

# Attitudes and Behaviour towards Alcohol Survey 2013/14 to 2015/16:

Methods report for the combination of three survey datasets

November 2016

ISBN: 978-1-927303-91-7

Prepared for the Health Promotion Agency by:  
Thewaporn (Wa) Thimasarn-Anwar, Susan Cook, and Holly Trowland

Citation: Health Promotion Agency (2016). *Attitudes and Behaviours towards Alcohol Survey 2013/14 to 2015/16: Methods report for the combination of three survey datasets*. Wellington: Health Promotion Agency.

This document is available at: <http://www.hpa.org.nz/research-library/research-publications>

Any queries regarding this report should be directed to the Health Promotion Agency at the following address:

Health Promotion Agency  
PO Box 2142  
Wellington 6140  
New Zealand  
[www.hpa.org.nz](http://www.hpa.org.nz)  
[research@hpa.org.nz](mailto:research@hpa.org.nz)

November 2016

## COPYRIGHT

The copyright owner of this publication is HPA. HPA permits the reproduction of material from this publication without prior notification, provided that fair representation is made of the material and HPA is acknowledged as the source.

## INTRODUCTION

---

Combining two or more datasets allows enough observations, in particular for respondents from various sub-population groups, for more specific analyses and to produce estimates with greater statistical precision. The combination of survey datasets over adjacent survey years is termed 'pooling'.

The 2013/14, 2014/15 and 2015/16 Attitudes and Behaviour towards Alcohol Survey (ABAS) datasets were pooled in order to allow researchers to analyse small sub-population groups within the dataset such as high risk drinkers and those living in small geographic regions.

These three surveys are suitable for pooling because the surveys were conducted using the same methodology and design and the questionnaire remained essentially the same over the three years. Although the surveys were conducted over three adjacent years, we can consider the three samples as representing the same population of New Zealanders ie, all adults aged 15+ years living in New Zealand.

## METHODS

---

The pooled dataset was weighted using benchmarking - it was treated as if all of the participants were drawn from a single survey year and the one large sample formed the population. The dataset was benchmarked to the most recent census data available (the 2013 Census) using gender, age and prioritised ethnicity. The 2013 Census was used for benchmarking because the implicit assumption in pooling survey data is that the underlying population of interest has not changed significantly during the pooled data time frame. If the population had changed drastically then it would be unclear what population the data describes. Because of this assumption, it is reasonable to benchmark the data to the population of the first year of the pooled dataset.

Before pooling the datasets, comprehensive data checking and screening procedures were performed. These included logic/routing checking, checking for unusual codes, and accounting for missing values and outliers.

Issues considered before pooling the datasets were as follows:

*a) Are the survey samples independent?*

The three surveys conducted in adjacent years are essentially statistically independent. In each survey approximately 4,000 households were randomly selected using a Random Digit Dialling (RDD) method. Because all possible telephone numbers were used as the sample frame each year, there is a chance that a participant selected in one survey year could also be selected in another survey year. However, because of the sample size of each survey, the chance is

extremely low. Therefore, proceeding as if the three cycles are simply three independent samples from the New Zealand population in different years is acceptable.

*b) Are survey designs the same?*

The survey methodology has remained unchanged over the period 2013-2016. The same fieldwork provider used the same approach and methods for sample selection and data collection in each of the three surveys.

*c) Are the questionnaires the same? Do the variables mean the same thing?*

The survey questionnaire format, topics, and content for 2013/14, 2014/15, and 2015/16 were the same, with the exception of minor changes for clarity, and minimal additions/deletions. Further, the data has the same format and the metadata (the variable and value labels) were essentially the same for all three survey years.

## CONSIDERATIONS WHEN ANALYSING THE POOLED DATASET

---

Descriptive statistics calculated from the pooled dataset might differ from an estimate that was calculated from separate surveys (such as using only 2014/15 survey data), with potentially different interpretations. Because the sample size and survey design have remained constant, estimates using the pooled dataset are essentially an average over the three survey years. Researchers can check if estimates have changed over time and report a time trend if an average value is not deemed appropriate.

When estimating model parameters, it is good practice to add a cycle term for each wave of the survey into the model, to adjust for any variation with time (in this instance, over the three survey years).

## REFERENCES

---

Cherpitel, C. J., & Ye, Y. (2015). Risky drinking, alcohol use disorders, and health services utilization in the US general population: data from the 2005 and 2010 National Alcohol Surveys. *Alcoholism: Clinical and Experimental Research*, 39(9), 1698-1704.

Knott, C. S., Coombs, N., Stamatakis, E., & Biddulph, J. P. (2015). All cause mortality and the case for age specific alcohol consumption guidelines: pooled analyses of up to 10 population based cohorts. *BMJ*, 350, h384.

Perreault, K., Bauman, A., Johnson, N., Britton, A., Rangul, V., & Stamatakis, E. (2016). Does physical activity moderate the association between alcohol drinking and all-cause, cancer and cardiovascular diseases mortality? A pooled analysis of eight British population cohorts. *British Journal of Sports Medicine*, bjsports-2016.