (see All Party Parliamentary Group on Alcohol Misuse Manifesto, 2015). Moreover, some countries (e.g. Thailand) have proposed legislation to implement a pictorial alcohol warning label policy (WHO, 2017). Nevertheless, there are continued discussions on the tension between the application of upstream and downstream measures in addressing such problems (Hoek and Jones, 2011).

Support for alcohol warning policies among citizens is high (Greenfield *et al.*, 2007; European Commission, 2010) yet some researchers have questioned the efficacy of alcohol warning labels and concluded that evidence of their influence on changing behaviour is very limited (e.g. Brennan *et al*., 2016; Coomber *et al*., 2015; Stockley, 2001). However, previous alcohol warning label reviews (e.g. Stockwell, 2006; Wilkinson and Room, 2009) have been based on studies conducted primarily to evaluate the USA warning label introduced in 1989.

This is a problem since the USA warning is very specific focusing on not drinking during pregnancy and alerting drinkers that alcohol impairs their ability to drive and operate machinery. The USA warning does state that consumption of alcoholic beverages may cause health problems but does not state which or to what extent the consumer may be at risk. Thus, the USA warning is limited and is not targeted at engendering responsible consumption beyond avoiding drinking alcohol in specific situations. There is thus a need to take account of evidence beyond that based on the USA warning by widening the coverage of studies that examined alcohol warnings across diverse message themes and formats.

The current research provides an up-to-date review of the findings reported and poses a research agenda for research on alcohol warning labels. Reviewing literature in the area of alcohol warning labels is important for two reasons. Firstly, researchers have cautioned against 'indiscriminate and uniformed use of warning messages' (Stewart and Martin, 1994, p. 13), therefore policy makers need to be appraised of the usefulness, limitations, and pitfalls in the use of warning labels. Secondly, Ringold (2002) reviewed potential unintended consequences of warnings finding that psychological reactance (a motivational state that occurs as a result of a freedom threat whereby the consumer feels a need to reassert control) is a commonly reported explanation for unintended effects. Such unintended consequences include the bolstering of positive attitudes towards drinking (Kozup *et al*., 2001).

The aim of this review is to evaluate more recent research (2000 to 2015) on the effectiveness of alcohol warning labels in order to provide an evidence base on factors that would contribute towards increased efficacy. Undertaking a systematic review makes a contribution to social marketing as it allows (1) the benchmarking of the current state of knowledge in the field and (2) highlights directions for future research. Furthermore, systematic reviews provide an evidence base to bridge the gap between theory and practice and as such make important contributions to the field of social marketing (Hastings, 2007).

Specifically, the objectives of this article are: to explore the key themes of the recent literature on alcohol warnings in the fields of marketing, public health, public policy, and psychology; to identify and discuss the central themes that impact the effectiveness of alcohol warnings in increasing consumers' acceptance of the warning message; and to make recommendations on future research directions. In this way, this article makes a contribution to social marketing literature by drawing together more recent findings and identifying the knowledge base on factors, regarding message design and format, that are central to the efficacy of alcohol warning labels. Social marketing has been identified as a useful approach in tackling individual behaviour in relations to alcohol (e.g., Gallopel -Morvan *et al*., 2011; Kubacki *et al*., 2015). Guided by sound social marketing principles, tools and techniques, and with a central focus on the target audience, such interventions may deliver innovative and lasting solutions to heterogeneous populations (Tapp and Rundle-Thiele, 2016). Gathering knowledge gained regarding the effective use of alcohol warning labels is timely as there are clear indications that governments are interested in research on this topic so as to inform them on the development of policies (e.g. European Commission, 2014; Rout and Hannan, 2016; WHO, 2009).

Alcohol warning labelling: Definition and efficacy assessment

A caution (e.g. drink responsibly) or a warning (e.g. may cause cancer) does not have a universally accepted definition but is usually associated with labelling that appears on a product, within product packaging, or at the product's place of use. A formal definition of a warning is provided by Lehto and Miller (1986, p.14) who stated that "warnings are those stimuli that alert people to hazardous conditions". A warning's principal function is information provision whereby those who pay attention to the warning are made aware of the potential risks, or about the magnitude of the risk, in using the product. Such information is intended to help the consumer weigh up whether or not they should use the product (e.g. take a medicine). Warnings may also include information about how to use the product safely

(e.g. operate a drill) or how to mitigate the risk associated with the use of the product (e.g. by keeping to a threshold). Beyond providing information, warnings have the potential to trigger behaviour change. However, the information needs to be processed and accepted in order that the consumer would, if pertinent, change behaviour to comply with the warning.

Based on information-processing frameworks, Argo and Main (2004) proposed five dimensions (attention, reading and comprehension, recall, judgments, and behavioural compliance) in assessing the efficacy of warning labels. Attention covers noticeability, awareness, attention, and recognition of the warning. Reading and comprehension relate to readability, understanding and comprehension of the warning message. Recall is a measure of the extent of consumers' recollection of the information contained in the message. Judgments reflect consumers' risk perceptions, believability, and attitudes towards complying with the message. Lastly, behavioural compliance examines the extent warning labels would engender compliance motivation that leads to behaviour change. The pertinence of these five dimensions in examining the efficacy of warning labels has been validated in more recent research (Purmehdi *et al*., 2017). Findings from the current study are discussed in relations to the Argo and Main (2004) framework specifically regarding the efficacy of warning labels in the discussion of the thematic areas identified in this review.

Methods

Only English language journal articles were considered covering the time period from January 2000 to September 2015 with the searches conducted during September/October 2015. Different search terms on the article full text were used across four databases: Web of Science (search term: alcohol AND "warning label*"); PubMed (search term: *alcohol* AND "*warning label*"); PsychInfo (search term: alcohol AND "warning label*"); and Cochrane (search term "alcohol warning label" OR "alcohol warning" OR "pictorial alcohol warning label" OR "graphic alcohol warning label" OR "alcohol warning message" OR "alcohol beverage

warning"). The databases searched were chosen because they had been used in prior systematic reviews on substance abuse issues (e.g., Lee et al., 2015; Loree et al., 2015). Journal articles that were available online (online first) up to the end of September 2015 were also included, resulting in 203 articles being identified containing 36 duplicates (see Figure 1 for details). Two additional articles were identified from the authors' knowledge of the literature. No contact was made with study authors to identify other publications or to gain additional information regarding their studies. The first author conducted the search and used article titles and abstracts to decide if an article fits the criteria of an empirical (qualitative or quantitative) study exploring or examining the effectiveness or design of alcohol warning labels. This led to the exclusion of 136 articles, 83 of which focused solely on tobacco warning labels with the remaining 53 articles having briefly mentioned alcohol warning labels without offering any discussion regarding their design or effectiveness. The remaining 33 eligible articles were subject to full-text examination independently by both authors. The focus and thus the inclusion criterion was that only studies that addressed the targeted individual consumer (consistent with downstream social marketing) were included. The full-text examination resulted in a further 18 articles being excluded leaving 15 articles for inclusion in the current review. A quality assessment of these 15 articles was undertaken independently by both authors based on the Mixed Methods Appraisal Tool (MMAT) developed by Pace et al. (2012). The MMAT is a tool for appraising the quality of studies across a broad range of methodologies (qualitative, quantitative, or mixed-method) on a small set of five generic criteria. The MMAT provides an overall quality assessment as well as the ability to look at specific aspects of potential weaknesses of prior studies dependent on the methods employed. This tool has been applied within other systematic reviews in the social marketing area (e.g. Almosa *et al*., 2017) and provides a quick quality assessment across multiple methodologies. The resulting MMAT

score for each article is given in Table 1 where the total highest possible quality score for an article is ****.

Results

Table 1 provides a brief outline of the 15 articles identified. Of the 16 studies reported in the 15 articles, seven were conducted in the USA, two in European countries (Germany and Luxembourg), with six conducted in Australia and one in New Zealand. The target participants covered young adults with five targeting students and two targeting adolescents, four articles targeting women, and the remaining four targeting the broader population. In terms of research method, five articles conducted experimental studies, eight implemented survey questionnaires, with five articles conducted focus group discussion – three of which were mixed-methods studies undertaken in conjunction with a survey. Of the five dimensions of effectiveness proposed by Argo and Main (2004), three articles (Parackal *et al*., 2010; Blume and Resor, 2007; Coomber *et al*., 2015) addressed the dimension of attention. Two articles addressed the dimensions of reading and comprehension (Kaskutas, 2000; Thomson *et al*., 2012) with one addressing recall (Blume and Resor, 2007). Nine articles addressed the dimension of judgments (Branco and Kaskutas, 2001; Creyer *et al*., 2002; Glock and Krolak-Schwerdt, 2013; Jones and Gregory, 2010; Kaskutas, 2000; Kozup *et al*., 2001; Krischler and Glock, 2015; Parackal *et al*., 2010; Pettigrew *et al*., 2014), with eight addressing behavioural compliance (Creyer *et al*., 2002; Glock and Krolak-Schwerdt, 2013; Jarvis and Pettigrew, 2013; Jones and Gregory, 2010; Kaskutas, 2000; Krischler and Glock, 2015; MacKinnon *et al*., 2000; MacKinnon *et al*., 2001). Across the nine articles addressing the dimension of judgments, a wide range of factors were examined covering - attitudes, believability, risk perceptions, information source, relevancy, convincingness, effectiveness and barriers, with the majority focused on attitudes and believability. The articles addressing behavioural

compliance covered outcomes including intentions, choice, drinking habits and patterns, drink size and overall alcohol consumption.

INSERT TABLE 1 HERE

Quality assessment

Table 1 also shows the results of the quality assessment of the 15 articles where only one achieved a 4* rating with seven rated as 3*, four as 2* and three as 1*. Agreement on quality ratings between the two authors was good (inter-rater reliability kappa = 0.82) based on the kappa statistic (Landis and Koch, 1977). In terms of the potential risk of bias within these articles, several issues are noted. Overall, studies failed to report information on the proportion of participants who agreed to take part against the total number of potential participants approached. The quantitative studies failed to take confounding factors and variables into account or to rule out order (carry-over) effects. Qualitative studies could have provided more detailed and precise information regarding their method of analysis and whether sampling was undertaken until redundancy was reached. Furthermore, qualitative studies failed to account for the role of the researcher in potentially influencing the research process. Lastly, mixed methods studies would benefit from a stronger justification and integration of the methods employed. At the review level, the set of 15 articles used diverse measures of the outcome variables (e.g. intentions to comply) and as a result introduced potential bias in this review because of a lack of ability to compare studies based the same outcome measures. Further, selective reporting of outcome measures within quantitative studies (e.g. failure to report non-significant results) may contribute toward additional bias. Lastly, none of the 15 articles reviewed captured an objective measure regarding behaviour, but relied on self-report

measures which have limitations because they might be biased towards under reporting due to socially desirable responding.

Initial findings on the warning messages employed in the studies

A number of themes regarding message framing and label design were identified from a narrative analysis of the tabular results. Studies that provided insights into more than one theme identified are discussed as appropriate. The themes were chosen to help identify significant areas of research endeavour and provide an opportunity to synthesize the research findings in order to identify research gaps. Therefore, the next section provides a discussion of the key findings along with the remaining gaps in knowledge. Where appropriate, studies published prior to the year 2000 are reflected on to provide either contrary or supporting evidence in the discussion of the findings from the 15 articles reviewed. But first it is important to understand the basic types of warnings that were used in the studies and if these are based on enacted policies.

Eight of the 15 articles reviewed were based on the USA warning label with some of these eight also contrasting the USA label against alternative designs. The remaining seven articles show diverse warning designs discussed further below. Alcohol warning labels may contain messages that target different themes such as health, social, or other negative consequences associated with alcohol consumption. Glock and Krolak-Schwerdt (2013) explored university students' reactions to five health-related warnings (e.g. "Alcohol damages your liver") in contrast to five 'social' warnings that countered expectations of positive social and relaxing utilities of alcohol (e.g. "Alcohol makes you feel alone"). Their findings showed that explicit attitudes towards drinking alcohol did not differ across the two (health vs. social) warning types. However, implicit attitudes differed with participants in the 'social' warning label condition reporting less positive (post- vs. pre- exposure) implicit attitudes towards alcohol, but those in the health warning label condition reporting more positive (post- vs. pre- exposure)

implicit attitudes. A study by Jarvis and Pettigrew (2013) explored negatively and positively framed messages (across both health and drink driving contexts) and found that for those who report higher consumption of alcohol, negative health messages ("Every drink of alcohol harms your brain") had the highest utility. Whereas a positive message about drink driving ("Make sure you're okay to drive") generated a boomerang effect that led to higher likelihood of choosing the alcohol beverage. Alcohol warning label messages may also vary in terms of tone. Thomson *et al*. (2012) compared consumer responses to serious and humorous alcohol warning labels of various formats. Their findings suggest that humorous messages are not effective because of a lack of comprehension of the core message by participants. Warnings can also be cast either as statements or posed as questions. Krischler and Glock (2015) evidenced experimentally that young adults respond better to warnings formulated as questions, yielding a significant increase in negative outcome expectancies related to alcohol consumption, but with no effect on drinking intentions.

Key thematic areas

The analysis of the articles resulted in five thematic areas pertaining to the general themes or styles used in designing the warning. These five areas are firstly, the use of imagery, secondly, the use of recommendations, with the remaining three areas focusing on the message contained within the warning regarding how specific the warning is, whether or not signal words are used, and the use of qualitative or quantitative message information.

Use of imagery. Jones and Gregory (2010) used focus groups to explore views on alcohol warning labels amongst students and found pictorial warnings to be more effective than text-only warnings. Coomber *et al*. (2015) examined the efficacy of the mainly text-based Australian warning label and recommend highly visible graphic warning labels be used to deliver behaviour change. Krischler and Glock (2015) used pictorial warnings in combination

with text to assess the effect of the text framed as a statement as opposed to a question, but did not examine the role of the pictorial part of their label. Overall, there is a need for research that compares text-based against pictorial messages in alcohol warning labels. Some evidence is available regarding other forms of warnings rather than labels. For example, in counter-alcohol research examining the effects of pairing text with pictorial images, Collymore and McDermott (2016) examined social and health (text with pictorial images) messages across gain, fear-loss, and disgust-loss conditions. Participants were shown a booklet containing six photographs each accompanied with a health persuasion message. These authors found that loss-framed messages (and in particular the health disgust-loss with a message about drinking more than government daily limits results in temporary facial imperfections) were most effective in increasing intentions to reduce alcohol consumption. Brown and Locker (2009) conducted a study to examine differences in responses to emotive (e.g. severe health consequences) and non-emotive imagery (e.g. intoxicated individuals) anti-alcohol messages in pamphlets. Their findings suggest that distressing (highly emotive) images might lead to defensive avoidance responses resulting in lower perceived risk estimates for participants vulnerable to alcohol-related problems. At present, limited research findings exist to shed light on the use of graphic imagery in alcohol warning labels per se. Research is also needed to assess different levels of "graphicness" similar to that undertaken in tobacco control (e.g. Kees *et al*., 2010). Lessons from tobacco control may inform research into the efficacy of including graphic images into alcohol warning labels. In particular, the 2007 WHO expert committee on problems related to alcohol consumption concluded that current alcohol warnings labels on containers do not lead to behaviour change. The committee noted that lessons can be learned from tobacco control policies where more graphic and larger warnings have affected behaviour (WHO, 2007).

Use of recommendations. Governments and health charities generally provide guidelines and recommendations on most aspects of health, including for example physical activity (e.g. WHO, 2010), healthy eating (e.g. NHS, 2016), healthy lifestyle (e.g. American Heart Association, 2016) and low risk alcohol consumption (e.g. Australian Department of Health, 2013). Pettigrew *et al*. (2014) evaluated general and cancer-related specific alcohol warning messages and found that two positively worded messages both phrased as a recommendation to reduce drinking alcohol were rated by respondents as more believable than negatively worded messages that relied on fear arousal and did not provide a recommendation. However, earlier research contradicts these findings. Specifically, Slater *et al*. (1998) examined the influence of providing a recommendation (in addition to a warning) and found no effect on the number of negative responses or risk perceptions across three messages (drink driving, health, alcohol & drugs). Effects on believability and number of positive responses were evidenced where the health and alcohol & drugs messages were more believable with a recommendation, yet an opposite effect was found for the drink driving message. Thus, the decision on providing a recommendation may depend on the message theme. Overall, limited research addresses the efficacy of recommendations and there is a need for more research regarding the influence of providing a recommendation across different message themes. Relatedly, limited research examined the impact of counterbalancing health claims on consumer attitudes and perceptions of disease-related risks. Kozup *et al*. (2001) found boomerang effects among drinkers and non-drinkers. Specifically, these authors found that in the absence of a health claim, wine drinkers had more favourable attitudes and risk perceptions regarding wine when a warning was present. Their results also showed that the presence of a health claim induced a boomerang effect among non-drinkers leading to more favourable attitudes and lower risk perceptions when the warning was also present. There is thus a need for future research to examine beyond just the ability of recommendations (e.g. UK 2-3 units per day) in influencing the efficacy of alcohol warning

labels. Given Kozup *et al*.'s (2001) findings, research should examine if and how positive marketing messages might interact with proposed government alcohol warning labels.

Message specificity. Pettigrew *et al*. (2014) found that, compared with warnings about specific types of cancer, general warnings about cancer were perceived as more believable, convincing, and personally relevant. Creyer *et al*. (2002) tested the USA warning against "Alcohol is a drug", in the USA and in Australia, and found the "Alcohol is a drug" warning led to greater risk perceptions than the USA warning. Creyer *et al*.'s (2002) results were explained by the associative memory perspective which suggests that because the word drug is associated in memory with negative concepts, such as addiction, presenting alcohol as a drug creates a link to these more negative concepts. Yet it can be argued that the "Alcohol is a drug" message is less specific than the USA warning (that targets pregnancy and driving). Nevertheless, Kaskutas (2000) showed that a large majority (over 85%) of high-risk urban minority women found the USA warning to be both believable and understandable though low in persuasiveness. Targeted and specific warning labels may be well received and potentially beneficial. For example, Parackal *et al*. (2010) found pregnant New Zealand women view alcohol warning labels as a source of information on the risks associated with drinking during pregnancy. However, Branco and Kaskutas (2001) found that warning labels that employ scare tactics can be perceived as overstating the risks and not as believable. Overall, more evidence is needed in comparing across types of message while taking into account of other design features such as the use of signal word. On balance, research so far indicates that less specific messages might work better in encouraging processing of the message but it is not clear which type of message might work best in achieving behavioural compliance.

Use of signal word(s) or qualifier(s). Signal words (e.g. health warning; government warning) and qualifiers (e.g. may cause) are inherent in the design of many warning messages. Thomson et al. (2012) undertook six focus groups predominantly with young people (two also

with parents of teenagers) and found that participants were more likely to accept the message if "Health Warning" rather than "Warning" or "Government Health Warning" was used. However, prior experimental research (Wogalter *et al*., 1999) that compared the use of different signal sources on consumers' perception of credibility and likelihood of compliance found that signals with a specific source (e.g. from medical/health bodies or government agencies) were more credible and more likely to be complied with than the less specific signals ("Warning"; "Health Warning"; "Important Health Warning"). Thus, there is some debate around the use of signal words. Regarding qualifiers, Pettigrew *et al*. (2014) compared warnings with the wording 'increases risk' versus 'can cause' and found that the 'increases risk' wording was more convincing and more believable for females than the 'can cause' wording. However, prior research on the student population found that the use of qualifiers such as 'may cause cancer' was associated with less avoidance of the message (MacKinnon *et al*., 1994). Thus, different target populations may perceive the use of signal words and qualifiers differently and further research is needed to clarify the effective use of signal words and qualifiers.

Use of qualitative versus quantitative information. Prior research has examined the use of quantitative information in alcohol warnings. Pettigrew *et al*. (2014) included two negative quantitative messages ("Alcohol causes around 5000 new cases of cancer each year" and "Alcohol causes 1 in 20 cancer deaths") to evaluate against ten other messages (e.g. "Alcohol increases your risk of cancer"; "Reduce your drinking to reduce your risk of cancer") and found that the quantitative messages performed poorly in terms of believability and were perceived differently across gender. Slater *et al*. (1998) found that recall of the message was improved with the use of quantitative (as opposed to qualitative) information in the warning label. However, Slater et al.'s research also found that quantitative information was less likely to be believed for two of the three messages tested (the effects of drinking on driving and on health). Overall, the limited evidence suggests that quantitative messages are of less utility than

qualitative messages but future research should test this tentative conclusion across a range of message types and frames.

Are alcohol warning labels effective?

Overall, findings are mixed on the efficacy of warnings with Scholes-Balog *et al*. (2012) reporting that beyond the adult population, alcohol warnings have little efficacy in affecting beliefs, risk perceptions, or alcohol consumption in adolescents. Coomber *et al*. (2015) further concluded that "current warning labels fail to effectively transmit health messages to the general public" (p. 816). In particular, limited research has examined the long-term effects of exposure to alcohol warning labels on consumption. MacKinnon *et al*. (2001) conducted a three-year longitudinal study on 649 students recruited initially when they were in their ninth or tenth grade. These authors found that the USA alcohol warning label is effective in informing and reminding students of the risks associated with alcohol consumption. However, the study failed to evidence a significant effect of the warning on consumption levels. Further, MacKinnon *et al*. (2000) in their five-year longitudinal study tracking over 30,000 tenth and twelfth grade students found no evidence of beneficial changes in alcohol related beliefs, consumption, or drink driving that might be attributable to the USA warning label. Interestingly, Blume and Resor (2007) found that awareness of nutritional information on warning labels is associated with reduced alcohol consumption during pregnancy, suggesting that additional insight is needed regarding the message and the design of alcohol warning labels for at least some targeted consumer groups.

Considering if warning labels are more or less effective for specific targeted audiences, Thomson *et al*. (2012) found that messages that were matched to gender and type of drink (e.g. wine, beer, vodka) were more relevant and acceptable to consumers. Further support for the need to take gender into account comes from Jarvis and Pettigrew (2013) who explored health

outcomes and drink driving warnings using a choice task and found that the messages with the greatest utility differed across gender. In a later study, Jarvis *et al*. (2015) also found gender to be a significant covariate in an alcohol choice task. Overall, the literature shows that gender needs to be considered carefully in the design of warning messages and that more research is also needed to assess the interaction between the type of drink and the warning message. In reporting results, studies that focus on placing messages on a particular drink need to be cautious in generalizing their results to other types of alcoholic beverage. Future research should examine the fit between the warning message and the type of drink across gender.

Examining the effect of alcohol warning labels across drinking status, Creyer *et al*. (2002) in their experimental study found main effects of drinking status (e.g., an abstainer, moderate drinker, heavy drinker) on attitudes, beliefs, and behavioural intentions, but did not find evidence that these (main) effects are affected by the type of warning (USA warning versus "Alcohol is a drug") shown. However, Brown and Locker (2009) in their experimental study did evidence a three-way interaction effect of warning type (high/low emotive images), level of denial and drinking status on risk estimates. Further, Coomber *et al*. (2017) in their large-scale online survey found the current Australian alcohol warning label failed to influence high-risk drinkers. Future research is needed to examine if drinking status should also be taken into account in assessing the efficacy of alcohol warning labels. Arguably, there is an imperative to develop and implement warnings that would benefit those who are in greater need of a behaviour change.

Most studies focus on judgements about the warnings or inclinations to behave in line with the warning. Many countries have alcohol warning label policies in place and the efficacy of these are currently unknown. Assessment of the efficacy of warning labels policies beyond the USA would allow a renewed focus on the dimensions of awareness and recall, as well as a stronger assessment of the key dimension of behavioural compliance in Argo and Main's

(2004) framework. Researchers should continue to explore different message designs and themes as well as examine potential mediators (e.g. perceived susceptibility) and moderators (e.g. believability) of the effects of messages on behavioural compliance. There is also a need to assess the potential for defensive reactions and other boomerang effects. Overall, the reviewed studies do not provide clear evidence to advocate a particular type of message theme or framing because little consistent findings exist. Further, without replication studies, confidence in the knowledge gained from past studies is limited. To deliver a stronger evidence base, this research field needs more studies that will in time allow a meta-analysis to be conducted to enable a more rigorous assessment of the utility of different types of messages. Lastly, most studies focus on judgements about the message but do not investigate factors such as the placement of the message on drinks, issues such as the length of the messages or use of boarders, or how pictorials or pictograms and text work together to be effective. As a result, there are many underexplored avenues for future research (see Table 2 for a summary).

INSERT TABLE 2 HERE

Conclusion

Unfortunately, large scale experimental studies on alcohol warnings do not exist and as a result at this point in time there is not enough evidence on which alcohol warning labels are more/less effective on behavioural compliance. Furthermore, the literature is only beginning to turn its attention to the potential of pictorial warnings and alternative warning label message themes and designs. Given limited current evidence in support of pictorial alcohol warning labels, the temptation is to generalize the findings of research on the use of graphic warnings within tobacco control, leading to the adoption of similar graphic warnings targeted at alcohol. Even though there is now a substantial evidence base (e.g. Hammond, 2011) affirming the ability of

highly graphic warnings in inducing fear among smokers thus leading to greater quit intentions, the findings for alcohol may well differ. One reason is that there is no longer a supporting argument regarding any health benefits arising from smoking, yet there exists public perceptions and continual debates regarding positive alcohol-related health claims, at least for moderate wine consumption (e.g. Chiva-Blanch *et al*., 2013; Kozup *et al*., 2001; Teresserra-Rimbau *et al*., 2015; Vecchio *et al*., 2017). Thus, alcohol warning label research needs to take a cautious stand when adopting procedures and findings from tobacco research. One such research is that of Brown and Locker (2009). Based on similar images used in past smoking studies, these authors adopted graphic and unsettling medical images in their study on alcohol warnings. Their results showed that such negative images can provoke defensive avoidance responses thus decreasing the efficacy of these graphic alcohol warnings. Further, Zahra *et al*. (2015) also urged caution in the use of negative imagery for alcohol warnings. Although their research was not focused on warning labels, these authors found positive (not negative) images related to consequences of abstinence are more effective in inducing accurate appraisals of the warning message. Other research has identified potential maladaptive reactions to fear appeals with a recommendation to take into consideration the target's level of self-efficacy and the role of self-affirmation (Awagu and Basil, 2016).

Many alcohol warnings are based on the assumption that inducing higher risk perceptions will reduce consumption. However, policy makers and social marketers should refrain from simply escalating risk perceptions in drinkers because, not least, ethical guidelines on health promotion emphasize informed choice (Turoldo, 2009). Furthermore, boundaries dividing safe and dangerous drinking are unclear and multifaceted (Rehm *et al*., 2008). Thus, the development of effective alcohol warning labels requires a nuanced and sound understanding of how drinkers think about risk. In particular, research on tobacco control (e.g. MacKinnon *et al*., 2002; Emery *et al*., 2014) has indicated the need to examine mediation designs to test the

theoretical basis regarding the underlying process for which warning labels might exert their effects on outcomes. Similarly, in line with sound social marketing principles, the development of theory-based overarching conceptual models of drinkers' engagement with such warnings will be beneficial to alcohol warning label research (Hastings, 2007).

Finally, given the lack of clear evidence in support of the efficacy of alcohol warnings, policy makers and social marketing researchers must bear in mind that individuals are embedded in an ecological system. As such, policy interventions must take into account pertinent historical, social, physical and environmental contexts (Brennan *et al*., 2016). This review has revealed a weak and fragmented evidence base to guide policy makers or practitioners. There is thus a need for policy makers and social marketing practitioners to take a cautious path in implementing alcohol warning labels in their attempt to effect positive behaviour change. In particular, policy makers should consider funding quality research that would provide clearer guidelines on whether and what alcohol warning label policy would deliver the desired downstream effects. Moreover, alcohol warning labels might serve as an effective communication strategy with an ability to deliver knock-on benefits. For example, the public becoming more aware of the dangers of alcohol consumption might lead to a greater acceptance of related policies such as minimum pricing and other control measures. Because diverse results have been reported across the article reviewed, care must be taken in the development and piloting of such interventions. In particular, social marketers must bear in mind the overall marketing mix, the need for segmentation and targeting, as well as the need to consider the exchange regarding what consumers gain and what they give up from making the behaviour change. Furthermore, there is a strong need for an early evaluation of any such intervention to ascertain that the desired behavioural change has been achieved without unintended consequences (such as boomerang effects).

Nevertheless, our review highlights the potential for warning labels to be acceptable to consumers as an information source. Furthermore, on-product warnings are practical in that they are viewed by users at a time close to consumption. As such, they are likely to exert greater influence when compared against alternative means of information provision (such as leaflets or other marketing communications). Warning labels, however, should not be the only strategy employed and that governments should consider other policy directions such as minimum pricing in the sale of alcohol products.

Limitations of the review

The current review has a number of limitations that would qualify the findings reported. Only English language journal articles were eligible for inclusion and thus the review failed to consider contributions from other sources and grey literature and thus may suffer from publication bias. The search strategy adopted was limited in terms of the number of databases searched as well as the search terms used. Furthermore, the inclusion criteria adopted focused solely on downstream social marketing and excluded studies that did not target the consumer directly. Further, no contact was made with authors in the field to get additional information on the studies reported. The paucity of studies compounded by the lack of consistency regarding outcome measures meant that a meta-analysis could not be undertaken.

References

Almosa, Y., Parkinson, J., and Rundle-Thiele, S. (2017), "Littering Reduction: A Systematic Review of Research 1995–2015", *Social Marketing Quarterly*, Vol. 23 No. 3, pp. 203-222.

All Party Parliamentary Group on Alcohol Misuse Manifesto (2015), available at: http://www.alcoholconcern.org.uk/wp-content/uploads/2014/10/APPG_Manifesto.pdf (accessed 5 December 2016).

American Heart Association (2016), available at: http://www.heart.org/HEARTORG/HealthyLiving/HealthyEating/Nutrition/The-American-Heart-Associations-Diet-and-Lifestyle-Recommendations_UCM_305855_Article.jsp#.V1V7T4-cH5o (accessed 5 December 2016).

Argo, J.J. and Main, K.J. (2004), "Meta-analyses of the effectiveness of warning labels", *Journal of Public Policy & Marketing*, Vol. 23 No. 2, pp. 193-208.

Australian Department of Health (2013) Reduce your risk: new national guidelines for alcohol consumption, available at: http://www.alcohol.gov.au/internet/alcohol/publishing.nsf/Content/guide-adult (accessed 5 December 2016).

Awagu, C. and Basil, D.Z. (2016), "Fear appeals: the influence of threat orientations", *Journal of Social Marketing*, Vol. 6 No. 4, pp. 361-376.

Blume, A.W. and Resor, M.R. (2007), "Knowledge about health risks and drinking behavior among Hispanic women who are or have been of childbearing age", *Addictive Behaviors*, Vol. 32 No. 10, pp. 2335-2339.

Branco, E.I. and Kaskutas, L.A. (2001), ""If it burns going down…": How focus groups can shape fetal alcohol syndrome (fas) prevention", *Substance Use and Misuse*, Vol. 36 No. 3, pp. 333-345.

Brennan, L., Previte, J., and Fry, M.-L. (2016), "Social marketing's consumer myopia: Embracing a systems view of social change markets", *Journal of Social Marketing*, Vol. 6 No. 3, pp. 219-239.

Brown, S. and Locker, E. (2009), "Defensive responses to an emotive anti-alcohol message", *Psychology & Health*, Vol. 24 No. 5, pp. 517-528.

Chiva-Blanch, G., Arranz, S., Lamuela-Raventos, R. M., and Estruch, R. (2013), "Effects of wine, alcohol and polyphenols on cardiovascular disease risk factors: evidences from human studies", *Alcohol and alcoholism*, Vol. 48 No. 3, pp. 270-277.

Collymore, N.N. and McDermott, M.R. (2016), "Evaluating the effects of six alcohol-related message frames on emotions and intentions: the neglected role of disgust", *Journal of Health Psychology*, Vol. 21 No. 9, pp. 1907-1917.

Coomber, K., Martino, F., Barbour, I.R., Mayshak, R., and Miller, P.G. (2015), "Do consumers 'Get the facts'? A survey of alcohol warning label recognition in Australia", *BMC Public Health*, Vol. 15 No. 1, pp. 816-824.

Coomber, K., Jones, S.C., Martino, F., and Miller, P.G. (2017), "Predictors of awareness of standard drink labelling and drinking guidelines to reduce negative health effects among Australian drinkers", *Drug and Alcohol Review*, Vol. 36 No. 2, pp. 200-209.

Creyer, E.H., Kozup, J.C., and Burton, S. (2002), "An experimental assessment of the effects of two alcoholic beverage health warnings across countries and binge-drinking status", *Journal of Consumer Affairs*, Vol. 36 No. 2, pp. 171-202.

Emery, L.F., Romer, D., Sheerin, K.M., Jamieson, K.H., and Peters, E. (2014), "Affective and cognitive mediators of the impact of cigarette warning labels", *Nicotine and Tobacco Research*, Vol. 16 No. 3, pp. 263-269.

European Commission (2010), EU citizens' attitudes towards alcohol, available at: http://ec.europa.eu/public_opinion/archives/ebs/ebs_331_en.pdf (accessed 10 December 2016).

European Commission (2014), State of play in the use of alcoholic beverage labels to inform consumers about health aspects: Action to prevent and reduce harm from alcohol, available at: https://ec.europa.eu/health//sites/health/files/alcohol/docs/alcohol_beverage_labels_full_report_en.pdf (accessed 10 September 2017).

Gallopel-Morvan, K., Gabriel, P., Le Gall-Ely, M., Rieunier, S., and Urien, B. (2011), "The use of visual warnings in social marketing: The case of tobacco", *Journal of Business Research*, Vol. 64 No. 1, pp. 7-11.

Glock, S. and Krolak-Schwerdt, S. (2013), "Changing outcome expectancies, drinking intentions, and implicit attitudes toward alcohol: a comparison of positive expectancy-related and health-related alcohol warning labels", *Applied Psychology: Health and Well Being*, Vol. 5 No. 3, pp. 332-347.

Greenfield, T.K., Ye, Y, and Giesbrecht, N.A. (2007), "Views of alcohol control policies in the 2000 National Alcohol Survey: what news for alcohol policy development in the US and its States?" *Journal of Substance Use*, Vol. 12 No. 6, pp. 429-445.

Hammond, D. (2011), "Health warning messages on tobacco products: a review", *Tobacco Control*, Vol. 20, pp. 327-337.

Hastings, G. (2007), *Social Marketing: Why should the devil have all the best tunes?* Elsevier/Butterworth-Heinemann, Amsterdam, London, Boston.

Hoek, J. and Jones, S. C. (2011), "Regulation, public health and social marketing: a behaviour change trinity", *Journal of Social Marketing*, Vol. 1 No. 1, pp. 32-44.

Jarvis, W., Pettigrew, S., and Olaru, D. (2015), "Alcohol warning statement effectiveness under different alcohol category conditions", in Kubacki, K. (Ed) *Ideas in Marketing: Finding the New and Polishing the Old. Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, Springer, Cham, pp. 243-252.

Jarvis, W. and Pettigrew, S. (2013), "The relative influence of alcohol warning statement type on young drinkers' stated choices", *Food Quality and Preference*, Vol. 28 No. 1, pp. 244-252.

Jones, S.C. and Gregory, P. (2010), "Health warning labels on alcohol products - the views of Australian university students", *Contemporary Drug Problems*, Vol. 37 No. 1, pp. 109-137.

Kaskutas, L.A. (2000), "Understanding drinking during pregnancy among urban American Indians and African Americans: health messages, risk beliefs, and how we measure consumption", *Alcoholism: Clinical and Experimental Research*, Vol. 24 No. 8, pp. 1241-1250.

Kees, J., Burton, S., Andrews, J.C., and Kozup, J. (2010), "Understanding how graphic pictorial warnings work on cigarette packaging", *Journal of Public Policy and Marketing*, Vol. 29 No. 2, pp. 265-276.

Kozup, J., Burton, S., and Creyer, E. (2001), "A comparison of drinkers' and nondrinkers' responses to health-related information presented on wine beverage labels", *Journal of Consumer Policy*, Vol. 24 No. 2, pp. 209-230.

Krischler, M. and Glock, S. (2015), "Alcohol warning labels formulated as questions change alcohol-related outcome expectancies: a pilot study", *Addiction Research and Theory*, Vol. 23 No. 4, pp. 343-349.

Kubacki, K., Rundle-Thiele, S., Pang, B., and Buyucek, N. (2015), "Minimizing alcohol harm: A systematic social marketing review (2000–2014)", *Journal of Business Research*, Vol. 68 No. 10, pp. 2214-2222.

Landis, J. R., and Koch, G. G. (1977), "The measurement of observer agreement for categorical data", *biometrics*, Vol. 33 No. 1, pp. 159-174.

Lee, N.K., Cameron, J., and Jenner, L. (2015), "A Systematic Review of Interventions for Co-occurring Substance Use and Borderline Personality Disorders", *Drug and Alcohol Review*, Vol. 34 No. 6, pp. 663–672.

Lehto, M.R. and Miller, J.M. (1986), *Warnings: Fundamentals, Design, and Evaluation Methodologies*, Fuller Technical Publications, Ann Arbor, MI.

Loree, A.M., Lundahl, L.H., and Ledgerwood, D.M. (2015), "Impulsivity as a Predictor of Treatment Outcome in Substance Use Disorders: Review and Synthesis", *Drug and Alcohol Review*, Vol. 34 No. 2, pp. 119–134.

MacKinnon, D.P., Nemeroffa, C., and Nohre, L. (1994), "Avoidance responses to alternative alcohol warning labels", *Journal of Applied Social Psychology*, Vol. 24 No. 8, pp. 733-753.

MacKinnon, D.P., Nohre, L., Cheong, J.W., Stacy, A.W., and Pentz, M.A. (2001), "Longitudinal relationship between the alcohol warning label and alcohol consumption", *Journal of Studies on Alcohol*, Vol. 62 No. 2, pp. 221-227.

MacKinnon, D.P., Nohre, L., Pentz, M.A., and Stacy, A.W. (2000), "The alcohol warning and adolescents: 5-year effects", *American Journal of Public Health*, Vol. 90 No. 10, pp. 1589-1594.

MacKinnon, D.P., Taborga, M.P., and Morgan-Lopez, A.A. (2002), "Mediation designs for tobacco prevention research", *Drug and Alcohol Dependence*, Vol. 90, pp. S69-S83.

NHS (2016) Healthy eating, available at: http://www.nhs.uk/LiveWell/healthy-eating/Pages/Healthyeating.aspx (accessed 5 January 2017).

Pace, R., Pluye, P., Bartlett, G., Macaulay, A. C., Salsberg, J., Jagosh, J., and Seller, R. (2012), "Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review", *International journal of nursing studies*, Vol. 49 No. 1, pp. 47-53.

Parackal, S.M., Parackal, M.K., and Harraway, J.A. (2010), "Warning labels on alcohol containers as a source of information on alcohol consumption in pregnancy among New Zealand women", *International Journal of Drug Policy*, Vol. 21 No. 4, pp. 302-305.

Pettigrew, S., Jongenelis, M., Chikritzhs, T., Slevin, T., Pratt, I.S., Glance, D., and Liang, W. (2014), "Developing cancer warning statements for alcoholic beverages", *BMC Public Health*, Vol. 14 No. 1, pp. 786-795.

Purmehdi, M., Legoux, R., Carrillat, F., and Senecal, S. (2017), "The effectiveness of warning labels for consumers: A meta-analytic investigation into their underlying process and contingencies", *Journal of Public Policy & Marketing*, Vol. 36 No. 1, pp. 36-53.

Rehm, J., Room, R., and Taylor, B. (2008), "Method for moderation: Measuring the risk of alcohol-attributable mortality as a basis for drinking guidelines", *International Journal of Methods in Psychiatric Research*, Vol. 17 No. 3, pp. 141-151.

Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y., and Patra, J. (2009), "Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders", *The Lancet*, Vol. 373 No. 9682, pp. 2223-2233.

Rehm J., Room R., Monteiro M., Gmel, G., Graham, K., Rehn, N., Sempos, C.T., and Jernigan, D. (2003), "Alcohol as a risk factor for the global burden of disease", *European Addiction Research*, Vol. 9 No. 4, pp. 157-164.

Ringold, D.J. (2002), "Boomerang effects in response to public health interventions: Some unintended consequences in the alcoholic beverage market", *Journal of Consumer Policy*, Vol. 25 No. 1, pp. 27-63.

Rout, J. Hannan, T. (2016), "Consumer awareness and understanding of alcohol pregnancy warning labels", *Wellington: Health Promotion Agency*, available at: https://www.hpa.org.nz/sites/default/files/Consumer%20awareness%20alcohol%20pregnancy%20warning%20label%20report%20FINAL.pdf (accessed 11 September 2017).

Scholes-Balog, K.E., Heerde, J.A., and Hemphill. S.A. (2012), "Alcohol warning labels: unlikely to affect alcohol related beliefs and behaviours in adolescents", *Australian and New Zealand Journal of Public Health*, Vol. 36 No. 6, pp. 524-529.

Slater, M.D., Karan, D., Rouner, D., Murphy, K., and Beauvais, F. (1998), "Developing and assessing alcohol warning content: responses to quantitative information and behavioral recommendations in warnings with television beer advertisements", *Journal of Public Policy & Marketing*, Vol. 17 No. 1, pp. 48-60.

Stewart, D.W. and Martin, I.M. (1994), "Intended and unintended consequences of warning messages: A review and synthesis of empirical research", *Journal of Public Policy & Marketing*, Vol. 13 No. 1, pp. 1-19.

Stockley, C.S. (2001), "The effectiveness of strategies such as health warning labels to reduce alcohol-related harms — an Australian perspective", *International Journal of Drug Policy*, Vol. 12 No. 2, pp. 153-166.

Stockwell, T.R. (2006), *A review of research into the impacts of alcohol warning labels on attitudes and behaviour*, British Columbia, Canada: University of Victoria, Centre for Addictions Research of BC.

Tapp, A., and Rundle-Thiele, S. (2016), Social marketing and multidisciplinary behaviour change. In F. Spotswood (Ed), *Beyond behaviour change: Key issues, interdisciplinary approaches and future directions*, Bristol, CT: Policy Press.

Thomson, L.M., Vandenberg, B., and Fitzgerald, J.L. (2012), "An exploratory study of drinkers views of health information and warning labels on alcohol containers", *Drug and Alcohol Review*, Vol. 31 No. 2, pp. 240-247.

Tresserra-Rimbau, A., Medina-Remón, A., Lamuela-Raventós, R.M., Bulló, M., Salas-Salvadó, J., Corella, D., Fitó, M., Gea, A., Gómez-Gracia, E., Lapetra, J. and Arós, F. (2015), "Moderate red wine consumption is associated with a lower prevalence of the metabolic syndrome in the PREDIMED population", *British Journal of Nutrition*, Vol. 113(S2), pp. S121-S130.

Turoldo, F. (2009), "Responsibility as an ethical framework for public health interventions", *American Journal of Public Health*, Vol. 99 No. 7, pp. 1197-1202.

Vecchio, R., Decordi, G., Grésillon, L., Gugenberger, C., Mahéo, M., and Jourjon, F. (2017), "European consumers' perception of moderate wine consumption on health", *Wine Economics and Policy*, Vol. 6, pp. 14-22.

Wilkinson, C. and Room, R. (2009), "Warnings on alcohol containers and advertisements: International experience and evidence on effects", *Drug and Alcohol Review*, Vol. 28 No. 4, pp. 426-435.

Wogalter, M.S., Kalsher, M.J., and Rashid, R. (1999), "Effect of signal word and source attribution on judgments of warning credibility and compliance likelihood", *International Journal of Industrial Ergonomics*, Vol. 24 No. 2, pp. 185-192.

World Health Organization (2007) WHO Expert Committee on Problems Related to Alcohol Consumption. Second report, WHO Technical Report Series No. 944. Geneva: WHO,

2007, available at: www.who.int/substance_abuse/expert_committee_alcohol_trs944.pdf (accessed 10 December 2016).

World Health Organization (2009) Evidence for the effectiveness and cost–effectiveness of interventions to reduce alcohol-related harm, available at: http://www.euro.who.int/en/publications/abstracts/evidence-for-the-effectiveness-and-costeffectiveness-of-interventions-to-reduce-alcohol-related-harm-2009 (accessed 11 September 2017).

World Health Organization (2010) Global recommendations on physical activity for health, available at: http://www.who.int/dietphysicalactivity/publications/9789241599979/en/ (accessed 6 February 2017).

World Health Organization (2014) Global status report on alcohol and health 2014, available at: http://apps.who.int/iris/bitstream/10665/112736/1/9789240692763_eng.pdf?ua=1 (accessed 10 February 2017).

World Health Organization (2017) *Bulletin of the World Health Organization* 2017; 95:487-488, available at: http://dx.doi.org/10.2471/BLT.17.020717 (accessed 10 September 2017).

Wymer, W. (2011), "Developing more effective social marketing strategies", *Journal of Social Marketing*, Vol. 1 No. 1, pp. 17-31.

Zahra, D. Monk, R.L. and Corder, E. (2015), "'IF you drink alcohol, THEN you will get cancer': Investigating how reasoning accuracy is affected by pictorially presented graphic alcohol warnings", *Alcohol and Alcoholism*, Vol. 50 No. 5, pp. 608-616.

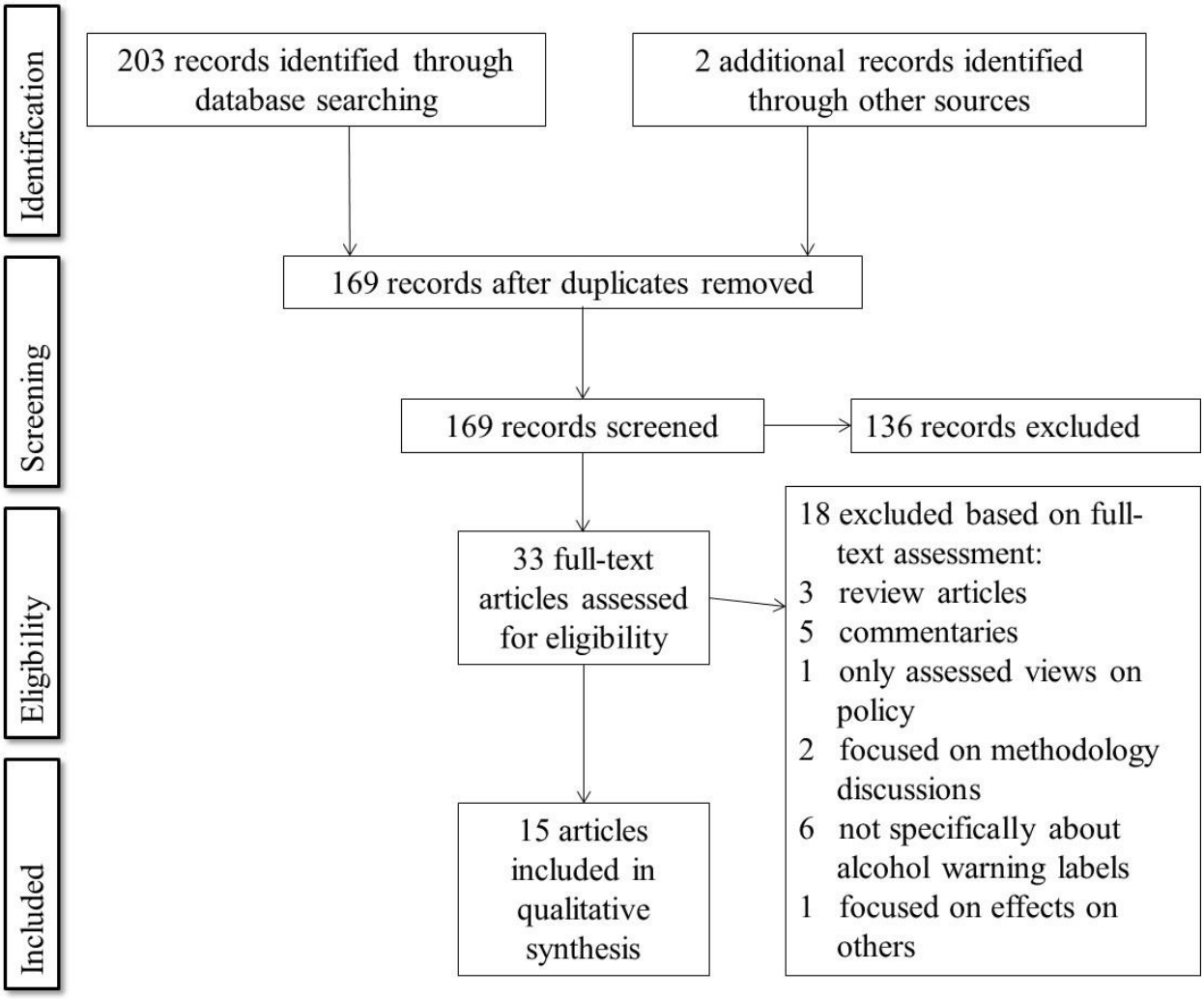

Figure 1 Search flow diagram.

Table 1 Summary of included articles

Author(s) [MMAT score]	Study population and Country	Main aim of study	Study design	Key findings and recommendations
Blume and Resor (2007) [*]	n = 99, Mexican American women colonia residents, USA	Examine association between alcohol use and individual factors.	Survey. Participants (aged 15 to 67) were first assessed regarding English language competence, alcohol consumption level and knowledge about health risks during pregnancy, followed by questions about what they remembered from the labels on alcoholic beverage bottles.	English competence associated with awareness of warning label content. Increased awareness of nutritional information, but not health warnings related to pregnancy, was found to significantly predict lesser amounts of drinking. Beliefs that drinking during pregnancy is not harmful, and perhaps could be helpful, may promote risky drinking during pregnancy. Increased certainty that chronic heavy drinking can cause cognitive and liver problems predicted lower alcohol consumption.
Branco and Kaskutas (2001) [**]	n = 11 pregnant and recent postpartum Native American and African American women, USA	To uncover relevant aspects of women's beliefs and opinions about drinking during pregnancy.	Two focus groups responded to broad questions about exposure and reactions to warning labels and other environmental messages.	Three themes emerged. Exposure and believability of warning label messages; perceptions about risks; and barriers to cutting down. Personal knowledge of women who drank but had healthy babies engenders disbelief of warning label messages. Importance of not overstating the risk in overly simplified (though powerful) warnings. Clarity needed regarding the value of cutting down during pregnancy, even later in the pregnancy.
Coomber *et al.* (2015) [***]	n = 561, online panel members aged 18-45, Australia	Assess alcohol consumption patterns, awareness of the 'Get the facts' logo and warning labels, and use of the associated industry-designed website.	Online survey based on panel members recruited through both online (e.g. banner ads, search engines) and offline (e.g. radio and print advertising) channels.	No participants recalled the 'Get the facts' logo, and the recall rate of warning labels was very low (under 2%) with a higher rate (16 %) for the not to drink during pregnancy warning. Older participants were less likely to be aware of any of the warning labels. Frequent binge drinking, consuming alcohol directly from a can or bottle, and those supporting the use of health warning labels were all significantly more likely to be aware of alcohol warning labels.
Creyer *et al.* (2002) [***]	n = 274, students, USA and Australia	Comparison of two warning labels across county with/without an alcohol warning labels policy.	Experiment with two types of warning: standard USA warning and 'alcohol is a drug' warning. Tested across countries and binge drinking status. USA (n = 168) and Australia (n = 106).	Mixed results across type of warning. The "alcohol is a drug" warning was more effective in generating greater risk perceptions than the standard USA warning.
Glock and Krolak-Schwerdt (2013) [*]	n = 40, students, Germany	Comparing the utility of health-warnings against warnings which contradict positive outcome expectancies.	Two-factorial mixed design with warning labels (health-related vs. positive-related) as a between subjects factor and time (before vs. after presentation of warning labels) as a within-subjects factor.	Greater efficacy of social warnings (countering positive alcohol social outcome expectancy) compared to health warnings. Explicit attitudes did not differ across types of warnings, but implicit attitudes differ with more negative attitudes for the social warnings group.
Jarvis and Pettigrew (2013) [**]	2 focus groups, students; n = 300 online panel (experiment), Australia	Novel method to assess potential effectiveness of four warning labels on bottles of pre-mixed drinks.	Mixed-methods: two focus groups (one male and one female), followed by choice experiment (4^3 design) based on 300 web panel members utilizing messages related to drink driving and health consequences of alcohol consumption.	Negatively framed messages were superior for those with higher alcohol consumption in potentially reducing consumption. Positive messages were found to have potential to increase consumption. Negative health messages had the highest utility for those who reported higher consumption of alcohol. Further, messages with the greatest utility differed across gender.
Jones and Gregory (2010)	n = 44, students, Australia	Examine attitudes and likely drinking behaviour as a result	Six focus groups (two each male/female and two mixed gender) discussed four text-	Limited evidence that any of the four text-based warnings shown would have an impact on alcohol consumption

[***]		from viewing alcohol warning labels	based warnings used in other countries currently as well as ideas for improvement of the warnings.	behaviour amongst participants. However, the research reported a consistent theme where participants in the focus groups agreed that graphic warnings were more effective than text.
Kaskutas (2000) [****]	n = 321 pregnant women, USA	To assess exposure and reactions to health warnings intended to encourage abstinence during pregnancy across ethnic groups.	Clinical samples solicited at prenatal clinics, augmented by outreach to women not necessarily seeking prenatal care. A total of 321 surveys completed with 102 American Indians, 185 African Americans, and 34 whites.	Over 85% of high risk urban minority women found the USA warning to be both believable and understandable, though low in persuasiveness. Women who drank one or more standard drinks per day during pregnancy were less likely than other pregnant drinkers to report seeing the warning label. No significant differences found in exposure rates by ethnicity.

Table 1 Continued

Author(s)	Study population and Country	Main aim of study	Study design	Key findings and recommendations
Kozup *et al.* (2001) [**]	n = 150 consumer panel, USA	Examine the effects of providing a health claim on wine across drinking status and presence of a warning label on attitudes and perceptions.	Mail survey based between-subjects experimental design manipulating health effect (claim, no claim), warning label (present, absent), and wine drinking status (drinker/nondrinker).	A boomerang effect is found amongst drinkers such that participants exposed to a warning without a health claim had more positive attitudes towards the wine product and inferred benefits from drinking moderate amounts of wine than when the warning was not present.
Krischler and Glock (2015) [*]	n = 122; mainly students, Luxembourg and Germany	Assess the utility of alcohol warning labels tailored toward young adults' positive outcome expectancies.	Experimental 3×2 mixed design with warning labels as a between-subjects factor (questions vs. statements vs. control) and expectancy category (positive vs. negative) as a within-subjects factor. With 60 students from Luxembourg, 50 students from Germany and 12 other young adults.	Warning labels presented as questions were more effective than warning labels presented as statements in inducing negative expectations from consuming alcohol. But no changes in drinking intentions were observed due to the labels.
MacKinnon *et al.* (2000) [***]	n = 16661 10th grade and n = 16856 12th grade students, USA	Examined the effects of the USA (1989) alcohol warning label on adolescents during the first 5 years that the warning was required.	Cross-sectional survey conducted in one Indiana county during each school year 1989-1990 through 1994-1995.	There were increases in warning awareness, exposure, and recognition memory. These effects levelled off approximately 3.5 years after the inclusion of the warning on alcohol beverage containers. There was no beneficial change attributable to the warning in beliefs, alcohol consumption, or driving after drinking.
MacKinnon *et al.* (2001) [**]	n = 649, adolescents, USA	Examine the longitudinal effects of warning label exposure on drinking behaviour among adolescents.	Survey conducted in three waves with the same cohort of high school students. Sample (n = 649) completed all three waves of survey measures.	Seeing the alcohol warning at a younger age does not significantly reduce alcohol consumption.
Parackal *et al.* (2010) [***]	n = 1129, non-pregnant female (16-40), New Zealand	Establishing whether a warning on alcoholic beverages would be seen as a source of information on risks associated with consuming alcohol when pregnant.	Nation-wide cross-sectional stratified random sample telephone survey.	Over half of respondents thought that a warning label on the risks associated with drinking during pregnancy would be a preferable source of information.
Pettigrew *et al.* (2014) [***]	n = 48 (focus group), n = 2168 (survey), Australia	Develop and test efficacy of different warning statements that may apply to Australian drinkers.	Mixed-methods: six focus groups and one cross-sectional survey evaluating 12 different warning statements where respondents evaluated 3 of the 12 warning statements. Statements varied in type with frame (positive/negative, whether numerical information was provided, type of appeal (fear/not fear) and whether the message was specific or general in its reference to cancer.	The general warning statement was assessed as most believable, convincing and relevant. The specific 'alcohol increases your risk of bowel cancer' warning performed best of all cancer warnings tested. Overall responses to the cancer messages were neutral to favourable with greater preference for positively framed messages. The use of qualifiers ("can cause") were viewed as less believable and less convincing.
Thomson *et al.* (2012) [***]	n = 45 (focus group), n = 1532 (survey), Australia	Assessing perceptions of specific warning label features and support for alcohol warning label policy.	Mixed-methods: six focus groups and cross-sectional population-based telephone survey.	Strong support (89%) for the use of "Health Warning" labels rather than "Warning" or "Government Health Warning" on alcoholic drinks. Support for matching messages to subpopulations (gender) and specific types of drink. Humorous warnings were found to be less effective than serious ones.

Study population covers general population unless otherwise stated. Stars represent the Mixed Methods Appraisal Tool (MMAT) quality score as described in Pace et al. (2012).

Table 2 Suggested research gaps and avenues for future research

Thematic area	Research gaps and future research opportunities
Use of imagery	• Need for experimental studies using between-subjects designs that compare the efficacy of messages with and without pictorials or pictograms • Need for experimental studies using between-subjects designs that assess the impact of (for example): the inclusion of a border; placement on bottle; and length of message • Needed to assess different levels of "graphicness" similar to that undertaken in tobacco control • Need for large scale longitudinal studies targeting different consumer segments to assess the long-term effects of graphic warning labels
Use of recommendations	• Need for experimental studies using between-subjects designs that compare messages with and without recommendations • Need for more research regarding the influence of providing a recommendation across different message themes (e.g., health; drink driving) • Need for research to examine if and how recommendations might interact with proposed government warning labels
Message specificity	• Need for research in assessing the effect of specific and targeted messages on behavioural compliance beyond potential mediators such as believability and willingness to process the message
Use of signal word(s) or qualifier(s)	• Need for experimental studies using between-subjects designs that assess the impact of signal words or qualifiers controlling for all other message and design factors • Need for research in clarifying the effective use of signal words and qualifiers across different target populations
Use of qualitative versus quantitative information	• Need for experimental studies using between-subjects designs that compare qualitative message statements versus quantitative message statements for the same theme (e.g. liver disease, drink driving)
Non-theme specific gaps identified	• Need for experimental studies using between designs that assess the effectiveness of warning messages with different themes (e.g. social consequences versus health, positively versus negatively framed messages). • Need to consider subgroup effects across psychological variables such as need-for-cognition and demographic factors • Need for research to consider a more stringent assessment of compliance based on repeated measures experimental (within-subject) designs that capture pre- and post- drinking behaviour or views on the harms of alcohol • Need for multi-level studies that compare the effectiveness of policies in operation across a wide range of countries focusing on Argo and Main's (2004) dimensions • Need for large scale longitudinal (panel) research on consumers in countries with mandatory or voluntary alcohol warning label policies (particularly outside the USA) focusing on Argo and Main's (2004) dimensions • Need for large scale surveys that assess judgement of warning label and potential behavioural compliance with warning label across segmentation variables such as gender, age, life stage, consumption level, socioeconomic status • Need for qualitative studies that would provide guidelines on the development of effective warning label messages or designs • Need for qualitative as well as experimental studies to examine the possible cross-over effects of industry alcohol advertising messages (particular on-product messages) on the effectiveness of (on-product) alcohol warning labels • Need for studies conducted beyond developed countries • Need for intervention studies that employ social marketing benchmark criteria • Need for studies that focus on up-stream as well as mid-stream social marketing