Personalised digital interventions showed no impact on risky drinking in young adults: a pilot randomised controlled trial

Title page

Emma L. Davies[1], Adam J. Lonsdale[1], Sarah E. Hennelly [1], Adam R. Winstock [2] ,David R. Foxcroft [1]

[1]Department of Psychology, Social Work and Public Health, Oxford Brookes University, Oxford, OX3 0BP, United Kingdom

[2]Global Drug Survey; University College London, United Kingdom

Corresponding author: Emma L Davies 01865 484056 edavies@brookes.ac.uk

Key words: alcohol, digital interventions, personalised feedback, social embarrassment

Funding: This project was funded by the Oxford Brookes University Central Research Fund.

Conflicts of interest: AW developed the Drinks Meter and OneTooMany applications.

Statement of contribution:

ED designed the study, analysed the data and drafted the manuscript. AL contributed to the design of the study measures and early version of the manuscript. SH contributed to the design of the study, recruited the participants and searching the literature. DF analysed the data and advised on study design. AW designed the interventions and contributed to the study design and measures. All authors contributed to and agreed on the final version of the manuscript

Word count: 3267 excluding tables and references

Personalised digital interventions showed no impact on risky drinking in young adults: a pilot randomised controlled trial

Abstract

Aim: To assess the effectiveness of two personalised digital interventions (OneTooMany and Drinks Meter) compared to controls.

Method: Randomised controlled trial (AEARCTR-0001082). Volunteers for the study, aged 18-30, were randomly allocated to one of two interventions or one of two control groups and were followed up four weeks later. Primary outcomes were AUDIT-C, drinking harms and preloading. Drinks Meter provided participants with brief screening and advice for alcohol in addition to normative feedback, information on calories consumed and money spent. OneTooMany presented a series of socially embarrassing scenarios that may occur when drinking, and participants were scored according to if/ how recently they had been experienced.

Results: The study failed to recruit and obtain sufficient follow-up data to reach a prior estimated power for detecting a difference between groups and there was no indication in the analysable sample of 402 subjects of a difference on the primary outcome measures (Drinks Meter; AUDIT-C IRR=0.98 (0.89-1.09); Pre-loading IRR=1.01 (0.95-1.07); Harms IRR=0.97 (0.79-1.20); OneTooMany; AUDIT-C IRR=0.96 (0.86-1.07); Pre-loading IRR=0.99 (0.93-1.06); Harms IRR=1.16 (0.94-1.43).

Conclusion: Further research is needed on the efficacy of such instruments and their ingredients. However, recruitment and follow-up are a challenge.

Personalised digital interventions showed no impact on risky drinking in young adults: randomised controlled trial

Introduction

In the United Kingdom (UK) young adults, and particularly university students, tend to drink at hazardous levels (Craigs et al., 2012; Davoren et al., 2016), putting themselves at risk from both short and longer term harms. Evidence for the effectiveness of interventions for this population is mixed. Traditional health campaigns intended to provide education about the harmful long term consequences of drinking (e.g. liver disease) have been shown to be largely ineffective (Logan et al., 2015). In addition, recent studies have shown that unit based advice about drinking is also not perceived as relevant to this group (De Visser and Birch, 2012) highlighting the need for novel approaches to reduce harms.

Brief alcohol interventions are often employed in primary care and other settings, and can help individuals to reduce their drinking (Kaner et al., 2007). A systematic review and meta-analysis by Tanner-Smith and Lipsey (2015)concluded that brief alcohol interventions can be effective in reducing both alcohol consumption and related harms in young adults aged 18-30. Digital interventions offer further advantages over such face to face interventions because of their potential to engage young people outside clinical settings and to reach large numbers of people relatively cheaply (Kaner et al., 2015). Given these advantages, it is understandable why researchers, public health professionals and software developers all across the world have developed a number of digital healthcare interventions or medical 'apps'. Thus, it is timely to explore whether the reported benefits of brief interventions can be translated into apps in young adults aged 18-30.

Several apps relating to alcohol reduction are available in the UK though none has had an evaluation of efficacy, to our knowledge. , We present here an attempt to evaluate two UK apps: Drinks Meter and OneTooMany, which differ in their underlying behaviour change techniques (BCTs). Drinks Meter uses a screening tool and gives users personalised feedback about their drinking as well as normative feedback on how their drinking compares with others who have used the tool. Using a taxonomy of BCTs to reduce excessive alcohol consumption (Michie et al., 2012), the main components can be described as 'providing information on consequences of excessive alcohol consumption and reducing excessive alcohol consumption' (BCT 1), and 'provide normative information about others' behaviour and experiences' (BCT 4). This approach is designed to correct the normative misperceptions about other people's drinking which are frequently seen in the general drinking population (Garnett et al., 2015). Despite the fact that evidence for norms-based interventions is weak (Foxcroft et al., 2015a), some evidence suggests that offering drinkers the chance to see how they compare to others could be a useful engagement strategy. For example over 75% of drinkers who responded to the Global Drug Survey 2014 wanted to see how they compared to other people (A Winstock, personal communication 3 February 2017).

The components of OneTooMany can be described using the Michie taxonomy as targeting the consequences of behaviour, specifically through 'providing information about social and environmental consequences' (BCT 5.3) and 'inducing anticipated regret' (BCT 5.5). Studies have found that inducing 'anticipated regret' (e.g. encouraging the contemplation of a hangover), may have a significant impact on student binge drinking (Cooke et al., 2007). Losing control, becoming loud and rude, or getting into trouble, have been identified as an important in young adults' reflections on their drinking (Epler et al., 2009). Thus, highlighting socially embarrassing consequences of excessive consumption might bring about a change in drinking behaviours by inducing anticipated regret about loss of control or damage to their personal image. The Global Drug Survey (GDS) 2015 specifically explored motivations for drinking less among different groups and identified social embarrassment as significant issues for some (Davies et al., 2017). Based on those data and results of focus groups with university students an ultra-brief self-assessment tool was developed focusing on short term negative consequence of drinking, called 'OneTooMany'. It was designed to engage young people in a process that would allow them to reflect on their drinking in a novel and entertaining way and which could signpost interested individuals to participate in the more comprehensive and traditional app, the Drinks Meter.

Thus, we set out to compare the effectiveness of the above two personalised digital interventions (OneTooMany and Drinks Meter) compared to 'control' interventions.

Method

We employed a randomised design of two interventions and two control groups. Follow up measures were sought four weeks after baseline. The protocol was registered on the American Economic Association's registry for randomised controlled trials under identification number AEARCTR-0001082 (Davies et al., 2016a). Posters were displayed at one academic institution, emails were sent to students at three academic institutions, and social media posts advertised the study as open to any young person aged 18-30 in the United Kingdom in order to recruit young people aged 18-30 who self-identified as a current drinker (does not answer 'never' to AUDIT question 1: How often do you have a drink containing alcohol?). The planned final sample size was 800, with 200 per condition based on a sample size calculation.

Interventions: Arm 1) Social embarrassment ("OneTooMany")

On the OneTooMany website participants were asked 20 questions about embarrassing situations that they may have experienced while drinking, for example posting an embarrassing photo on social media, being sick in public, or getting into fights with friends (see above, and http://onetoomany.co/). The website asks participants to indicate whether each of the 20 experiences have occurred 'in the last month', 'in the last year' or 'never/not in the last year'. On completion, the participant is presented with a social embarrassment score (0-40). These scores are broken down into 4 groups, each category being given a label (e.g. Culus Major). Feedback is offered on the the types of drinker that he/she falls into and the associated risks and consequences (to act as motivators to reduce consumption).

Arm 2) Personalised feedback about drinking ("Drinks Meter")

Drinks Meter is a free, anonymous, smart phone and online digital app that offers a traditional 'identification and brief advice' approach in which users receive personalised feedback about their own drinking and are informed about how this compares with other people with similar demographic / geographic characteristics (see http://www.drinksmeter.com/). They also receive information about the amount of calories they consume when drinking and the amount of money they spend on alcohol, in comparison to others. It was designed as a tool to map data for public health organisations and health service providers and it also allows signposting into treatment. The Drinks Meter app has been downloaded by over 30,000 people and has transparent intervention content (offering assessment, identification of drinkers at high risk and advice on how to reduce). In a recent analysis of app store reviews the Drinks Meter was the most highly praised app overall (Milward et al., 2016).

Arms 3 and 4) Control groups

In one arm of the trial, participants were directed to a placebo condition, where they were asked to imagine they are exposed to (reading / watching / listening) information about alcohol misuse, without actually receiving any alcohol information. This condition was intended to act as a more robust placebo condition, as the active ingredients of the intervention will have been removed (e.g. feedback on drinking), but other aspects e.g. being directed to a website and being given some information about alcohol, remained (albeit in an imaginary sense). The final arm of the trial was a no-treatment control group who were asked for baseline and follow up measures only.

Outcomes: Primary outcome measures:

Follow up outcome measures were sought four weeks following the baseline questionnaire and randomisation. Alcohol consumption data were requested using AUDIT-C (Babor et al., 2001) which has been empirically assessed as the best measure for hazardous drinking in young people (Foxcroft et al., 2015b). Also, a scale to rate drinking-related harms scale was adapted from (Davies et al., 2016b) based on a validated scale used in a national survey conducted in the UK (Fuller, 2013). This scale was piloted with a young adult sample and minor changes were made to ensure the drinking harms were relevant to this population, which meant the addition of items about drug use and missing university/work. The items were; being sick; embarrassed; missing university/ work; trouble with police; injury; being taken to hospital; having a fight; taking an illegal drug or 'legal high'; losing a personal item such as a phone; unprotected sex; regretted sex; not knowing where you are when you woke up; having an embarrassing photo posted on social media.

In addition we used as an outcome measure 'pre-loading', (consuming alcohol at home before a night out) which is associated with high subsequent levels of consumption, intoxication and harms (Foster and Ferguson, 2012): '*How often in the last four weeks have you drank alcohol at home before going out on a night out?* (1 = never - 5 = always).

Randomisation: Random allocation into one of the 4 groups was via Qualtrics Survey Software. Researchers were blind to the allocation.

Incentivising and Ethics: After randomisation, participants were told that if they completed the baseline and follow up questionnaire measures they would be entered into a draw to win one of four iPad mini tablets, or a £50 or a £15 Amazon voucher. The study received approval from Oxford Brookes University Ethics Committee (reference 150944).

Statistical methods: Quantitative data were analysed in the form of descriptive statistics (e.g. item-specific means and standard deviations), and generalized linear models using the R statistical language (R Core Team, 2016) were used to estimate incident rate ratios and 95%

confidence intervals associated with intervention effects. As outcomes were all count data, we used a Poisson regression with a log link function. Baseline measures of age, gender and behaviour corresponding to the outcome variable being analysed were included in the model, but are not presented here for clarity.

Results

In total, 790 people expressed an interest in taking part and were sent an email inviting them to complete the baseline questionnaire. Six hundred and forty one people gave their consent to take part and started to complete the baseline measures. Of those, 11 participants completed the baseline measures twice, and we excluded their second attempt. Six non-drinkers started the questionnaire, and once indicating that they never drank alcohol (AUDIT question one) they were directed to a landing page thanking them for their interest and advising them that their participation was complete. Fifty five participants reported their age as 31 or over and so were excluded from the current analysis. A further 81 participants did not complete baseline measures. A total of 488 participants were randomised to one of the arms of the trial and were therefore included in the intention to treat analysis (*M* age = 21.70, *SD* = 3.28; 67.2% female; 83.4 % students; 75% from one academic institution). Four weeks later follow up data were collected from 402 participants, a response rate of 82.3% for the complete case analysis (*M* age = 21.76; SD = 3.40; 69.2 % female; 82.8 % students). Figure 1 shows the flow of participants through the intervention protocol.

[Insert Figure 1]

Randomisation check: There were no substantive differences (e.g. age, gender & alcohol intake) between participants in each of the four arms of the trial at baseline (Table 1).

[Insert Table 1]

Descriptive statistics related to participant demographics for the whole sample, and broken down for each condition are presented in Table 1. While the sample was predominantly from the student population, young people in employment or looking for work also took part.

Primary outcome measures

In the sample as a whole, significant reduction were found in the scores on all of the primary outcomes measures between baseline and follow up. Average AUDIT-C score reduced from 6.46 (*SD* =2.48) to 5.94 (*SD* = 2.40) ($t(384) = 7.29$, p<.001), median harms decreased from 2.5 (*IQR* = 4) to 1 (*IQR* =2) ($t(401) = 16.17$, $p<.001$) and mean preloading frequency decreased from 3.77 (SD =1.22) to 3.14 (*SD* = 1.62) ($t\ (378) = 9.84$, $p<.001$).

For all outcomes (AUDIT-C, harms, and pre-loading) the point estimates and confidence intervals were very similar in both control questions and thus they were collapsed to present the results with clarity. In all analyses, gender was not a significant predictor, but lower age and higher baseline behaviour were significant predictors in the generalized linear models. We only report incident risk ratios (IRR) and confidence intervals associated with intervention effects (Table 2). There were no differences between the interventions and controls on any of the primary outcome measures.

Post-hoc analyses

Our recruitment strategy may have failed to attract young people who were at risk of harm from their drinking and thus our null results may be attributable to the number of low risk drinkers in the sample, who were not the intended targets of the interventions. Thus, in a post-hoc analysis, we decided to compare risky and non-risky drinkers to determine if there were any differences in the intervention effects. Because there has previously been a lack of evidence for optimal cut offs we used two approaches and compared respondents who were classified according to their baseline AUDIT score as non-risky (<8) and risky (8+) drinkers (Babor et al., 2001), and using cut offs of 4+ female and 9+ for males (Foxcroft et al., 2015b).

There were no significant differences between risky and non-risky drinkers who received either intervention on AUDIT-C or pre-loading, compared to controls, at follow-up. However in subgroup 1 non-risky drinkers who were in the OneTooMany condition were 2.58 times more likely to report increased harms at baseline compared to non-risky drinkers in the control conditions (95% CI 1.29 to 5.18). Because this was unexpected, we recoded the harms in order to explore this further. Harms were re-coded by the authors, and the classifications discussed and agreed with a class of students. Being in trouble with the police, suffering an injury, being taken to hospital, taken an illegal drug, taken a 'legal high', and having unprotected sex were classified as 'longer term or health related' harms. Being sick; being embarrassed; missing university or work; having a fight; losing something; having sex and then regretting it; not knowing where they were when they woke up; or had an embarrassing photo posted on social media were classified as 'social or short term' harms. The subsequent analysis showed that the subgroup 1 non-risky drinkers sub-sample in the OneTooMany condition reported more social or short term harms at follow up than controls (IRR= 2.52, 95% CI 1.61 to 5.51).

In subgroup 2 those who were classified as risky drinkers at baseline, reported more social or short term harms at follow up than controls (IRR= 1.29, 95% CI 1.01 to 1.61). It is likely that this difference in subgroups can be accounted for by gender differences. In sub group 1 the same cut off is applied to male and female drinkers, whereas in subgroup 2 there is a different cut off for risky drinking for males (9+)and females (4+) and so more female drinkers were classed as risky. There was no difference between any of the groups in terms of the longer term or health related harms.

[Insert Table 2]

Discussion

The aim of the current study was to assess the effectiveness of two personalised digital interventions (OneTooMany and Drinks Meter) compared to controls. We found no effects of the interventions compared to controls on alcohol consumption measured by AUDIT-C or risky drinking indicators of harm and pre-loading from baseline to four week follow up. Although reductions in these measures were observed across the sample, neither intervention significantly changed behaviour over and above controls. We were underpowered, however, to show such a difference.

There are large numbers of alcohol related apps on the market, but only a minority have a health focus, and of these few contain recognised BCTs (Crane et al., 2015). The current study explored two apps which use specific BCTs that aim to reduce risky drinking, but found no benefits over controls. One explanation is that completing measures of drinking and behaviours may lead to changes in behaviour themselves; this sometimes called the 'mere measurement' or 'question behaviour' effect (Sprott et al., 2006). If this is a consistent effect, it means that rather than seeking the optimum intervention format, instead we might consider searching for an optimum combination of questionnaire measures which could be a much more cost effective way of changing behaviour. However, while some systematic reviews of the mere-measurement phenomena suggest that this might be a powerful public health tool, others suggest the effects are spurious (Rodrigues et al., 2016). While we cannot conclude that this was the case in our study, it is certainly an interesting phenomenon, which should be explored further to determine the need for 'active' intervention strategies to reduce alcohol consumption among risky drinkers.

Conversely it does appear that apps do engage younger drinkers, and those at higher levels of risk (Garnett et al., 2017). A recent review highlighted the importance of tailoring apps to young people's interests, and ensuring they are engaging and easy to use (Milward et al., 2016). Future research should focus on the level of additional engagement such apps can bring to the delivery of brief advice to groups of young drinkers who may otherwise be unreached by traditional methods.

However, given absence of significant benefits above a 'control' intervention, perhaps it is time to question the apparent shift away from face-to-face to digital interventions. This may be based on naïve and untested assumptions of 'technological utopianism' or economic cost-effectiveness, rather than a rigorous and systematic evaluation of the available research evidence. There have been calls to ensure more rigorous testing of all freely available medical apps, as well as to educate app consumers to make choices about the health apps that they use (Krieger, 2016; Wicks and Chiauzzi, 2015). Recently, the UK National Health Service has withdrawn its 'app library' after reviews discovered concerns with data protection and privacy (Huckvale et al., 2015),

Post-hoc sub group analysis found increased reporting of harms in the OneTooMany condition when exploring differences between risky and non-risky drinkers. It may be that OneTooMany raises people's awareness of negative consequences and makes them more likely to recall subsequent events at follow up. However, these increased harms were accompanied by a decrease in alcohol consumption, so it is possible that there is an alternative explanation for this finding. It is likely that exposure to the intervention either normalised or heightened awareness of these consequences, and made it more socially acceptable to report them at follow-up.

As well as the loss of power already mentioned it is important to state that we recruited a number of low risk drinkers into the study, who were not the intended target group for the interventions, whereas drinking apps do tend to attract higher risk drinkers in the real world (Garnett et al., 2017). It is possible that our participants did actually reduce their alcohol consumption, but for reasons other than taking part in our study. A further limitation is that our sample were predominantly students, thus the generalisability of the findings to non-student populations is unclear. Linked to this, our time two data were collected during a later part of the university semester (late Autumn/ early Winter) when they may have had more work to complete for their courses, this increased workload may interfered with students' drinking behaviour.

In conclusion we found no effect of two digital interventions on risky drinking in young people compared to controls. These findings highlight the need for a comprehensive understanding of what constitutes an effective digital healthcare intervention to reduce alcohol consumption within this population.

References

Babor, TF, Higgins-Biddle, JC, Saunders, JB, Monteiro, MG (2001) The Alcohol Use Disorders Identification Test, Guidelines for Use in Primary Care. Geneva: World Health Organization.

Cooke, R, Sniehotta, F, Schuz, B (2007) Predicting binge-drinking behaviour using an extended TPB: Examining the impact of anticipated regret and descriptive norms. *Alcohol and Alcoholism* **42**: 84-91.

Craigs, CL, Bewick, BM, Gill, J, O'May, F, Radley, D (2012) UK student alcohol consumption: A cluster analysis of drinking behaviour typologies. *Health Education Journal* **71**: 516-26.

Crane, D, Garnett, C, Brown, J, West, R, Michie, S (2015) Behavior Change Techniques in Popular Alcohol Reduction Apps: Content Analysis. *Journal of Medical Internet Research* **17**.

Davies, EL, Conroy, D, Winstock, AR, Ferris, J (2017) Motivations for reducing alcohol consumption: An international survey exploring experiences that may lead to a change in drinking habits. *Addictive Behaviors*.

Davies, EL, Lonsdale, A, Foxcroft, DR (2016a) A randomised controlled trial to investigate the effects of two personalised digital interventions on alcohol risky drinking and alcohol related harms in young people. *AEA RCT Registry* **March 10**.

Davies, EL, Martin, J, Foxcroft, DR (2016b) Age differences in alcohol prototype perceptions and willingness to drink in UK adolescents. *Psychology, Health & Medicine* **21**: 317-29.

Davoren, MP, Demant, J, Shiely, F, Perry, IJ (2016) Alcohol consumption among university students in Ireland and the United Kingdom from 2002 to 2014: a systematic review. *BMC Public Health* **16**: 1-13.

De Visser, RO and Birch, JD (2012) My cup runneth over: Young people's lack of knowledge of low-risk drinking guidelines. *Drug and Alcohol Review* **31**: 206-12.

Epler, AJ, Sher, KJ, Piasecki, TM (2009) Reasons for Abstaining or Limiting Drinking: A Developmental Perspective. *Psychology of Addictive Behaviors* **23**: 428-42.

Foster, JH and Ferguson, CS (2012) Home Drinking in the UK: Trends and Causes. *Alcohol and Alcoholism* **47**: 355-58.

Foxcroft, DR, Moreira, MT, Almeida Santimano, NML, Smith, LA (2015a) Social norms information for alcohol misuse in university and college students. *Cochrane Database of Systematic Reviews*.

Foxcroft, DR, Smith, LA, Thomas, H, Howcutt, S (2015b) Accuracy of Alcohol Use Disorders Identification Test for Detecting Problem Drinking in 18–35 Year-Olds in England: Method Comparison Study. *Alcohol and Alcoholism* **50**: 244-50.

Fuller, E (2013) Smoking, Drinking and Drug Use in Young People in England in 2012. London: Health and Social Care Information Centre.

Garnett, C, Crane, D, West, R, Michie, S, Brown, J, Winstock, A (2015) Normative misperceptions about alcohol use in the general population of drinkers: A cross-sectional survey. *Addictive behaviors* **42**: 203-6.

Garnett, C, Crane, D, West, R, Michie, S, Brown, J, Winstock, A (2017) User characteristics of a smartphone app to reduce alcohol consumption. *Translational Behavioral Medicine*.

Hackley, C, Bengry-Howell, A, Griffin, C, Szmigin, I, Mistral, W, Hackley, RA (2015) Transgressive drinking practices and the subversion of proscriptive alcohol policy messages. *Journal of Business Research* **68**: 2125-31.

Huckvale, K, Prieto, JT, Tilney, M, Benghozi, P-J, Car, J (2015) Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Medicine* **13**: 1-13.

Kaner, EFS, Beyer, FR, Brown, J *et al.* (2015) Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. *Cochrane Database of Systematic Reviews*.

Kaner, EFS, Dickinson, HO, Beyer, FR *et al.* (2007) Effectiveness of brief alcohol interventions in primary care populations. *Cochrane Database of Systematic Reviews*.

Krieger, WH (2016) When are medical apps medical? Off-label use and the Food and Drug Administration. *Digital Health* **2**.
Logan, DE, Kilmer, JR, King, KM, Larimer, ME (2015) Alcohol interventions for mandated students: behavioral outcomes from a randomized controlled pilot study. *Journal of studies on alcohol and drugs* **76**: 31-7.
Michie, S, Whittington, C, Hamoudi, Z, Zarnani, F, Tober, G, West, R (2012) Identification of behaviour change techniques to reduce excessive alcohol consumption. *Addiction* **107**: 1431-40.
Milward, J, Khadjesari, Z, Fincham-Campbell, S, Deluca, P, Watson, R, Drummond, C (2016) User Preferences for Content, Features, and Style for an App to Reduce Harmful Drinking in Young Adults: Analysis of User Feedback in App Stores and Focus Group Interviews. *JMIR mHealth uHealth* **4**: e47.
Rodrigues, AM, French, DP, Sniehotta, FF (2016) Commentary: The Impact of Asking Intention or Self-Prediction Questions on Subsequent Behavior: A Meta-Analysis. *Frontiers in Psychology* **7**.
Sprott, DE, Spangenberg, ER, Block, LG, Fitzsimons, GJ, Morwitz, VG, Williams, P (2006) The question–behavior effect: What we know and where we go from here. *Social Influence* **1**: 128-37.
Tanner-Smith, EE and Lipsey, MW (2015) Brief Alcohol Interventions for Adolescents and Young Adults: A Systematic Review and Meta-Analysis. *Journal of Substance Abuse Treatment* **51**: 1-18.
Wicks, P and Chiauzzi, E (2015) 'Trust but verify' – five approaches to ensure safe medical apps. *BMC Medicine* **13**: 1-5.

Tables and Figures

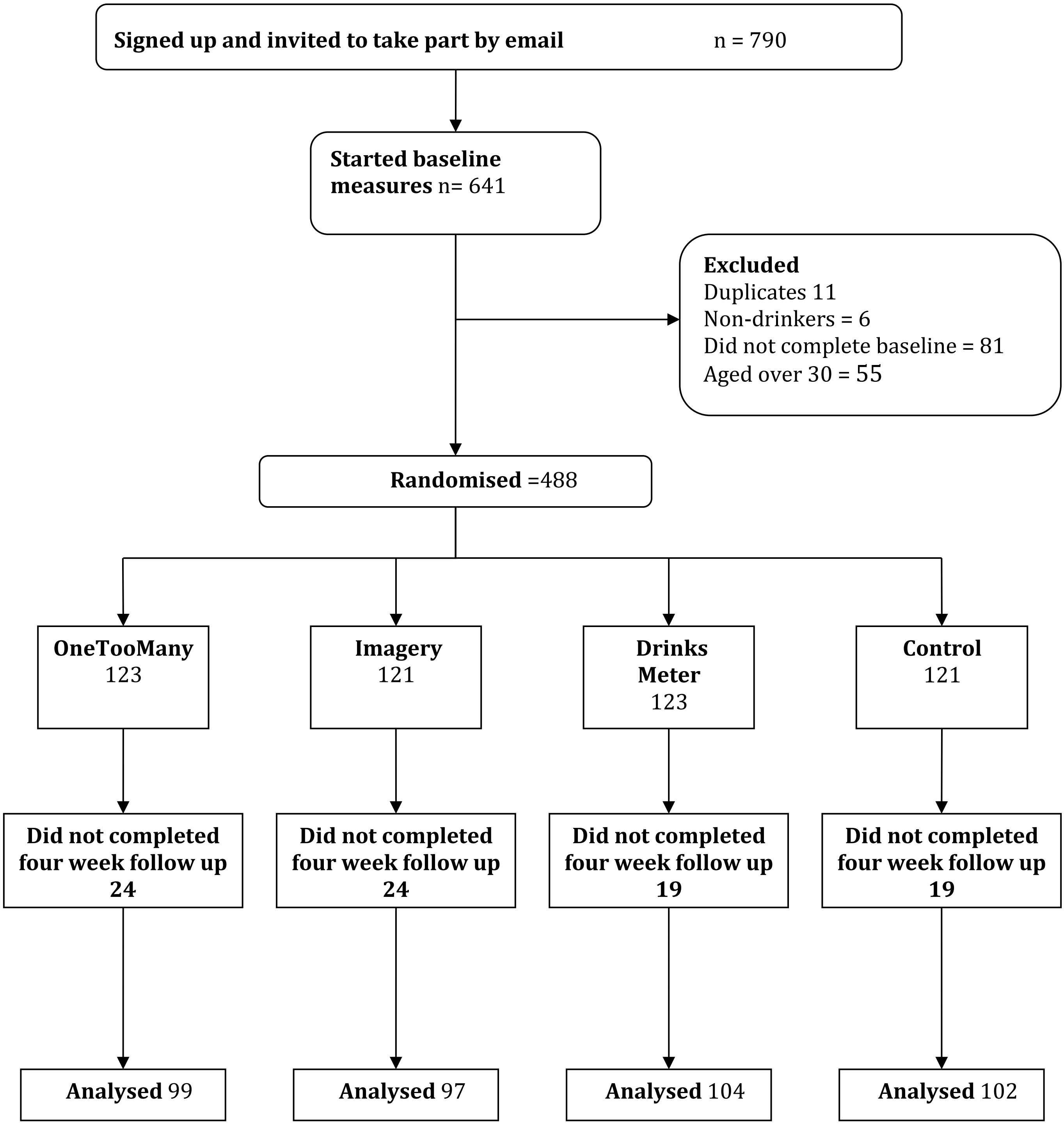

Figure 1 Diagram to show flow of participants through the study protocol

Table 1 *Baseline characteristics of participants in the whole sample and for each intervention condition*

Characteristic	Full sample (N= 488)	OneTooMany (N =123)	Drinks meter (N=123	Controls (N=242)	p-value
Female, N (%)	328 (67.2)	84 (68.3)	84 (68.3)	160 (66.1)	p=.372
White	446 (91.4)	111(90.2)	113 (91.9)	222 (90.2)	p=.331
Age mean (SD)	21.70 (3.28)	21.94 (3.38)	21.76 (3.37)	21.54 (3.18)	p=.517
Student N (%)	407 (83.4)	103 (83.7)	105 (85.4)	205 (84.7)	p=.870
AUDIT C mean (SD)	6.52 (2.53)	6.20 (2.54)	6.55 (2.62)	6.66 (2.47)	p=.259
Pre-loading	3.76 (1.20)	3.69 (1.22)	3.78 (1.20)	3.78 (1.20)	p=.775
Harms median (IQR)	3 (4)	2 (2)	2 (3)	3 (4)	p=.140
Short term/social harm	2 (4)	2 (3)	2 (2)	2 (3)	p=.121
Long term/health harm	0 (1)	0 (1)	0 (1)	0 (1)	p=.372

Chi squared test for association conducted for gender, ethnicity and occupation, ANOVA conducted for age, AUDIT C, and pre-loading, Kruskal-Wallis test conducted for harms

Table 2 *Four week outcomes comparing intervention groups to control groups for the whole sample, and comparing those who were risky and non-risky drinkers at baseline; generalised linear models (poisson or quasi-poisson regression) at four weeks follow-up. Incident risk ratio (IRR) and 95% confidence intervals*

	AUDIT-C	Pre-loading	Total Harms	Long term/ health	Short term social

Full Sample Control = 199					
Drinks Meter N = 101	0.98 (0.89-1.09)	1.01 (0.95-1.07)	0.97 (0.79-1.20)	0.76 (0.47-1.19)	1.09 (0.85-1.38)
OneTooMany N = 96	0.96 (0.86-1.07)	0.99 (0.93-1.06)	1.16 (0.94-1.43)	1.00 (0.65-1.51)	1.30 (1.02-1.64)
Subgroup 1: non-risky drinker Control N = 67					
Drinks Meter N= 28	0.98 (0.77-1.23)	1.01 (0.90-1.14)	1.44 (0.51-3.63)	1.69 (0.08-15.64)	1.30 (0.41-3.55)
OneTooMany N= 35	0.92 (0.73-1.15)	1.01 (0.90-1.13)	2.58 (1.29-5.18)	1.06 (0.10-8.42)	2.52 (1.16-5.51)
Subgroup 1: risky drinkers Control N = 131					
Drinks Meter N= 61	0.99 (0.88-1.11)	1.00 (0.93-1.08)	0.86 (0.68-1.07)	0.65 (0.39-1.04)	0.95 (0.74-1.23)
OneTooMany N = 69	0.98 (0.87-1.10)	0.98 (0.91-1.06)	1.07 (0.86-1.33)	0.96 (0.61-1.47)	1.16 (0.90-1.50)
Subgroup 2: non-risky drinkers Control = 29					
Drinks Meter N = 13	1.04 (0.71-1.50)	1.03 (0.86-1.23)	0.52 (0.06-4.44)	0.00 (0.00-Inf)#	0.67 (0.08-5.54)
OneTooMany N = 17	0.93 (0.65-1.33)	1.01 (0.85-1.20)	0.78 (0.18-3.36)	0.00 (0.00-Inf)#	1.14 (0.28-4.66)
Subgroup 2: risky drinkers Control = 169					
Drinks Meter N = 84	0.98 (0.88-1.10)	1.00 (0.93-1.07)	0.93 (0.75-1.16)	0.72 (0.44-1.15)	1.04 (0.81-1.33)
OneTooMany N = 79	0.97 (0.87-1.09)	0.99 (0.93-1.06)	1.16 (0.94-1.44)	1.01 (0.64-1.54)	1.29 (1.01-1.64)

Complete Case analysis. Subgroup 1: Baseline AUDIT 8+; Subgroup 2: Baseline AUDIT 4+ (females) or 9+ (males) from Foxcroft et al. (2015);
 # not estimated reliably because of sparse data