

This is the peer reviewed version of the following article:

Barratt, M.J., Ferris, A., Zahnow, R., Palamar, J., Maier, L., & Winstock, A. (2017). Moving on From Representativeness: Testing the Utility of the Global Drug Survey. *Substance Abuse*, 11.

which has been published in final form at
<https://doi.org/10.1177/1178221817716391>

This article may be used for non-commercial purposes in accordance with
[SAGE's Archiving and Sharing Policy](#)

© 2017. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Moving on from representativeness: testing the utility of the Global Drug Survey

Monica J. Barratt^{a, b, c}, Jason A. Ferris^d, Renee Zahnow^d, Joseph J. Palamar^{e, f},
Larissa J. Maier^g, Adam R. Winstock^{h, i}

^a Drug Policy Modelling Program, National Drug and Alcohol Research Centre, UNSW Australia, Sydney NSW 2052, Australia.

^b National Drug Research Institute, Faculty of Health Sciences, Curtin University, GPO Box U1987, Perth WA 6845, Australia.

^c Centre of Population Health, Burnet Institute, 85 Commercial Road, Melbourne Vic 3004, Australia.

^d Institute for Social Science Research, University of Queensland, St Lucia Qld 4072, Australia.

^e New York University Langone Medical Center, Department of Population Health, New York, NY, US.

^f Center for Drug Use and HIV Research, New York University College of Nursing, New York, NY, US.

^g University of Zurich, Zurich, Switzerland.

^h University College London, London, UK.

ⁱ Global Drug Survey Ltd, London, UK.

Corresponding author: Dr Monica Barratt, m.barratt@unsw.edu.au

Substance Abuse: Research and Treatment

This article is submitted as part of a special issue:
Emerging Topics and Innovative Methodologies in Cannabis Research.

Abstract

A decline in response rates in traditional household surveys, combined with increased internet coverage and decreased research budgets, has resulted in increased attractiveness of web survey research designs based on purposive and voluntary opt-in sampling strategies. In the study of hidden or stigmatised behaviours, like cannabis use, web survey methods are increasingly common. However, opt-in web surveys are often heavily criticised due to their lack of sampling frame and unknown representativeness. In this paper, we outline the current state of the debate about the relevance of pursuing representativeness, the state of probability sampling methods and the utility of non-probability, web survey methods especially for accessing hidden or minority populations. Our paper has two aims: (1) to present a comprehensive description of the methodology we use at Global Drug Survey (GDS), an annual cross-sectional web survey, and (2) to compare the age and sex distributions of cannabis users who voluntarily completed (a) a household survey or (b) a large web-based purposive survey (GDS), across three countries: Australia, the US and Switzerland. We find that within each set of country comparisons the demographic distributions among recent cannabis users are broadly similar: demonstrating that the age and sex distributions of those who volunteer to be surveyed are not vastly different between these non-probability and probability methods. We conclude that opt-in web surveys of hard-to-reach populations are an efficient way of gaining in-depth understanding of stigmatized behaviours, and are appropriate, as long as they are not used to estimate drug use prevalence of the general population.

Keywords: Web Survey, Sampling, Representativeness, Hard-to-reach, Hidden Population, Drug Use

Manuscript Type: Original Research

Introduction

[J]ust because a survey is based on a probability sample, does not mean it is a valid and reliable reflection of the population it purports to measure. In the same way, just because a survey is based on self-selected methods does not automatically disqualify it from attention or invalidate its findings.

American Association for Public Opinion Research taskforce on non-probability sampling (2013) ^{1, p. 13}

Research surveys that do not adopt probability based sampling frames have generally been considered inherently inferior to surveys that adopt probability sampling frames ¹⁻³. Our research group conducts the world's largest web survey on drug use, the Global Drug Survey (GDS). The GDS data are drawn from a population of self-selected respondents and, to date, despite the fact that the sampling is not probability-based, our papers have been published in a variety of high-impact journals ⁴⁻⁸. Nonetheless, during the submission process, some GDS manuscripts have been rejected due to concerns about the lack of a probability-based sampling frame. Indeed, some journals in our field, anecdotally, do not send papers out to review if they are based on survey data derived from purposive samples.

In this paper, we aim to address the assumptions about sample representativeness that may underlie some of these rejections, thereby better explaining and justifying the methodology we use at GDS. We concur with others ^{2,9-13} in arguing that sample representativeness is typically only necessary when answering research questions about population prevalence estimates. When the research has other aims, including measuring relationships between variables ¹⁴ or in-depth profiling of sub-populations, the use of probability-based sampling frameworks is often inefficient, may be unnecessary or even better avoided ^{2,10}. Furthermore, the 'gold standard' household survey is more like 'tarnished gold' ¹⁵ when response rates and volunteer bias still affect what data are obtained ^{16,17}. In this paper we use predicted probability models to compare age-and-sex-related probabilities of self-reported cannabis use between matched GDS and household survey data, making within-country comparisons for Australia, the US and Switzerland. We aim to demonstrate that the GDS can produce samples of similar age and sex distribution to equivalent household surveys, while offering a level of detail that is usually not possible from general household surveys.

Sampling people who use drugs

In the area of substance use research, it can be time-consuming, incredibly costly and sometimes impossible to access people who report engaging in stigmatized behaviours through representative sampling frames ¹⁸. Probability sampling methods are limited in many regards. Nonresponse bias limits the representativeness of household surveys. Landline telephone use is declining and biased against the younger and more transient populations one wishes to target in much substance use research. Even when recruiting via mobile phones, several biases will occur: for example, geographic inaccuracies in sampling frames are more common when using mobile phone numbers ¹⁹, mobile phone users are less likely to expect or respond to research requests ²⁰ and cannabis and tobacco use are more commonly reported through mobile phone sampling frames compared with landline samples ²¹. People are generally less likely to report illicit drug use on the phone when compared to anonymous web surveys. In addition, probability methods are particularly expensive when researchers are targeting rare practices for in-depth analysis ²². The changing context of survey research ²³ and the growth of internet access has prompted greater use of purposive sampling of otherwise hard-to-reach populations via web surveys and internet recruitment ^{24,25}. Many studies of drug using populations now use internet recruitment and/or survey methods with purposive sampling ^{25,26}. Some well-documented advantages of conducting surveys of hidden populations online include: large and geographically and linguistically diverse samples can be obtained relatively easily; responses can be gathered more rapidly; costs and other resource demands are relatively low; transcription and data-entry is automated; and flexibility and convenience are enhanced for both respondents and researchers ²⁷⁻²⁹. Because opt-in web-surveys are increasingly popular for accessing hidden populations of people who use drugs, it is important that researchers in our field 'take off the blindfold' ²² and work with, rather than against, web samples.

What is representativeness good for?

There have been lively debates in the field of epidemiology regarding the pursuit of sample representativeness, with some concluding that sample representativeness is only required where population

prevalence estimates are sought, and should even be actively avoided where other research questions are primary^{2,11}. But what is representativeness? In a series of papers published in 1979-1980³⁰⁻³³, Kruskal and Mosteller provide an account of the various meanings of the term ‘representative sampling’. They describe 6 categories of meaning in the non-scientific and non-statistical scientific literature, including “(1) general acclaim for data, (2) absence of selective forces, (3) miniature of the population, (4) typical or ideal case(s), coverage of the population, and (6) vague term, to be made precise”^{32, p. 244}. When examining the statistical literature specifically, they found the above meanings plus 3 additional categories, including “representative sampling as a specific sampling method, (8) representative sampling as permitting good estimation, (9) representative sampling as good enough for a particular purpose”^{32, p. 244}. Kruskal and Mosteller’s account of the meanings and histories of the idea of ‘representativeness’ in sampling demonstrates that the seeming blind faith in a particular idea of representative sampling, and its status as an ideal, is not a given. Even if the ideal statistical sampling model is followed, there are long-standing concerns about its assumptions and logic. For example, classic statistical theory states that representativeness must be achieved on each variable measured in the research, not just for a few choice demographics, such as age and sex³⁴. In practice, a household survey with 100 variables will only create weights on variables that can be matched to population distributions. It seems almost impossible to actually follow the classical statistical model for conducting a probability sampling design in the social sciences.

So, does this matter? It depends on the kinds of inference we wish to make from our research. Pasek explores the utility of non-probability samples in answering 3 types of questions: (1) How are variables distributed in society? (2) How are variables related to one another? and (3) How do variables change over time?¹⁴. Keiding and Lewis write about the ‘perils and potentials of self-selected entry to epidemiological studies and surveys’ in an article which elicited a wide-ranging debate among epidemiologists and statisticians on the issue of the utility of opt-in sampling¹⁵. In both publications, we can see that there is greater support for deriving inferences from non-probability samples for questions of the relationships between variables and trends in variables over time, under particular conditions. In contrast, the use of non-probability sampling methods for point estimation is usually considered inappropriate, although there are a number of statistical and modelling methods applicable only to some kind of non-probability sampling methods that may be used for prevalence estimation, for example, in the case of respondent driven sampling^{35,36}. Prevalence estimates are important. For example, they are used to direct investment into service delivery and to estimate the size and distribution of a particular problem in the wider population. But just knowing the prevalence of a particular behaviour in a population tells us very little about the patterns and practices of that behaviour and its health impact among large cohorts of affected populations. Understanding these patterns and practices of behaviour are equally important when trying to inform public health responses. The GDS is a response to this very need in relation to drug use.

Aims of this paper

Our paper has two aims: (1) to present a comprehensive account of the history, orientation and methodology we use at GDS, and (2) to compare the age and sex distributions of cannabis users who voluntarily completed (a) a household survey or (b) a large web-based purposive survey (GDS), across three countries: Australia, the US and Switzerland. These two aims are addressed in two parts below, with the discussion and conclusion bringing them together.

Part 1 – Global Drug Survey

Global Drug Survey runs the largest annual anonymous web survey of people who use licit and/or illicit psychoactive drugs. GDS partners with global media organisations who help promote the survey and share the findings with their readers usually 4-6 months after closure of the survey each year. GDS uses a cross-sectional design. Data have been collected annually for the past 6 years (see Table 1). Prior to the launch of GDS2012¹, the previous annual surveys led by Adam Winstock were targeted only at a UK audience, supported by the UK clubbing magazine *Mixmag* and the UK news organisation *The Guardian*. These

¹ The GDS naming convention refers to the year that the findings were released, not the year that the survey began data collection. That is, GDS2012 data were collected in late 2011, and released in mid-2012.

Mixmag surveys had been annually conducted between 1999 and 2004. They were originally distributed in print as part of the *Mixmag* magazine and later also available as a web survey targeted at the *Mixmag* reading population through their sister website 'Don't stay in'. This survey was relaunched after a 5-year hiatus in 2009 and formally became known as the Global Drug Survey in 2011. The most recent surveys have been approved by human research ethics committees at Kings College London (and the University of Zurich in 2014).

[Insert Table 1 here]

Orientation and business model

The *Mixmag* surveys, and later, the GDS surveys, adopted a specific orientation towards the target population, with a focus on non-treatment-seeking drug users and an emphasis on curating information that was useful to those who used drugs. While the *Mixmag* era specifically targeted participants in the nightlife economy, the population targets were broadened to include anyone with a drug-use history (including legal drugs) as the GDS widened its scope of media partners and increased its reach.² Nevertheless, the philosophy of engagement remained stable—we make clear to potential participants that we are independent of government, self-funded and do not assume that all drug use is harmful. Instead, we acknowledge drug-related pleasure and assume that most people who use drugs are interested in reducing their risk of drug-related harm. Also, we engage in honest conversations about drugs through our research, website and harm reduction tools. Our main goal in conducting the GDS is to make drug use safer, regardless of the legal status of the drug³⁷. The approach to engagement includes a casual voice, humour, and above all, a statement about the orientation of our survey: that we are not beholden to funding bodies (government or otherwise) that often skew research towards examining the harms and ignoring the benefits of drug use, that we share our findings rapidly with the broader communities in accessible ways (e.g. media stories and opinion pieces, not just academic articles). Perhaps most importantly we accept that the most credible source of drug information for people who use drugs are not doctors, academics or government authorities but other people who use drugs. For GDS, the participants are viewed as the experts and we invite them to share their expertise with us. In return, we are offering community benefits, including free access to digital harm reduction tools, including the Highway Code, Safer Use Limits, Drinksmeter and Drugsmeter (see: <https://www.globaldrugsurvey.com/free-online-resources/>) and regular public engagement into pertinent debates around drug policy reforms. However, no financial incentives are provided for participation. For example, Box 1 shows the call for participation in GDS2014. This text emphasised that participants can contribute to changing the inaccurate and incomplete stereotypes that often pervade public discourse around illicit drugs.

[Insert Box 1 here]

As mentioned above, as part of our efforts to engage people who use drugs (who may be suspicious of the motives of researchers), we draw on our business model to convince them of our independence. The business model we use is unusual in the academic world. Typically, research projects are funded through competitive grants, tenders, or consultancies with government or corporate entities. Research proposals are submitted, assessed and a selection of them are funded. Research begins once the funding is received. There are three problems with the existing funding model. Firstly, it is typically relatively slow, with lag-times of months to years between submitting a project application and the commencement of funding. This lag greatly impedes rapid research and the identification of new trends. Secondly, project ideas must be written to match the predispositions of the funding bodies. Thirdly, reviewers are typically established in their fields and may not welcome orientations that do not necessarily coincide with the traditional worldview held. The GDS business model works in reverse to the usual funding model. The GDS operates without core funding. It can do this because of initial seed funding from Dr Adam Winstock who owns GDS and the mutual benefit to academic researchers who volunteer their expertise and time to the survey operation in exchange for access to the data, from which they can publish academic papers—an activity which is supported by their university and is essential to their success as researchers. GDS is also able to do this because of the

² Including to populations who inject drugs through partnership with the International Network of People Who Use Drugs (INPUD) in 2015 and 2016.

relatively low cost of running web surveys. While some media partners agree to promote the survey freely in the recruitment stage in exchange for exclusive stories in the dissemination stage, many media partners are willing to compensate GDS for the privilege of access to tailored news stories. It should be noted that none of the media partners have influence over the content of the survey, and we also have no control over what our media partners choose to present. While GDS has developed a number of free digital harm reduction tools which are available on its website, it also receives payment from government entities and companies that use its harm reduction tools in clinical applications such as delivering brief screening and intervention for alcohol (e.g., Drinksmeter, ³⁸). GDS also provides bespoke data reports to health and corporate organisations, and is beginning to receive funding from universities and other organisations that wish to include survey modules within GDS. GDS does not accept money from the alcohol and tobacco industries. The funds GDS raise arise from the value of the data already collected, and as these funds grow, GDS may begin to be able to compensate its core team for its labour in more than just data access. This unique business model has caused some confusion; for example, for consumers of research reports who are used to being able to access them for free. Instead, GDS operates a ‘freemium’ model, where basic findings are publicly accessible as quickly after data collection as possible, while more detailed data reports must be commissioned.

Preparation of the survey

There is a core set of questions GDS uses annually which provide continuity and comparability between different data collection years. As well as monitoring changes in drug trends, GDS includes specialist topics each year addressing areas of current interest identified by an international expert advisory group and the academic network that now spans 22 countries. Unique modules may be proposed by research group partners to rapidly address emerging issues. For example, when the online drug market Silk Road first emerged, the GDS was the first to survey Silk Road users in late 2012 ⁷, and has been at the forefront of surveying users of cryptomarkets in the years since ³⁹. Through the administration of the ‘New Drugs’ core module, GDS has authored papers that are the first to recruit large samples of the users of new or emerging substances, including mephedrone ⁴, synthetic cannabinoids ^{6,40,41}, the NBOMe series ⁴², DMT ⁴³, nitrous oxide ⁴⁴, and methoxetamine ^{45,46}. Modules on alcohol and prescription medication use have resulted in papers on alcohol’s harm to others ⁸ and profiles of the prescription medication Tramadol ^{47,48}.

An important aspect of the survey design is participant anonymity. The GDS website is encrypted and no IP addresses are collected meaning that data cannot be matched with specific individuals. Storing IP addresses is not ethically appropriate for a population who report illegal behaviours and therefore values anonymity ⁴⁹. IP addresses can be used by multiple people validly (e.g., share houses, university dorms) ⁵⁰. The GDS offers an option at its completion for participants to provide contact details, stored separately in an encrypted database, should they give permission for researchers to recontact them for further research opportunities. Missing data are allowed in GDS; rather than forcing respondents to provide an answer to all questions, we have erred on the side of reducing annoyance for the respondent who may wish to skip a section without being forced to answer it. GDS2017 was the first GDS wave to use proprietary survey software, Survey Gizmo, which offers cross-platform accessibility (mobile, tablet and computer). Prior to GDS2017, GDS waves were written for the web directly via PHP code. Since GDS2014, the survey has been translated into multiple languages; GDS2015, GDS2016 and GDS2017 were available in English and 10 other languages. Participants choose their language on the front page and are then directed to the survey. Translation has relied on the partner organisations in the respective countries providing translations to fragments of text, which were then entered into the survey. As shown in Table 1, the number of partner countries has grown exponentially over 6 years.

Before launch, the extended GDS network pilots the survey and provide lists of errors, inconsistencies and items that lack clarity or reflect unfounded assumptions. This iterative process continues as the survey is improved. Inevitably, some problems are not identified through piloting, and we always invite participants to contact GDS directly once the survey launches to identify problems, which are either dealt with immediately if possible, or recorded to fix for next year. This is another example of adopting community approach and learning from our participants.

Recruiting the participants

GDS has become the largest web survey of drug use in the world through its collaboration with some of the world's most well-known global media organisations (see Table 1). These partnerships have evolved and increased over the years and once initiated have proved remarkably stable. GDS2017 had media partners in over 20 countries. The survey typically launches in the second week of November and runs for a period of about 6-8 weeks until the end of the year. Our media partners act as hubs, providing initial direct access to our target group. In return for their support we offer exclusive data reports that are attractive to this group. Successful partnerships between media organisations and GDS involve working together prior to survey launch, determining the drug-related media topics of interest, and the expertise the GDS team can offer. The media pieces that recruit for GDS are typically the result of independent journalism which may be based on backgrounding GDS experts (e.g., ^{51,52}), but may also include GDS-written opinion pieces (e.g., ^{53,54}). Stand-alone media content will include advertising for participants and calls for participation within the article text. While the media partnerships provide direct survey recruitment, secondary recruitment occurs through social media sharing of media partner content that mentions GDS on Facebook, Twitter, Reddit and other web discussion groups, including drug discussion forums. Key members of the GDS network also share this content and motivate participation through their networks with nightlife settings where drug use is overrepresented. The GDS network also includes non-government organisations and harm reduction groups who also promote the survey to their networks.

The success of recruitment through media stories is heavily influenced by world events that may dominate the media cycle, crowding out the media stories that lead to recruitment of participants to GDS. For example, in November 2015, the Paris bombing occurred during the GDS launch, which hindered recruitment, resulting in the need to extend the recruitment period beyond its usual 6-8 weeks, see Table 1.

Writing up and disseminating results

Once recruitment is complete, data cleaning and checks are conducted prior to any preliminary analysis. A key validity issue for web surveys is ensuring that participants only complete the survey once ⁵⁵. The raw data are screened for duplicate responses that may result from data glitches or respondents completing the survey more than once (either accidentally or intentionally). Duplicate removal is two staged. The first stage removes all records that are complete matches; specifically, two or more of records which are completely identical. The second stage removes all records where duplication is present on a series of demographic variables and drug use variables captured in the drug screen module. Demographic variables include: age, sex, ethnicity, educational attainment, employment type, income, country, height, weight, clubbing activity, exercise and standard of living. Drug screen variables include ever, past year, past month, age of first use and frequency of use. Second and subsequent duplicate records are removed. Because there are no material incentives offered (e.g., lotteries, prizes, payments), and because the survey typically takes 15 minutes to 1 hour to complete (depending upon drug experience), we believe it is unlikely that anyone would deliberately complete the survey more than once for personal gain. Other data checks that may result in the removal of cases at this stage include: participants reporting the use of a fictitious drug and participants who report no psychoactive drug use at all. Once the number of cases in the dataset is finalized, data modules are checked, cross-checked and cleaned. Questions which should be dependent on an answer to an earlier question are cleaned according to sets of rules: for example, if a respondent reports first trying LSD through an online drug market, but does not report ever use of LSD in the earlier drug screen, the later data is removed to ensure data consistency, given primacy to the accuracy of the earliest response.

Following data cleaning, country reports are produced to a set template. These are distributed confidentially to the media partners of that country. The media partners use the reports to write exclusive stories which are released on a pre-agreed launch date, usually in May or June, or 5 to 6 months after data collection ends (e.g., ⁵⁶). This time period ensures that the trends are timely. More detailed analyses are published as academic papers and produced within tailored data reports.

Limitations and strengths

Being a non-probability sample that excludes people who do not use licit and/or illicit psychoactive drugs and who do not use the internet, GDS does not seek to answer questions about population drug use prevalence. Non-response bias, where there are inherent differences between those who participate and those

who do not¹⁷, and volunteer bias, where people are more likely to respond if they are interested in the topic⁵⁷, may both influence the composition of GDS samples. Both of these biases may also affect probability samples, but in theory, such biases can be adjusted for through weightings that are based on knowledge about the people who do not respond. With GDS, we cannot estimate the characteristics of non-responders. It is important to note, though, that weighting probability samples cannot fully account for sampling bias^{15,17,34}. For example, the data from young males who complete a household survey will be weighted more heavily to make up for the lower participation rates of their cohort, yet the young males that respond are likely to have different characteristics to the young males who do not respond. That is, there are unmeasured confounders¹⁵. This is one of the reasons why it is highly likely that household surveys underestimate the prevalence of illicit drug use, as the segments of the population that use drugs are less likely to be reached or if reached, to agree to participate. Stigma has also been shown to influence prevalence estimates from household surveys, as individuals who use drugs become more concerned about admitting to such behaviours, especially to a government-run survey⁵⁸. There may also be underreporting of drug use using face-to-face or telephone survey modes⁵⁹, perhaps due to fear of being identified. Because GDS targets people who use drugs through web recruitment, we access a relatively high proportion of younger respondents who can be difficult to capture on telephone or in face-to-face surveys. So, while the sample is self-selected and may suffer from biases, many of these biases are also present in the so-called “gold standard” of the household survey, which is increasingly less likely to be able to access the young and mobile populations¹⁷ that are the most likely to have lived experiences of illicit drug use. Another strength comparing GDS to household surveys is the cost. Probability sampling methods are expensive, particularly when rare behaviours are targeted. At a fraction of the cost, GDS can access large numbers of respondents who report relatively rare behaviours. We also have the unique advantage of being able to present the same set of questions in multiple languages, enabling country comparisons. It is difficult to achieve this level of standardisation through pooling data from household surveys from multiple countries, which inevitably use different question wording.

Part 2 – Testing the utility of Global Drug Survey

Given the limitations and strengths of GDS compared with household surveys using probability sampling methods, in this paper, we compare the similarities between predicted probabilities of recent (past 12-month) and current (past 30-day) use of cannabis amongst subsamples of self-reported lifetime cannabis users. The GDS sample used for analysis deliberately excludes people who report no drug use ever (including alcohol), so we would not expect the probabilities of lifetime use to be similar. However, if we find similar patterns of past year and past month use *within samples reporting lifetime use*, we can have greater confidence that the characteristics of those who volunteer to complete the GDS reflect the characteristics of those who volunteer to complete household surveys, conducted in the same locality and within a similar timeframe. In this paper we conduct this analysis with cannabis users from Australia, the US and Switzerland, and report results from a similar analysis on alcohol use in the supplementary appendix.

Method

Global Drug Survey (GDS)

GDS is an annual, international, web survey of drug use which is self-completed, largely by younger individuals, on a self-nominating, anonymous basis. In this paper we employ data from the GDS2014 (see Table 1 column 3 for more information). In the current study, the sample was restricted to respondents living in Australia (N=5,789), the US (N=6,419) and Switzerland (N=4,971). All respondents reported being 16 years or older. Of the total 16,828 respondents that reported their age from the selected countries, 10,708 (63.6%) were male and 6,120 (36.4%) were female. The mean reported age was 34.2 years (SD=13.9). The GDS asked participants a range of questions related to the use of alcohol and illicit substances. To assess lifetime and recent alcohol use, all participants were asked if they had ever used alcohol and if they had used alcohol in the past 12 months. All GDS participants were also asked if they had ever used or used in the past year: Cannabis/marijuana- herbal-high potency (hydro); Cannabis/marijuana – herbal-normal weed (e.g., bush weed/pressed); Cannabis (resin/hash); and Cannabis oil. GDS participants who reported using at least one type of cannabis in the past 12 months are asked “How many days have you used this type (the type used most commonly) of cannabis in the last month?” This item was recoded into a dichotomous indicator

of past month cannabis use. Participants were coded as “1” on this variable if they reported using cannabis on at least one day in the past month and a “0” if they reported no use in the past 30 days.

National Drug Strategy Household Survey (NDSHS, Australia)

The National Drug Strategy Household Survey (NDSHS) is a national survey conducted by the Australian Institute of Health and Welfare every three years⁶⁰. The 2013 sample comprised 23,855 people aged 12 years or older residing in private residences in Australia. Sampling was conducted via a multi-stage, random approach stratified by geographic area to best provide for national representativeness. The survey was conducted via a self-completion, drop and collect method. Three attempts were made to establish contact with households at drop-off and again at collection. The response rate for in-scope households was 49.1%. Despite careful sampling design and implementation the NDSHS sample over-represents two-parent families, individuals with higher levels of education and those not currently employed while under-representing non-English speaking and highly socio-economically disadvantaged residents.

The NDSHS asks participants to report on their use of a range of substances including alcohol, tobacco, cannabis and other illicit drugs. To assess lifetime and recent alcohol use, all participants in the NDSHS were asked “Have you ever had a full serve of alcohol? For example, a glass of wine, a whole nip of spirits, a glass of beer etc.”; “Have you had an alcoholic drink of any kind in the last 12 months?” To assess lifetime and recent use of cannabis, all participants were asked: “Have you ever used Marijuana/Cannabis?”; “Have you used Marijuana/Cannabis in the last 12 months?”; “Have you used Marijuana/Cannabis in the last month?” Different forms of cannabis, such as leaf, head, resin, oil or other, are not clarified until after the above questions are answered. All participants in the sample were also asked standard questions regarding demographic characteristics including current age (in years) and sex.

National Survey on Drug Use and Health (NSDUH, US)

Data were obtained from the 2013 year of data collection from the National Survey on Drug Use and Health (NSDUH), an ongoing cross-sectional survey of non-institutionalized individuals in the 50 US states and in the District of Columbia⁶¹. NSDUH is a nationally representative probability sample of individuals residing in households obtained through four stages. First, Census tracts were selected within each state, then segments in each tract were selected, then dwelling, and finally respondents were selected for final sample. Surveys were administered via computer-assisted interviewing and audio computer-assisted self-interviewing. Sampling weights were provided by NSDUH to address unit- and individual-level non-response and to ensure that estimates are consistent with estimates provided by the US Census Bureau; however, we did not utilize sampling weights for the following analyses. The (unweighted) interview response rate in 2013 was 76.4%⁶².

Participants were asked their sex (i.e., male, female) and age. NSDUH provided pre-coded categorical variables indicating participant age, which we recoded into: age 16-20, 21-25, 26-34, 35-49, and 50 and older. Participants were asked about use of alcohol and cannabis (“marijuana”). An alcoholic drink was defined as a can or bottle of beer, a wine cooler, a glass of wine, a shot of liquor, or a mixed drink containing liquor. Participants were notified that having only a sip or two was not defined as having an alcoholic drink. They were then asked if they had ever, even once, had any type of alcoholic beverage. Among those who reported lifetime use, they were then asked about recency of past use, which were coded as lifetime use, but no use in the past 12 months, 12-month use, but not in the past 30 days, and use within the past 30 days. The same questions were then asked with regard to marijuana. Participants were reminded that marijuana is also called pot or grass, and that it is usually smoked in joints or a pipe, and it sometimes comes in food. Hashish (“hash”) was also included in this definition and it was noted that hash is usually smoked in a pipe and it also comes in oil form. Among those who reported lifetime use, they were then asked about recency of past use, which were coded as lifetime use, but no use in the past 12 months, 12-month use, but not in the past 30 days, and use within the past 30 days.

Addiction Monitoring in Switzerland (AMIS, Switzerland)

The Addiction Monitoring in Switzerland (AMIS) is a cross-sectional national survey that has been conducted by the Swiss Federal Office of Public Health (FOPH) every year from 2011 to 2016⁶³. Data were obtained from telephone interviews conducted during a one year period starting from July 2013 to June 2014 (wave 6/7). The sample comprised 12,008 Swiss residents aged 15 years or older. AMIS uses a dual frame approach to increase representativeness: 91.5% of the calls used landline numbers and sampling was

conducted via a multi-stage, random approach stratified by geographic area to best provide for national representativeness. On the other hand, 8.5% of the calls used mobile phone numbers that were generated by Random Digit Dialing (RDD). The survey was conducted via a self-completion, drop and collect method. In 2013 people aged 15-29 years were intentionally overrepresented. Up to 20 attempts within 30 days were made to establish contact with households at drop-off and again at collection. The response rates in 2013 were 55% of landline calls and 15% of mobile-phone calls. In 2014, the response rates were 44% and 13%, respectively. The landline interviews were 25-30 minutes long, while the mobile phone interviews contained just the core questionnaire and were, therefore, shorter (10-15 minutes long). Despite careful sampling design and implementation the AMIS sample over represents couple families, individuals with higher levels of education and those not currently employed while under representing highly socio-economically disadvantaged residents. Sampling weights were provided by AMIS to address unit- and individual-level non-response and to ensure that estimates are consistent with estimates provided by the Swiss Population and Households Statistics. The Population and Households Statistics are part of the surveys conducted within the framework of the Federal population census. Sampling weights were not considered for the following analyses.

The AMIS asks participants to report on their use of a range of substances used including alcohol, tobacco, cannabis and other illicit drugs. The survey language was dependent on the language region and could be either German, French or Italian. To assess lifetime and recent alcohol use, all participants in the AIMS were asked “Have you ever had a full serve of alcohol? For example, a glass of wine, a whole nip of spirits, a glass of beer etc.”; “How often have you consumed alcoholic drinks such as beer, wine or liquors in the last 12 months?” To assess lifetime and recent use of cannabis, all participants were asked: “Have you ever used Hash/Marijuana/Cannabis?”; “Have you used Hash/Marijuana/Cannabis in the last 12 months?”; “Have you used Hash/Marijuana/Cannabis in the last 30 days?”. Participants were also asked their sex (i.e., male, female) and age. AMIS provided pre-coded categorical variables indicating participant age: age 15-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75 or older.

Analysis

The GDS2014 data were restricted to Australian, US and Swiss residents for comparison with the NDSHS, NSDUH and AMIS, respectively. The analytic sample in each dataset comprised respondents who recorded responses to questions about cannabis use (yes/no), sex (male/female) and age.³ Age in the GDS and the NDSHS is a continuous variable. Age in the US and Swiss national survey is a categorical variable. Age categories are listed in Table 2. Unweighted data for each of the national surveys was used for comparison with GDS respondents. While it would be correct to weight the data if the intention was to use the survey to estimate population prevalence, in this analysis, we are not interested in prevalence because we cannot estimate prevalence from the GDS. Instead, we are interested in comparing the characteristics of the people who actual complete the household survey with those who complete the non-probability survey (GDS). Predicted probability analyses were conducted using Stata 14⁶⁴. Descriptive statistics are presented in Tables 2 and 3. Although valid percentages are used throughout, for data transparency, missing data counts are presented in Table 2. Results of the predicted probability analyses are presented in Figures 1-3.

Results

[Insert Tables 2 and 3 here]

Australia

As shown in Tables 2 and 3, the analytic sample from the NDSHS (N=23,855) comprised 10,624 (44.5%) males and 13,231 (55.5%) females with a mean age of 47.5 years (SD 16.7). With respect to cannabis use, 8,101 (34.4%) indicated they had used cannabis in their lifetime, 2,174 (26.8% of lifetime users) reported using cannabis in the past 12 months and 1,143 (52.7% of past year users) reported use in the past month.

A total of 5,789 Australians participated in GDS2014. This sample comprised 3,479 (61.0%) males and 2,223 (39.0%) females with a mean age of 37.13 years (SD 13.1). With respect to cannabis use, 4,063

³ The analytic sample for supplementary analyses examining predicted probability of alcohol use comprised all participants who recorded responses to questions about alcohol use (yes/no); sex (male/female) and age.

(70.2%) reported using cannabis in their lifetime, 2,037 (50.1% of lifetime users) reported using cannabis in the past 12 months and 1,379 (67.7% of past year users) reported using the drug in the past month.

Figure 1 presents the age-based predicted probabilities of reporting cannabis use across the lifetime; within the past 12 months and within in the past month by sex for each of the Australian datasets. For both the GDS and the NDSHS data, regardless of age, males are typically more likely to report using cannabis than females: this can be observed for lifetime use and past 12 month use in both datasets. However, the bottom figures show some crossover between the sexes for reported cannabis use in the past month with females at particular ages more likely to report cannabis use compared to males. For the GDS sample this crossover occurs when the respondents are aged 69 years or older. For the NDSHS sample, this crossover occurs when respondents are aged between 44 years of age and 64 years of age.

For respondents 40 years and over, overall, males and females show similar trends of a decreasing probability of ever using cannabis in their lifetime. Notably, the probability of GDS respondents ever using cannabis is typically higher than the NDSHS sample likely due to the sampling methodology. Furthermore, younger male respondents (those less than 40 years) in the GDS were more likely to report ever using cannabis compared to their counterparts in the NDSHS sample. The patterns for cannabis use within the past year, from both samples, are much more similar across the ages. For respondents who had used cannabis in the past month, the pattern for females was similar, especially for those between the ages of 25-30 years and 55-60 years of age. For both datasets, as the age of female respondents increased so did the probability of using cannabis within the past month. For females over 55-60 years of age, in the GDS sample, the slope indicating the predicted probability of using cannabis in the past month was not as great as observed in the NDSHS sample. By contrast, except for the younger males in the GDS sample (those 30 years of age or less), the patterns were more similar.

While the probability of ever using cannabis is higher in the GDS sample, once the datasets are restricted to respondents who have used cannabis in their lifetime, the probability of reporting cannabis use in the past year and, in to some degree, the past month are similar across the GDS and the NDSHS samples.

[Insert Figure 1 here]

US

As presented in Tables 2 and 3, the analytic sample from the NSDUH (N=43,465) comprised 20,302 (46.7%) males and 23,163 (53.3%) females. The greatest proportion of respondents was aged 16-20 years (29.3%) and 21-25 years (26.3%). With respect to cannabis, 21,075 (48.5%) reported lifetime use, 9,498 (45.1% of lifetime users) reported using cannabis in the past 12 months and 5,664 (59.6% of past year users) reported cannabis use in the past month.

A total of 6,419 US residents participated in GDS2014. The US sample of GDS respondents comprised 3,846 (61.2%) males and 2,439 (38.8%) females. The distribution of respondents was similar across age categories with 18.9% aged 16-20 years, 19.8% aged 21-25 years, 20.4% aged 26-34 years, 20.0% aged 35-49 years and 20.9% aged over 50 years. With respect to cannabis, 5,774 (90.0%) reported they had used it sometime during their lifetime, 4,519 (78.3% of lifetime users) indicated using in the past 12 months and 3,879 (85.8% of past year users) had used the drug in the past month.

Figure 2 presents the age-based predicted probabilities of reporting cannabis use across the lifetime, within the past 12 months, and within the past month by sex for each of the US datasets. For both the GDS and the NSDUH data, regardless of age, males were typically more likely to report using cannabis than females. However, the bottom figures depict some crossover between the sexes for reported cannabis use in the past month for respondents between the ages of 36-49 and 50+ years. This is observed in both the GDS and the NSDUH data. In both cases, females were more likely to report cannabis use compared to males.

In both the GDS and the NSDUH data, males and females typically show similar trends of a decreasing probability of lifetime and past year cannabis use with age. Alternately, the probability of cannabis use in the past month among those who reported using cannabis in the past year remains relatively flat across the age categories; although as the age-group of females from the NSDUH sample increases, there is a slight increase in the probability of reporting cannabis use in the past month (this is only observed in the GDS data for females 26-34 years and older).

While the probability of ever using cannabis is higher in the GDS sample, the probability of using cannabis in the past year among lifetime users, and using within the past month among past year users, is similar across the GDS and NSDUH datasets.

[Insert Figure 2 here]

Switzerland

As presented in Tables 2 and 3, the analytic sample from the AMIS (N=12,008) comprised 5,542 (46.2%) males and 6,466 (53.8%) females. The greatest proportion of respondents was aged over 50 years (38.3%) and 20-29 years (20.3%). With respect to cannabis, 3,465 (28.9%) reported lifetime use, 1,057 (30.5% of lifetime users) reported using cannabis in the past year and 481 (45.5% of past year users) reported cannabis use in the past month.

A total of 4,972 Swiss residents participated in GDS2014. The Swiss sample of GDS respondents consisted of 3,381 (69.9%) males and 1,456 (30.1%) females. The greatest proportion of respondents was aged 20-29 years (43.5%) and 30-39 years (22.0%). With respect to cannabis use, 3,236 (65.1%) reported lifetime use, 2,056 (63.5% of lifetime users) reported using cannabis in the past year and 1,599 (77.8% of past year users) reported using cannabis in the past month.

Figure 3 presents the age-based predicted probabilities of reporting cannabis use across the lifetime, within the past 12 months and within the past month by sex for the Swiss datasets. For both the GDS and the AMIS data, regardless of age, males were typically more likely to report using cannabis than females. Only in the GDS did females report higher prevalence of cannabis use than males for use in the past month. As shown in the GDS figure for past month use, there is a constant, slightly increasing probability of males reporting use of cannabis within the past month across the age groups. However, for females, the probability of using cannabis within the past month increases substantially between the ages of 20-29 year and 40-49 years at which point the probability of using cannabis in the past month is greater for females than for males.

Overall, males and females show similar trends of a decreasing probability of lifetime and past-year cannabis use with age (especially after 20-29 years of age). Alternately, the probability of reporting cannabis use in the past month among those who reported using cannabis in the past year trends upwards across the age categories.

The probability of reporting cannabis use is similar across the GDS and the AMIS datasets, particularly for those aged over 20 years. For those who have used cannabis in the past year, the probability of having reported using cannabis in the past month is lower for the AMIS sample. This is likely owing to the small number of participants from Swiss GDS sample who were aged over 50 years and reported cannabis use in the past year.

[Insert Figure 3 here]

Discussion

While the probability of ever using cannabis is higher in the GDS sample, as would be expected, once the datasets are restricted to respondents who have used cannabis in their lifetime, the probability of reporting cannabis use in the past year and, in to some degree, the past month are similar across the GDS and the equivalent probability samples, across the three countries. That is, we can see in Figures 1-3 a decreasing probability of reporting past-year use with age among lifetime users, present across both kinds of samples, and a stable or flatter probability of reporting past-month use with age among past-year users, present across both kinds of samples. These findings give us some confidence that samples of past-year or past-month cannabis users recruited through the GDS do not contain a highly skewed group in terms of their age or their sex. These findings show that the age and sex distribution of those who volunteer to be surveyed about their cannabis use is not vastly different between non-probability and probability samples, across the three countries of Australia, the US, and Switzerland.

The efficiency of accessing past-month users of cannabis through the GDS is also illustrated in Table 2. For questions other than prevalence of use in the general population, it is important to access a large and diverse sample. In the case of Australia, the GDS recruited 1,379 past-month cannabis users from surveying 5,789, compared with 1,143 past-month cannabis users from surveying 23,855 via the household survey. While in the US, a greater number of past-month cannabis users was recruited through the probability survey (5,664)

compared with the GDS (3,879), the total sample surveyed by the US probability survey (43,465) was almost 7 times the total number surveyed by GDS (6,419). In Switzerland, the GDS recruited 1,599 past-month cannabis users compared with 481 recruited by the probability sampling. While it is beyond the scope of this paper to estimate the exact cost comparisons, it is clear that the additional costs associated with accessing these past-month cannabis users samples by probability sampling methods would be considerable. Furthermore, it is unlikely that the level of detail that GDS can obtain would be possible through household surveys, where the time and space needed for extra questions is at such a premium.

With reference to our earlier discussion about representativeness and whether it matters, again, we need to return to the question: what kind of inferences do we wish to make from our research? ^{14,15} The analysis presented in this paper provides support for the use of non-probability sampling methods where large samples are collected to find out about a hidden practice or the characteristics of a hidden population for the purposes of understanding these practices or populations in greater depth. Relationships between variables within affected populations and trends over time can be answered, while questions about estimating population prevalence cannot. If the people who are reached and choose to complete population-based surveys are not vastly different from the people who are reached and choose to complete non-probability surveys, we can have greater confidence in employing methodologies like GDS. Further work of this kind should be conducted, e.g. looking at characteristics of populations of novel substances users reached by probability vs non-probability sampling. Such work is dependent on probability-based surveys including the relevant variables for comparison.

Limitations

There are several limitations of the analyses presented in this paper. First, the use of age categories in the NSDUH and AMIS prevented a more nuanced probability analysis as was conducted with the Australian data. Further, as the GDS tends to attract younger participants, the number of GDS respondents in the age 50+ years category who recently used cannabis was small. This was particularly the case in the Swiss sample. We also recognise that differences in the method of survey administration (e.g., GDS uses a web interface, NDSHS is drop and collect, NSDUH uses computer-assisted personal interviewing (CAPI) and AMIS is conducted through landline and mobile telephone) and wording of survey items may influence participants' responses to questions about drug use. For example, while the NSDUH refers to marijuana, the GDS asks participants about cannabis. Further, while the NDSHS asks participants whether they have used cannabis in the past month, the GDS asks how many days in the past month participants have used. To minimise the effect of differences in survey item wording we have focussed on cannabis, a relatively common substance that would be familiar to most participants, and have limited the analysis to examine dichotomous indicators of cannabis use.

In terms of the GDS methodology, no method for mapping and understanding drug use patterns is without limitations. Increasingly, it is accepted that triangulation of data sources is needed to best represent the 'dynamic epidemiology' that characterises drug use patterns across the world ⁶⁵. Data can be gleaned from waste water analyses, sentinel populations, police seizures, treatment services, toxicology services, border control or customs, harm reduction services in nightlife settings (including drug checking or testing data), as well as surveys of general and targeted populations. What is required is the time and capacity to synthesize these diverse data sources into the most robust and valid picture of drug use within our communities. GDS provides one part of this complex picture, and therefore should not be understood and interpreted in isolation. In addition to the need for triangulation, there are improvements GDS can make in its procedures to be more inclusive. For example, we could recruit booster samples from groups less well represented (e.g. females, those from non-white backgrounds) by partnering with more focused media groups and improving utilisation of targeted social media recruitment. The aim is to increase the knowledge-base about the drug use patterns and behaviours associated with these cohorts to allow further segmentation of drug using populations to inform both public policy and health promotion. We could also implement more standardised procedures around translation, including back translation (e.g. translation from English, to Spanish, then back to English), to better ensure validity of measures across our languages. Another limitation of GDS is that dropout rates have typically not been tracked: that is, we are unable to calculate the difference between the number of respondents who began participating versus those who completed participation.

Conclusion

In this paper, we have assessed the utility of GDS to recruit samples of cannabis users with similar age and sex characteristics to those produced by probability sampling. With cannabis being the most commonly reported illegal drug globally, we find similar age-based predicted probabilities of recent and regular cannabis use compared to equivalent household surveys, across three separate country comparisons. Unfortunately, it is not possible to conduct the same comparisons for the distributions of novel drugs, because they are typically not measured in household surveys until many years after their emergence. However, the similar patterns among cannabis users give us greater confidence to the capacity of the GDS to produce samples of sufficient diversity to provide useful and meaningful results, while offering a depth of content analyses that is unlikely from general household surveys, with their expense and time constraints.

This paper provides the necessary level of procedural detail on the GDS methodology to complement the interpretation of our results-oriented papers. It answers the questions raised by many of those who have reviewed our papers. We hope that, along with other publications led by the first author^{22,25,49,66,67}, this paper helps to demystify the conduct of internet research with groups that are understandably suspicious of researchers, who have historically misrepresented their issues and perpetuated, rather than challenged, stereotypes of drug use and drug users. We conclude that opt-in web surveys of hard-to-reach populations are an efficient way of gaining in-depth understanding of stigmatized behaviours, and are appropriate, as long as they are not used to estimate drug use prevalence of the general population.

Acknowledgements

We would like to thank the participants who gave so generously of their time to complete the GDS. We are grateful for the promotion of GDS by a long list of world media partners, see Table 1. We are also indebted to Stuart Newman for his programming skills, advice and patience and Chris Parsons for his hard work in managing the translation site for the survey. We thank the Australian Institute of Health and Welfare (AIHW) and the Australian Data Archive (ADA) for providing access to the NDSHS data, and the Swiss Federal Office of Public Health: Addiction Monitoring in Switzerland 2013-2014 for providing access to the AMIS dataset (see <http://suchtmonitoring.ch/>). We also thank RTI International for funding the open access fee for this paper. RTI International is an independent, non-profit research organisation dedicated to conducting research that improves the human condition.

Funding sources

The authors received no specific financial support for the research, authorship, and/or publication of this study. MB and JF were supported by fellowships from the Australian National Health and Medical Research Council (APP1070140, APP1089395). JP was supported by the US National Institutes of Health (NIH) (K01 DA-038800, PI: Palamar). The National Drug and Alcohol Research Centre and the National Drug Research Institute were supported by funding from the Australian Government under the Substance Misuse Prevention and Service Improvement Grants Fund. We also acknowledge the contribution of the Victorian Operational Infrastructure Support Program received by the Burnet Institute.

Author Contributions

MB, JF, LM and AW conceived of and developed the paper and together form the GDS Core Research Team. MB wrote the first draft of the manuscript. RZ, JF, LM and JP conducted statistical analyses. All authors reviewed, made critical revisions and approved of the final manuscript.

Declaration of interest

AW is founder and managing director of Global Drug Survey.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Baker R, Brick JM, Bates NA, et al. *Report of the American Association for Public Opinion Research (AAPOR) task force on non-probability sampling*. Oakbrook Terrace, IL: AAPOR;2013.
2. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012-1014.
3. Gobo G, Mauceri S. *Constructing survey data. An interactional approach*. London: Sage; 2014.
4. Winstock AR, Mitcheson LR, Deluca P, Davey Z, Corazza O, Schifano F. Mephedrone, new kid for the chop? *Addiction*. 2011;106:154-161.
5. Winstock AR, Mitcheson L, Gillatt DA, Cottrell AM. The prevalence and natural history of urinary symptoms among recreational ketamine users. *BJU Int*. 2012;110(11):1762-1766.
6. Winstock AR, Barratt MJ. Synthetic cannabis: A comparison of patterns of use and effect profile with natural cannabis in a large global sample. *Drug Alcohol Depend*. 2013;131:106–111.
7. Barratt MJ, Ferris JA, Winstock AR. Use of Silk Road, the online drug marketplace, in the UK, Australia and the USA. *Addiction*. 2014;109(5):774–783.
8. Bellis MA, Quigg Z, Hughes K, Ashton K, Ferris J, Winstock A. Harms from other people's drinking: an international survey of their occurrence, impacts on feeling safe and legislation relating to their control. *BMJ Open*. 2015;5(12).
9. Rothman KJ, Gallacher JE, Hatch EE. Rebuttal: When it comes to scientific inference, sometimes a cigar is just a cigar. *Int J Epidemiol*. 2013;42(4):1026-1028.
10. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. *Int J Epidemiol*. 2013;42(4):1018-1022.
11. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years—has the time come to move on? *Int J Epidemiol*. 2013;42(4):1016-1017.
12. Elwood JM. Commentary: On representativeness. *Int J Epidemiol*. 2013;42(4):1014-1015.
13. Ebrahim S, Davey Smith G. Commentary: Should we always deliberately be non-representative? *Int J Epidemiol*. 2013;42(4):1022-1026.
14. Pasek J. When will nonprobability surveys mirror probability surveys? Considering types of inference and weighting strategies as criteria for correspondence. *International Journal of Public Opinion Research*. 2016;28:269-291.
15. Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2016;179(2):319-376.
16. Kreuter F. Facing the Nonresponse Challenge. *The Annals of the American Academy of Political and Social Science*. 2013;645(1):23-35.
17. Zhao J, Stockwell TIM, Macdonald S. Non-response bias in alcohol and drug population surveys. *Drug Alc Rev*. 2009;28(6):648-657.
18. Wiebel WW. Identifying and gaining access to hidden populations. In: Lambert EY, ed. *The collection and interpretation of data from hidden populations (NIDA Research Monograph No. 98)*. Rockville, MD: National Institute on Drug Abuse; 1990:4-11.
19. Skalland B, Khare M. Geographic inaccuracy of cell phone samples and the effect on telephone survey bias, variance, and cost. *Journal of Survey Statistics and Methodology*. 2013;1(1):45-65.
20. McBride O, Morgan K, McGee H. Recruitment using mobile telephones in an Irish general population sexual health survey: challenges and practical solutions. *BMC Med Res Methodol*. 2012;12(1):45.

21. Livingston M, Dietze P, Ferris J, Pennay D, Hayes L, Lenton S. Surveying alcohol and other drug use through telephone sampling: A comparison of landline and mobile phone samples. *BMC Medical Research Methods*. 2013;13:41.
22. Barratt MJ, Ferris JA, Lenton S. Hidden populations, online purposive sampling, and external validity: Taking off the blindfold. *Field Methods*. 2015;27:3-21.
23. Groves RM. Three Eras of Survey Research. *Public Opin Q*. 2011;75(5):861-871.
24. Fricker RD. Sampling Methods for Web and E-mail Surveys. In: Fielding N, Lee RM, Blank G, eds. *The Sage Handbook of Online Research Methods*. London: Sage; 2008:195-216.
25. Barratt MJ, Lenton S. Beyond recruitment? Participatory online research with people who use drugs. *International Journal of Internet Research Ethics*. 2010;3:69-86.
26. Miller PG, S nderlund AL. Using the internet to research hidden populations of illicit drug users: A review. *Addiction*. 2010;105:1557-1567.
27. Tuten TL. Conducting online surveys. In: Gosling SD, Johnson JA, eds. *Advanced methods for conducting online behavioral research*. Washington, DC: American Psychological Association; 2010:179-192.
28. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: The future in epidemiology? *Am J Epidemiol*. 2010;172:1292-1298.
29. Kays KM, Keith TL, Broughal MT. Best practice in online survey research with sensitive topics. In: Sappleton N, ed. *Advancing Research Methods with New Technologies*. Hershey, PA: Information Science Reference; 2013:157-168.
30. Kruskal W, Mosteller F. Representative Sampling, I: Non-Scientific Literature. *International Statistical Review / Revue Internationale de Statistique*. 1979;47(1):13-24.
31. Kruskal W, Mosteller F. Representative Sampling, II: Scientific Literature, Excluding Statistics. *International Statistical Review / Revue Internationale de Statistique*. 1979;47(2):111-127.
32. Kruskal W, Mosteller F. Representative Sampling, III: The Current Statistical Literature. *International Statistical Review / Revue Internationale de Statistique*. 1979;47(3):245-265.
33. Kruskal W, Mosteller F. Representative Sampling, IV: The History of the Concept in Statistics, 1895-1939. *International Statistical Review / Revue Internationale de Statistique*. 1980;48(2):169-195.
34. Gobo G. Sampling, representativeness and generalizability. In: Seale C, Gobo G, Gubrium JF, Silverman D, eds. *Qualitative research practice*. London: Sage; 2007:405-426.
35. McCreesh N, Copas A, Seeley J, et al. Respondent Driven Sampling: Determinants of Recruitment and a Method to Improve Point Estimation. *PLoS ONE*. 2013;8(10):e78402.
36. Volz E, Heckathorn DD. Probability Based Estimation Theory for Respondent Driven Sampling. *Journal of Official Statistics*. 2008;24(1):79-97.
37. Global Drug Survey. Global Drug Survey Brochure. 2016; https://www.globaldrugsurvey.com/wp-content/uploads/2016/09/GDS_Brochure_2016.pdf.
38. Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. 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*. 2016;4(2):e47.
39. Barratt MJ, Ferris JA, Winstock AR. Safer scoring? Cryptomarkets, social supply and drug market violence. *International Journal of Drug Policy*. 2016;35:24-31.
40. Winstock A, Lynskey M, Borschmann R, Waldron J. Risk of emergency medical treatment following consumption of cannabis or synthetic cannabinoids in a large global sample. *J Psychopharmacol (Oxf)*. 2015;29(6):698-703.

41. Winstock AR, Barratt MJ. The 12-month prevalence and nature of adverse experiences resulting in emergency medical presentations associated with the use of synthetic cannabinoid products. *Human Psychopharmacology: Clinical and Experimental*. 2013;28(4):390-393.
42. Lawn W, Barratt M, Williams M, Horne A, Winstock A. The NBOMe hallucinogenic drug series: Patterns of use, characteristics of users and self-reported effects in a large international sample. *J Psychopharmacol (Oxf)*. 2014;28:780-788.
43. Winstock AR, Kaar S, Borschmann R. Dimethyltryptamine (DMT): Prevalence, user characteristics and abuse liability in a large global sample. *J Psychopharmacol (Oxf)*. 2014;28:49-54.
44. Kaar SJ, Ferris J, Waldron J, Devaney M, Ramsey J, Winstock AR. Up: The rise of nitrous oxide abuse. An international survey of contemporary nitrous oxide use. *J Psychopharmacol (Oxf)*. 2016;30:395-401.
45. Lawn W, Borschmann R, Cottrell A, Winstock A. Methoxetamine: prevalence of use in the USA and UK and associated urinary problems. *Journal of Substance Use*. 2016;21(2):115-120.
46. Winstock AR, Lawn W, Deluca P, Borschmann R. Methoxetamine: An early report on the motivations for use, effect profile and prevalence of use in a UK clubbing sample. *Drug Alc Rev*. 2016;35(2):212-217.
47. Winstock A, Bell J, Borschmann R. Friends, doctors, and tramadol: we might have a problem. *BMJ : British Medical Journal*. 2013;347.
48. Winstock AR, Borschmann R, Bell J. The non-medical use of tramadol in the UK: findings from a large community sample. *Int J Clin Pract*. 2014;68(9):1147-1151.
49. Barratt MJ, Potter GR, Wouters M, et al. Lessons from conducting trans-national Internet-mediated participatory research with hidden populations of cannabis cultivators. *International Journal of Drug Policy*. 2015;26:238-249.
50. Bauermeister JA, Pingel E, Zimmerman M, Couper M, Carballo-Diéguez A, Strecher VJ. Data quality in HIV/AIDS web-based surveys: Handling invalid and suspicious data. *Field Methods*. 2012;24:272-291.
51. Aubusson K. Drug users swap stories, share warnings online in search for safety. *Sydney Morning Herald*. 16 November, 2016.
52. Aubusson K, Spooner R. Global Drug Survey set to peel back curtain on professionals microdosing on LSD and magic mushrooms. *Sydney Morning Herald*. 2 December, 2016.
53. Winstock A, Barratt M. Dark web markets: The good, the bad and the ugly. *High Times*. 12 December, 2016.
54. Winstock A, Barratt M. Zero tolerance for drugs doesn't work. Let's embrace zero harm. *Sydney Morning Herald*. 29 December, 2016.
55. Bowen A, Daniel C, Williams M, Baird G. Identifying multiple submissions in Internet research: Preserving data integrity. *AIDS and Behavior*. 2008;12:964-973.
56. von Lüssi M. Etwa jeder 10. Konsument hat ein Drogenproblem [About every 10th consumer has a drug problem]. *20 Minuten*. 14 June, 2016.
57. Eysenbach G, Wyatt J. Using the Internet for Surveys and Health Research. *J Med Internet Res*. 2002;4(2):e13.
58. Chalmers J, Lancaster K, Hughes C. The stigmatisation of 'ice' and under-reporting of meth/amphetamine use in general population surveys: A case study from Australia. *International Journal of Drug Policy*. 2016;36:15-24.
59. Delaney-Black V, Chiodo LM, Hannigan JH, et al. Just say "I don't": lack of concordance between teen report and biological measures of drug use. *Pediatrics*. 2010;126(5):887-893.
60. Australian Institute of Health and Welfare. *2013 National Drug Strategy Household Survey. Detailed report*. Canberra: Author;2014.

61. Substance Abuse and Mental Health Services Administration (SAMHSA). *National Survey on Drug Use and Health, 2013*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research;2014.
62. Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*. Rockville, MD: Author;2014. H-48.
63. Gmel G, Kuendig H, Notari L, Gmel C. *Suchtmonitoring Schweiz - Konsum von Alkohol, Tabak und illegalen Drogen in der Schweiz im Jahr 2015 [Addiction monitoring Switzerland - Consumption of alcohol, tobacco and illicit drugs in Switzerland in 2015]*. Lausanne, Schweiz: Sucht Schweiz;2016.
64. *Stata Statistical Software: Release 14* [computer program]. College Station, TX: StataCorp LP; 2015.
65. Mounteney J, Griffiths P, Sedefov R, Noor A, Vicente J, Simon R. The drug situation in Europe: an overview of data available on illicit drugs and new psychoactive substances from European monitoring in 2015. *Addiction*. 2016;11:34–48.
66. Barratt MJ. The efficacy of interviewing young drug users through online chat. *Drug Alc Rev*. 2012;31:566–572.
67. Barratt MJ, Maddox A. Active engagement with stigmatised communities through digital ethnography. *Qualitative Research*. 2016;16(6):701–719.

Tables

Table 1. Global Drug Survey: characteristics of the survey over 6 years

	GDS2012	GDS2013	GDS2014	GDS2015	GDS2016	GDS2017
Date range of data collection	22 Nov 2011 – 22 Dec 2011	15 Nov 2012 – 2 Jan 2013	11 Nov 2013 – 29 Dec 2013	9 Nov 2014 – 3 Jan 2015	8 Nov 2015 – 10 Feb 2016	15 Nov 2016 – 18 Jan 2017
Total N sample after cleaning	15,095	21,575	74,864	97,855	96,900	119,075 ^a
Tagline	NA	NA	Informing change	Informing change	Separating fact from fiction	Change starts with experience
List of translated languages	English	English	English, German, Spanish, Dutch, French, Hungarian, Portuguese, Danish, Slovak	English, German, Spanish (+SA), Dutch, French, Hungarian, Portuguese (+SA), Danish, Italian, Greek, Flemish, Polish	English, German, Spanish (+SA), Dutch, French, Hungarian, Portuguese (+SA), Danish, Italian, Greek, Flemish, Polish	English, German, Spanish (+SA), Dutch, French, Hungarian, Portuguese (+SA), Danish, Italian, Swedish
List of countries with active recruitment	UK	Australia, UK	Australia, Belgium, Denmark, France, Germany, Hungary, Netherlands, New Zealand, Switzerland, UK, US	Australia, Belgium, Brazil, Colombia, Denmark, France, Germany, Greece, Hungary, Netherlands, New Zealand, Poland, Portugal, Spain, Switzerland, UK, US	Australia, Belgium, Brazil, Colombia, France, Germany, Hungary, Ireland, Italy, Netherlands, New Zealand, Norway, Portugal, Hungary, Spain, Switzerland, UK, US	Australia, Austria, Belgium, Brazil, Canada, Colombia, Denmark, Finland, Germany, Hungary, Iceland, Ireland, Italy, Mexico, Netherlands, New Zealand, Portugal, Spain, Switzerland, UK, US
List of media partners	Mixmag, The Guardian	Mixmag, The Guardian, Fairfax Media, GT	Mixmag, The Guardian, Fairfax Media, GT, The Huffington Post, Zeit Online, 20 Minuten, Liberation, stuff.com.nz, Hot Press, The Herald,	Mixmag, The Guardian, Fairfax Media NZ and AU, The Huffington Post, Zeit Online, 20 Minuten, Liberation, stuff.co.nz, Hot Press, The Herald, Studio	Mixmag, The Guardian (UK, and AU), Fairfax Media -NZ, Zeit Online, 20 Minuten, Liberation, stuff.co.nz, Hot Press, The Herald, Vice, Thump Vice, der Morgen co.nz	Mixmag, The Guardian (USA), Fairfax Media -NZ+AU, Zeit Online, 20 Minuten, Liberation, stuff.co.nz, Hot Press, Studio Brussels, Vice UK and Canada, Thump Vice,

			Blitz, neon, Publico, DMorgen, BNN, Studio Brussel	Brussels, Vice, Pink News, The journal.ie, attitude.co.uk, Folha de S.Paulo, Super, BNN, Interessante	Motherboard, i-D, Dazed, Super Interessante, Daag Blatt, La Repubblica (It)	La Repubblica (It), The Independent UK, High Times USA, Der Standard, El Universal Mexico, Hu444
Other recruitment methods	Facebook, Twitter	Facebook, Twitter, Reddit, Web forums	Facebook, Twitter, Reddit, Web forums	Facebook, Twitter, Reddit, Web forums	Facebook, Twitter, Reddit, Web forums	Facebook, Twitter, Reddit, Web forums

Note. SA = South American. ^a This total excludes respondents aged over 80 years and respondents who did not nominate either male or female sex.

Table 2. Demographics and alcohol and cannabis use patterns in the Australian, US and Swiss GDS and national survey samples

	GDS N (†%)	National Survey N (†%)
Australia	N=5,789	N=23,855
<i>Sex</i>		
Male	3,479 (61.0)	10,624 (44.5)
Female	2,223 (39.0)	13,231 (55.5)
Missing	87	0
<i>Age</i>		
mean (SD)	37.1 (13.1)	47.5 (16.7)
Missing	75	1,157
<i>Cannabis use</i>		
Cannabis ever	4,063 (70.2)	8,101 (34.4)
Missing	0	330
Cannabis past 12 months*	2,037 (50.1)	2,174 (26.8)
Missing	0	0
Cannabis past month**	1,379 (67.7)	1,143 (52.7)
Missing	0	4
<i>Alcohol use</i>		
Alcohol ever	5,711 (98.7)	20,881 (87.6)
Missing	0	11
Alcohol past 12 months*	5,403 (94.6)	19,161 (91.8)
Missing	0	0
US	N=6,419	N=43,465
<i>Sex</i>		
Male	3,846 (61.2)	20,302 (46.7)
Female	2,439 (38.8)	23,163 (53.3)
Missing	134	0
<i>Age</i>		
16-20 yrs	1,194 (18.9)	12,750 (29.3)
21-25 yrs	1,251 (19.8)	11,433 (26.3)
26-35 yrs	1,287 (20.4)	5,446 (12.5)
36-49 yrs	1,263 (20.0)	7,511 (17.2)
50+ yrs	1,323 (20.9)	6,325 (14.5)
Missing	101	0
<i>Cannabis use</i>		
Cannabis ever	5,774 (90.0)	21,075 (48.5)
Missing	0	0
Cannabis past 12 months*	4,519 (78.3)	9,498 (45.1)
Missing	0	0
Cannabis past month**	3,879 (85.8)	5,664 (59.6)
Missing	0	0
<i>Alcohol use</i>		
Alcohol ever	6,292 (98.0)	35,811 (82.4)
Missing	0	0
Alcohol past 12 months*	5,621 (89.3)	30,758 (85.9)
Missing	0	0
Switzerland	N=4,971	N=12,008
<i>Sex</i>		
Male	3,381 (69.9)	5,542 (46.2)
Female	1,456 (30.1)	6,466 (53.8)

Missing	134	0
<i>Age</i>		
15-19 yrs	775 (16.0)	2,306 (19.2)
20-29 yrs	2,108 (43.5)	2,438 (20.3)
30-39 yrs	1,067 (22.0)	1,043 (8.7)
40-49 yrs	533 (11.1)	1,625 (13.5)
50+ yrs	358 (7.4)	4,596 (38.3)
Missing	130	0
<i>Cannabis use</i>		
Cannabis ever	3,236 (65.1)	3,465 (28.9)
Missing	0	40
Cannabis past 12 months*	2,056 (63.5)	1,057 (30.5)
Missing	0	2
Cannabis past month**	1,599 (77.8)	481 (45.5)
Missing	0	0
<i>Alcohol use</i>		
Alcohol ever	4,913 (98.8)	11,136 (92.7)
Missing	0	0
Alcohol past 12 months*	4,679 (95.2)	10,353 (93.5)
Missing	0	58

* past year only include respondents who indicated they had used cannabis/alcohol in their lifetime.

**past month only includes respondents who indicated they had used cannabis in the past 12 months.

† Percentages are of valid cases (excluding missings) and may sum to greater than 100 due to rounding.

Table 3. Age categories by sex in the Australian, US and Swiss GDS and national surveys

	GDS		National Survey	
	N (†%)		N (†%)	
	Male	Female	Male	Female
Australia	N=3,479	N=2,223	N=10,357	N=12,976
<i>Age</i>				
16-20 yrs	249 (7.2)	135 (6.1)	622 (5.9)	686 (5.2)
21-25 yrs	465 (13.4)	380 (17.1)	616 (5.8)	763 (5.8)
26-34 yrs	1,029 (29.6)	716 (32.2)	1,494 (14.1)	2,183 (16.5)
35-49 yrs	1,011 (29.1)	590 (26.5)	2,374 (22.4)	3,228 (24.4)
50+ yrs	725 (20.8)	402 (18.1)	5,251 (49.4)	6,116 (42.2)
US	N=3,837	N=2,430	N=20,302	N=23,163
<i>Age</i>				
16-20 yrs	921 (24.0)	267 (11.0)	6,380 (31.4)	6,370 (27.5)
21-25 yrs	797 (20.8)	445 (18.3)	5,233 (25.8)	6,200 (26.8)
26-34 yrs	736 (19.2)	547 (22.5)	2,448 (12.1)	2,998 (12.9)
35-49 yrs	679 (17.7)	575 (23.7)	3,426 (16.9)	4,085 (17.6)
50+ yrs	704 (18.5)	596 (24.5)	2,815 (13.9)	3,510 (15.2)
Switzerland	N=3,361	N=1,451	N=5,542	N=6,466
<i>Age</i>				
15-19 yrs	537 (15.9)	237 (16.3)	1,177 (21.2)	1,129 (17.5)
20-29 yrs	1,374 (40.6)	728 (50.1)	1,186 (21.4)	1,252 (19.4)
30-39 yrs	767 (22.7)	293 (20.2)	440 (7.9)	603 (9.3)
40-49 yrs	408 (12.1)	119 (8.2)	708 (12.8)	917 (14.2)
50+ yrs	275 (8.2)	74 (5.1)	2,031 (36.7)	2,565 (39.7)

† Percentages are of valid cases (excluding missings)

Figure legends

Figure 1. Cannabis use in Australia – GDS and NDSHS

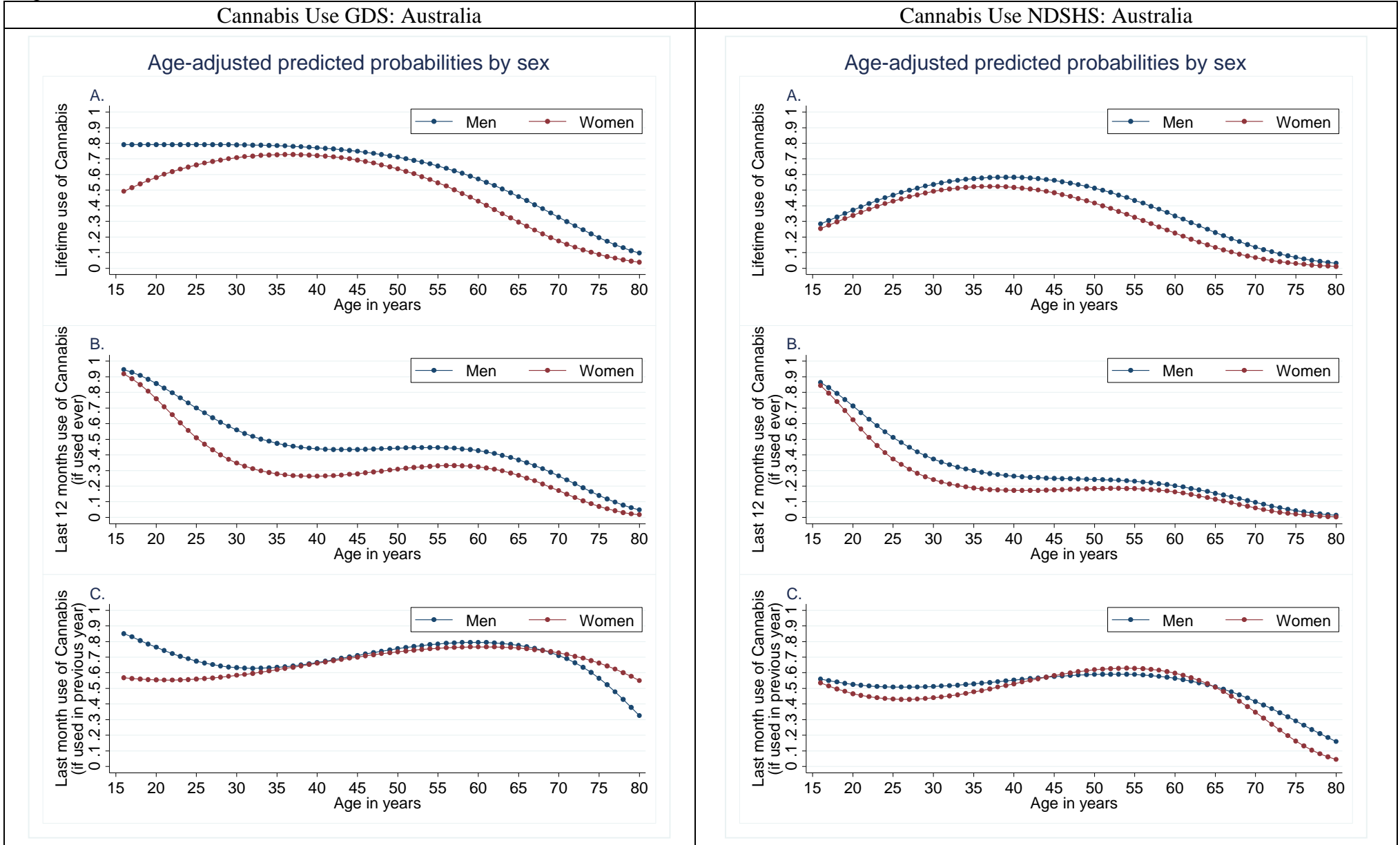


Figure 2. Cannabis use US – GDS and NSDUH

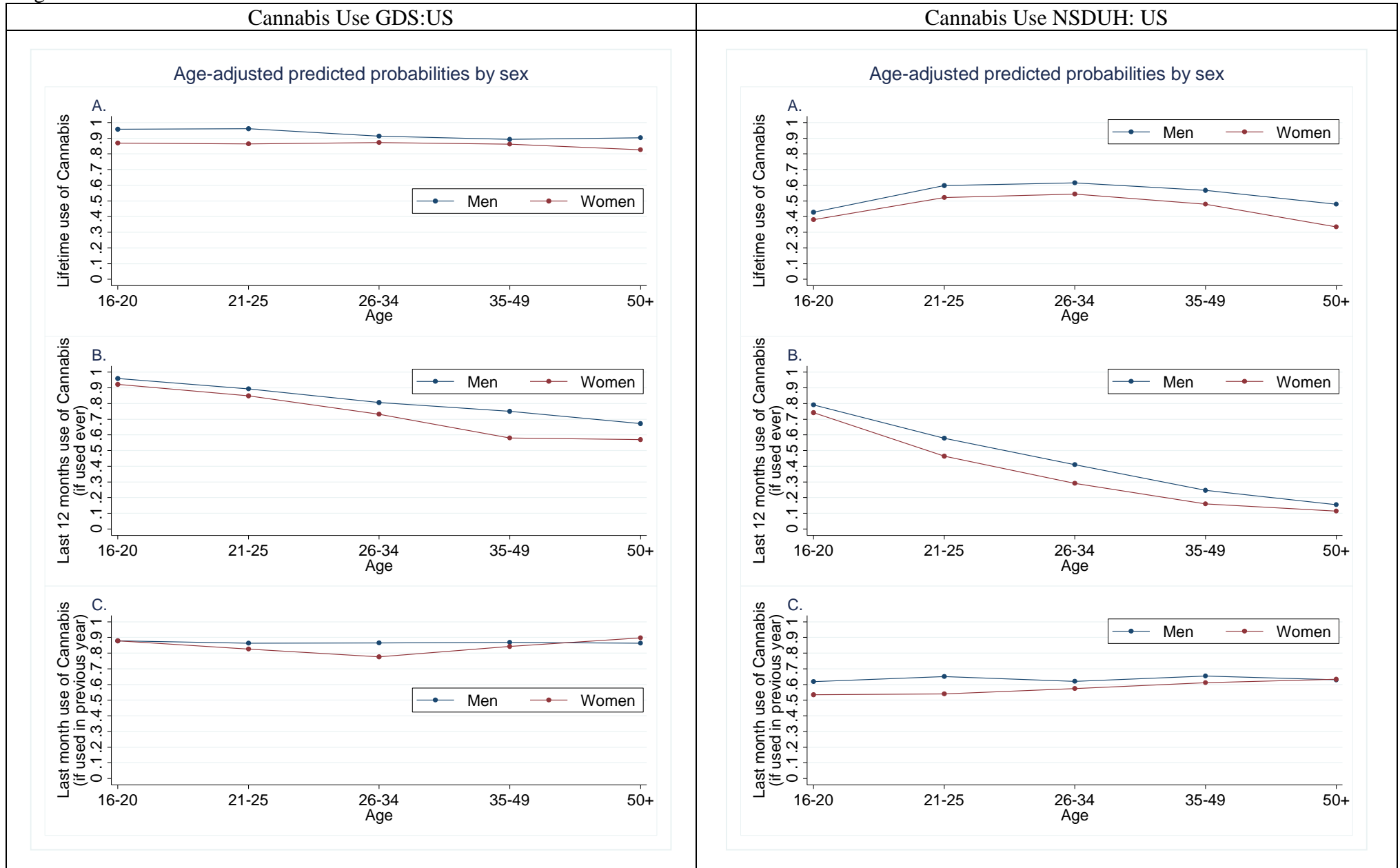
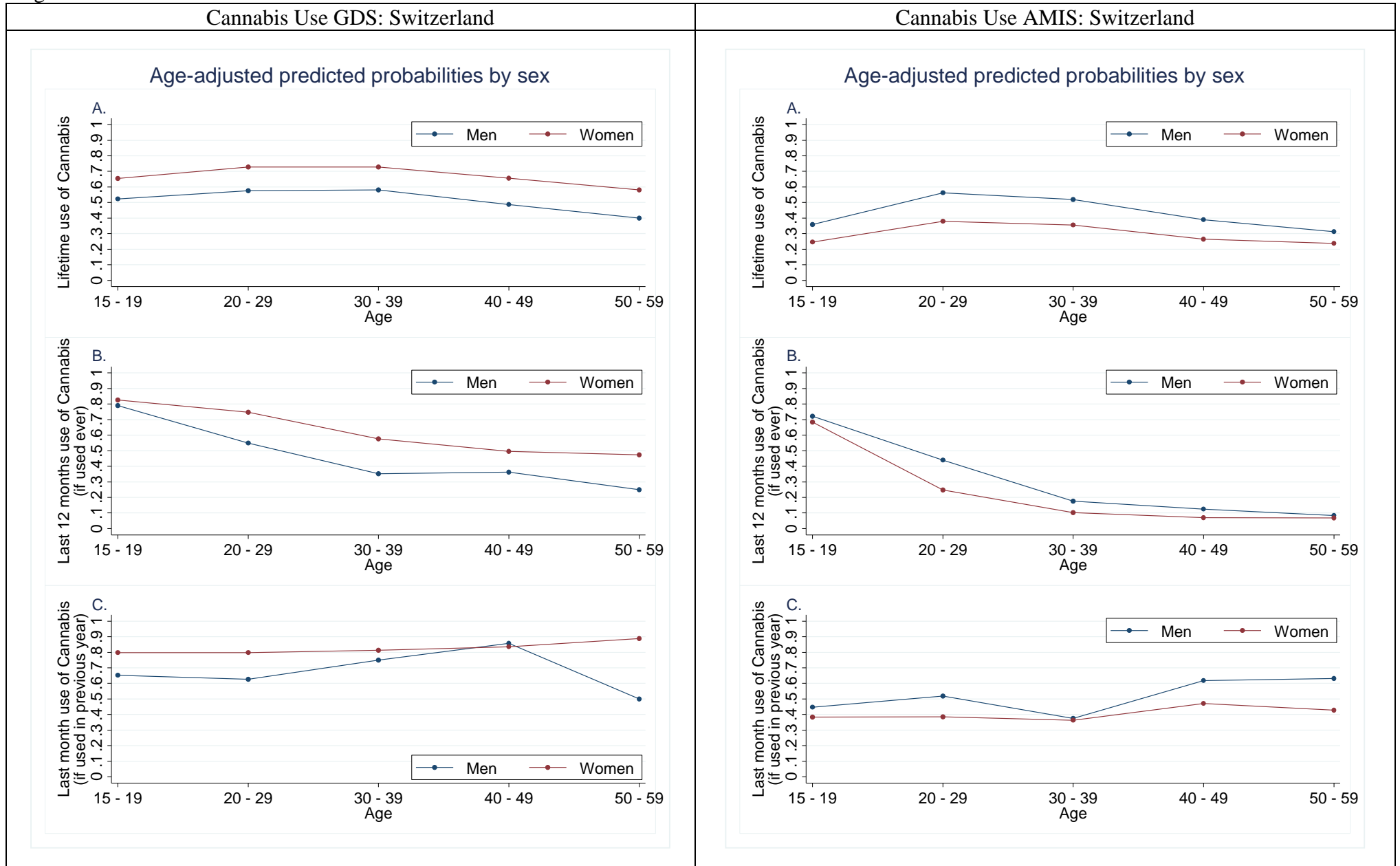


Figure 3. Cannabis use in Switzerland – GDS vs AMIS



Box 1 – GDS2014 recruitment call

#GDS2014 'informing change'

Drugs are a funny thing. Lots of people take them, lots of people talk about them – often not knowing much at all about what they are talking about. The media sometimes does a fantastic reporting on drug issues, other times it's just not useful or accurate. Sometimes it's just plain rubbish. Researchers do some interesting stuff but most of it's focused on the harm that drugs cause – not the pleasure they can bring to people's lives. Global Drug Survey hopes to change that. Collaborating with over 20 media partners and host of researchers and harm reduction networks across the world, translated into 10 languages and running with hubs in 17 countries we are about to undertake the biggest survey of everyday drug use (including alcohol, tobacco and of course prescription medication – they are drugs too!) that the world, has ever seen. From what would be the perfect cannabis and life after the Silk Road, to what happened when you're caught with drugs or being drug-tested in work place; from nitrous oxide balloons and ecstasy pill testing to value for money and finding out what people do to keep themselves safe while high – this is going to be huge and relevant for people who use drugs, who drink, drop, get high, get low, gets scripts, grow their own or are just curious about the world around them, With a focus on drug policy and the rising tide of new drugs and the internet – if you are interested in drugs help us get the best information out to everyone. Our results will be published exclusively with our global media partners in March and April 2014. Take part in Global Drug Survey 2014 at [\[link\]](#). Everything is anonymous and confidential.

Supplementary data

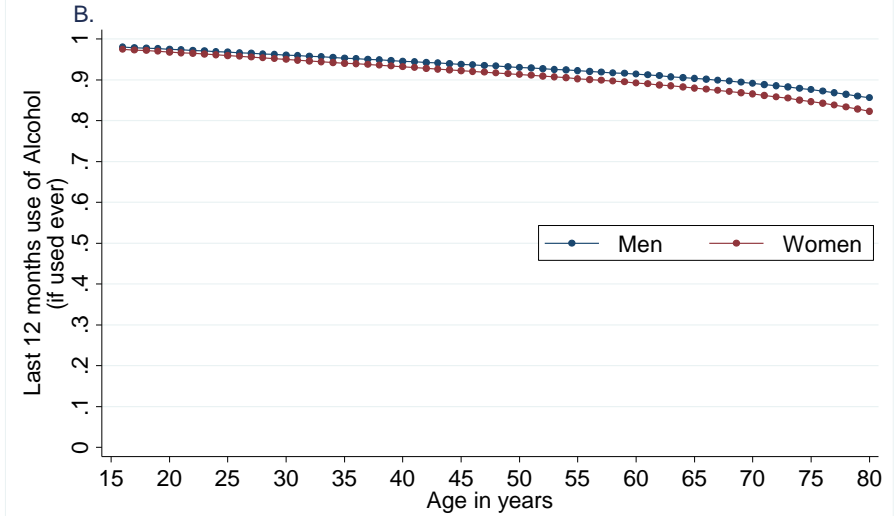
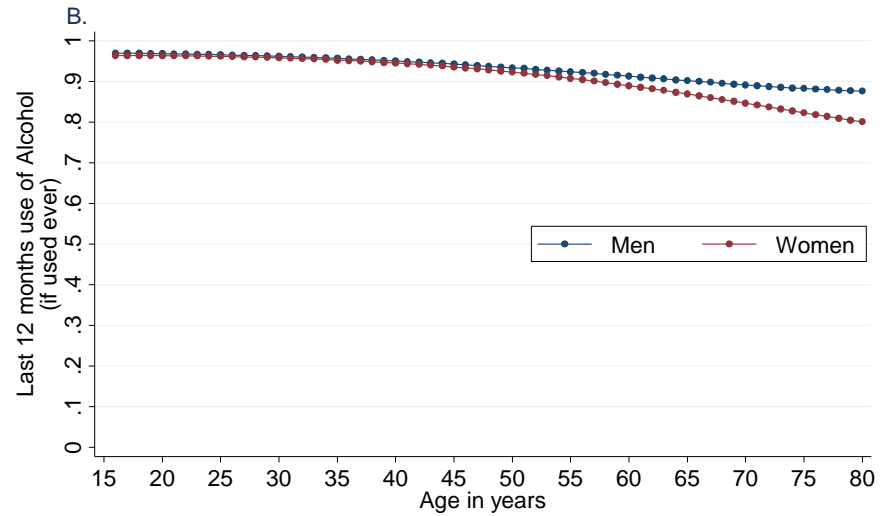
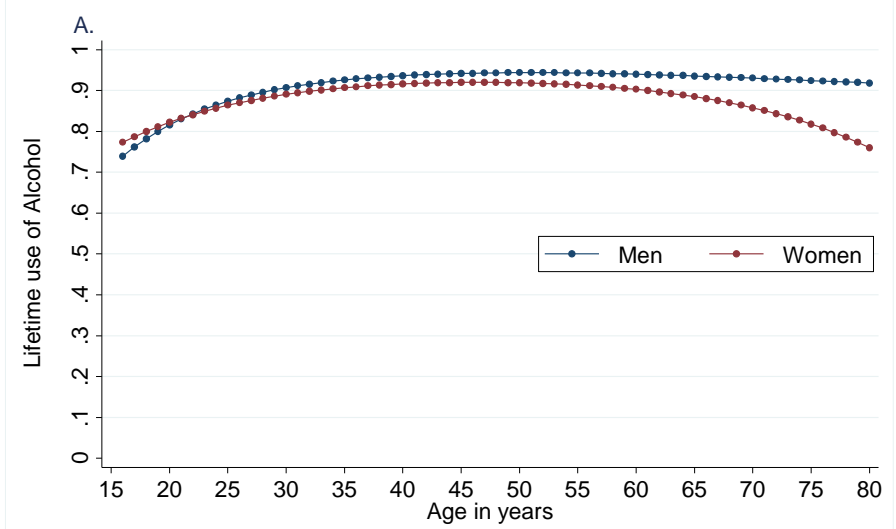
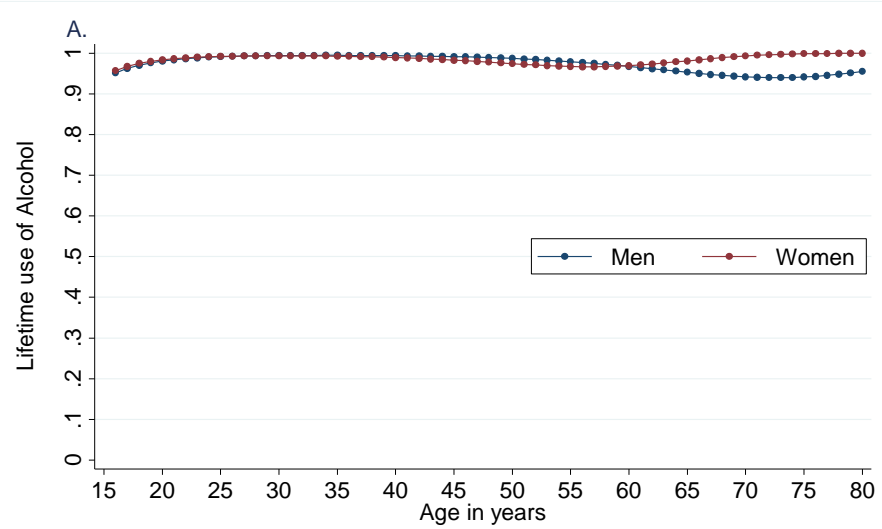
If supplementary data is in separate files, list and briefly describe the files here - Appendix 1. Alcohol use in Australia – GDS vs NDSHS

Alcohol Use GDS: Australia

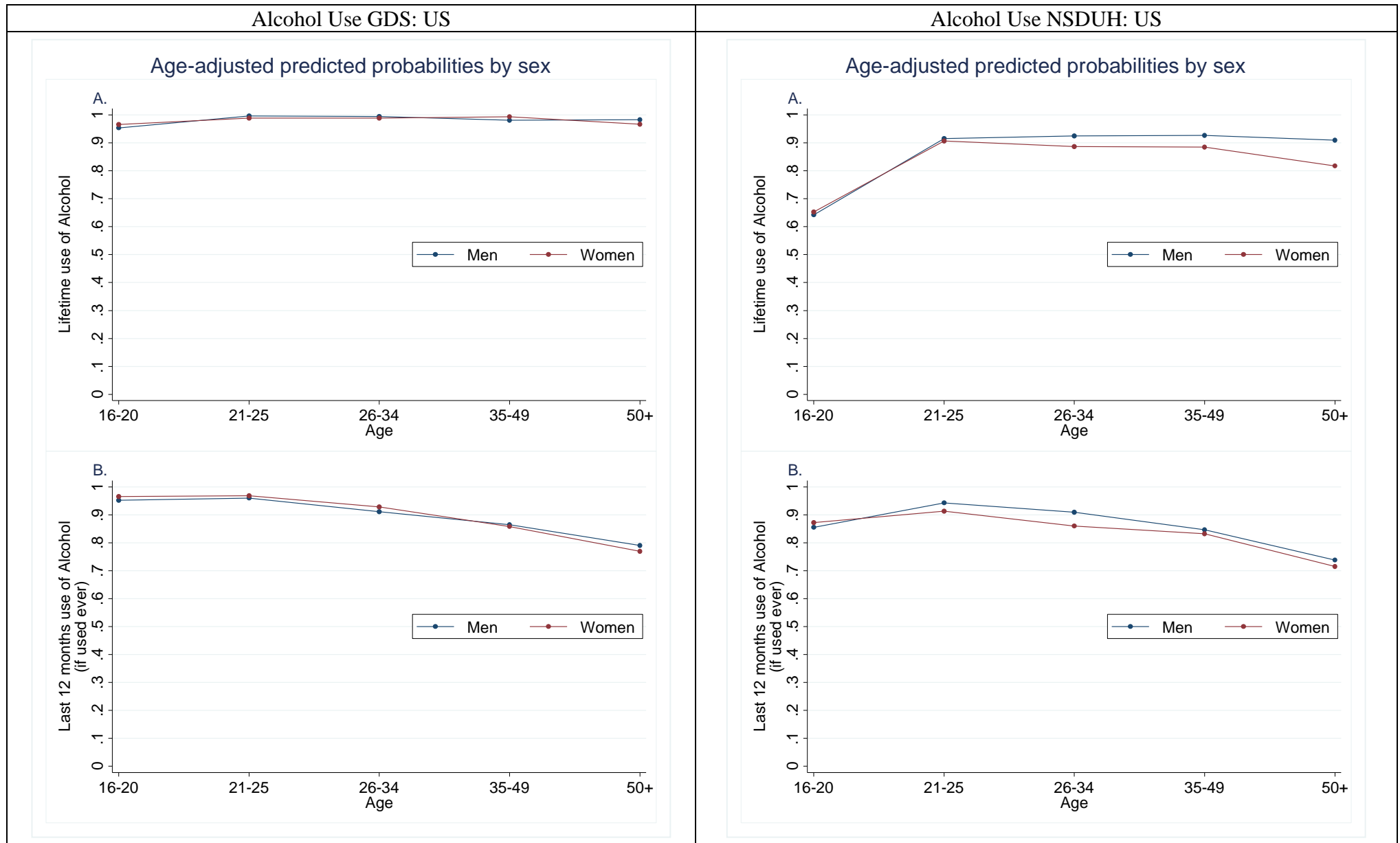
Alcohol Use NDSHS: Australia

Age-adjusted predicted probabilities by sex

Age-adjusted predicted probabilities by sex



Appendix 2. Alcohol use in US – GDS vs NSDUH



Appendix 3. Alcohol use in Switzerland – GDS vs AMIS

Alcohol Use GDS: Switzerland

Alcohol Use AMIS: Switzerland

