

# Moving on From Representativeness: Testing the Utility of the Global Drug Survey

Monica J Barratt<sup>1,2,3</sup>, Jason A Ferris<sup>4</sup>, Renee Zahnow<sup>4</sup>,  
Joseph J Palamar<sup>5,6</sup>, Larissa J Maier<sup>7</sup> and Adam R Winstock<sup>8,9</sup>

<sup>1</sup>Drug Policy Modelling Program, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, Australia. <sup>2</sup>National Drug Research Institute, Faculty of Health Sciences, Curtin University, Perth, WA, Australia. <sup>3</sup>Behaviours and Health Risks Program, Burnet Institute, Melbourne, VIC, Australia.

<sup>4</sup>Institute for Social Science Research, The University of Queensland, St Lucia, QLD, Australia.

<sup>5</sup>Department of Population Health, NYU Langone Medical Center, New York, NY, USA. <sup>6</sup>Center for Drug Use and HIV Research, NYU Rory Meyers College of Nursing, New York, NY, USA. <sup>7</sup>University of Zurich, Zurich, Switzerland. <sup>8</sup>Institute of Epidemiology & Health Care, Faculty of Population Health Sciences, University College London. <sup>9</sup>Global Drug Survey Ltd, London, UK.

Substance Abuse: Research and Treatment  
Volume 11: 1–17  
© The Author(s) 2017  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1178221817716391



**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, such as 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 article, 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 article 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 United States, 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 stigmatised 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

**RECEIVED:** January 20, 2017. **ACCEPTED:** May 22, 2017.

**PEER REVIEW:** Six peer reviewers contributed to the peer review report. Reviewers' reports totalled 2920 words, excluding any confidential comments to the academic editor

**TYPE:** Original Research

**FUNDING:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: M.J.B. and J.A.F. were supported by fellowships from the Australian National Health and Medical Research Council (APP1070140, APP1089395). J.J.P. was supported by the US National Institutes of Health (NIH) (K01 DA-038800, J.J.P.). 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. This work is also supported by the contribution of the Victorian Operational Infrastructure Support Program received by the Burnet Institute.

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: A.R.W. is the founder and managing director of Global Drug Survey.

**CORRESPONDING AUTHOR:** Monica J Barratt, Drug Policy Modelling Program, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW 2052, Australia. Email: m.barratt@unsw.edu.au

## 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<sup>1(p13)</sup>)*

Research surveys that do not adopt probability-based sampling frames have generally been considered inherently inferior to surveys that adopt probability sampling frames.<sup>1–3</sup> 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.<sup>4–8</sup> 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 article, 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<sup>2,9–13</sup> 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<sup>14</sup> or in-depth profiling of sub-populations, the use of probability-based sampling frameworks is often inefficient, may be unnecessary or even better avoided.<sup>2,10</sup> Furthermore, the 'gold standard' household survey is more like 'tarnished gold'<sup>15</sup> when response rates and volunteer bias still affect what data are obtained.<sup>16,17</sup> In this article, 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 United States, 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 stigmatised behaviours through representative sampling frames.<sup>18</sup> Probability sampling methods are limited in many regards. Non-response 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,<sup>19</sup> mobile phone users are less likely to expect or respond to research requests,<sup>20</sup> and cannabis and tobacco use are more commonly reported through mobile phone sampling frames compared with landline samples.<sup>21</sup> People are generally less likely to report illicit drug use on the phone when compared with anonymous web surveys. In addition, probability methods are particularly expensive when researchers are targeting rare practices for in-depth analysis.<sup>22</sup> The changing context of survey research<sup>23</sup> and the growth of internet access have prompted greater use of purposive sampling of otherwise hard-to-reach populations via web surveys and internet recruitment.<sup>24,25</sup> Many studies of drug-using populations now use internet recruitment and/or survey methods with purposive sampling.<sup>25,26</sup> Some well-documented advantages of conducting surveys of hidden populations online include the following: 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 are automated, and flexibility and convenience are enhanced for both respondents and researchers.<sup>27–29</sup> 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’<sup>22</sup> 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.<sup>2,11</sup> But what is representativeness? In a series of papers published in 1979 to 1980,<sup>30–33</sup> Kruskal and Mosteller provide an account of the various meanings of the term ‘representative

sampling’. They describe six 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’.<sup>32(p244)</sup> When examining the statistical literature specifically, they found the above meanings plus three additional categories, including (7) ‘representative sampling as a specific sampling method, (8) representative sampling as permitting good estimation, (9) representative sampling as good enough for a particular purpose’.<sup>32(p244)</sup> 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 sex.<sup>34</sup> 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 three 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?<sup>14</sup> Keiding and Louis<sup>15</sup> 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 sampling. 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.<sup>35,36</sup> 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 is 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 Article

Our article 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 United States, 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. Global Drug Survey partners with global media organisations who help promote the survey and share the findings with their readers usually four to six months after closure of the survey each year. Global Drug Survey uses a cross-sectional design. Data have been collected annually for the past six years (see Table 1). Prior to the launch of GDS2012, the previous annual surveys led by Adam Winstock were targeted only at UK audience, supported by the UK clubbing magazine *Mixmag* and the UK news organisation *The Guardian*. (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.) These *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 web site ‘Don’t stay in’. This survey was relaunched after a five-year hiatus in 2009 and formally became known as the GDS 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).

### *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. Although 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. (Including to populations who inject drugs through partnership with the International Network of People Who Use Drugs (INPUD) in 2015 and 2016.) 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, web site, and harm reduction tools. Our main goal in conducting the GDS is to make drug use safer, regardless of the legal status of the drug.<sup>37</sup> 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 (eg, 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, Drinks Meter, and Drugs Meter (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.

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. First, 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. Second, project ideas must be written to match the predispositions of the funding bodies. Third, 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. Global Drug Survey is also able to do this because

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

	GDS2012	GDS2013	GDS2014	GDS2015	GDS2016	GDS2017
Date range of data collection	November 22, 2011-December 22, 2011	November 15, 2012-January 2, 2013	November 11, 2013-December 29, 2013	November 9, 2014-January 3, 2015	November 8, 2015-February 10, 2016	November 15, 2016-January 18, 2017
Total no. of sample after cleaning	15 095	21 575	74 864	97 855	96 900	119 075 <sup>a</sup>
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, Blitz, neon, Publico, De Morgen, BNN, Studio Brussel	Mixmag, The Guardian, Fairfax Media NZ and AU, The Huffington Post, Zeit Online, 20 Minuten, Liberation, stuff.com.nz, Hot Press, The Herald, Vice, Thump Vice, De Morgen co. journal.ie, attitude.co.uk, Folha de S. Paulo, Super, BNN, Interessante	Mixmag, The Guardian (UK and AU), Fairfax Media – NZ, Zeit Online, 20 Minuten, Liberation, stuff.com.nz, Hot Press, The Herald, Vice, Thump Vice, De Morgen co. nz, Motherboard, i-D, Dazed, Superinteressante, Daag Bladt, La Repubblica (It)	Mixmag, The Guardian (USA), Fairfax Media – NZ + AU, Zeit Online, 20 Minuten, Liberation, stuff.com.nz, Hot Press, Studio Brussel, Vice UK and Canada, Thump Vice, 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

Abbreviation: SA, South American.

<sup>a</sup>This total excludes respondents aged more than 80 years and respondents who did not nominate either male or female sex.

**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.

of the relatively low cost of running web surveys. Although 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. Although GDS has developed a number of free digital harm reduction tools which are available on its web site, 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.<sup>38</sup> Global Drug Survey 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. Global Drug Survey 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<sup>27</sup> and has been at the forefront of surveying users of cryptomarkets in the years since.<sup>39</sup> 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,<sup>4</sup> synthetic cannabinoids,<sup>6,40,41</sup> the NBOMe series,<sup>42</sup> dimethyltryptamine,<sup>43</sup> nitrous oxide,<sup>44</sup> and methoxetamine.<sup>45,46</sup> Modules on alcohol and prescription medication use have resulted in papers on alcohol's harm to others<sup>8</sup> and profiles of the prescription medication tramadol.<sup>47,48</sup>

An important aspect of the survey design is participant anonymity. The GDS web site is encrypted, and no internet protocol (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.<sup>49</sup> The IP addresses can be used by multiple people validly (eg, share houses, university dorms).<sup>50</sup> The GDS offers an option at its completion for participants to provide contacts 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 six years.

Before launch, the extended GDS network pilots the survey and provides 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*

Global Drug Survey 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 more than 20 countries. The survey typically launches in the second week of November and runs for a period of about six to eight 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<sup>51,52</sup> but may also include GDS-written opinion pieces.<sup>53,54</sup> Stand-alone media content will include advertising for participants and calls for participation within the article text. Although 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 over-represented. The GDS network also includes non-governmental 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 six to eight 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.<sup>55</sup> 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 the following: age, sex, ethnicity, educational attainment, employment type, income, country, height, weight, clubbing activity, exercise, and standard of living. Drug screen variables include the following: 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 (eg, lotteries, prizes, payments), and because the survey typically takes 15 minutes to one hour to complete (depending on 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 the following: 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 data set is finalised, 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 lysergic acid diethylamide (LSD) through an online drug market but does not report ever use of LSD in the earlier drug screen, the later data are 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 five to six months after data collection ends.<sup>56</sup> 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 not,<sup>17</sup> and volunteer bias, where people are more likely to respond if they are interested in the topic,<sup>57</sup> 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, however, that weighting probability samples cannot fully account for sampling bias.<sup>15,17,34</sup> For example, the data from young men who

complete a household survey will be weighted more heavily to make up for the lower participation rates of their cohort, yet the young men who respond are likely to have different characteristics to the young men who do not respond. That is, there are unmeasured confounders.<sup>15</sup> 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 survey.<sup>58</sup> There may also be underreporting of drug use using face-to-face or telephone survey modes,<sup>59</sup> 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, although 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 populations<sup>17</sup> that are the most likely to have lived experiences of illicit drug use. Another strength comparing GDS with 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 wordings.

## Part 2 – Testing the Utility of GDS

Given the limitations and strengths of GDS compared with household surveys using probability sampling methods, in this article, we compare the similarities between predicted probabilities of recent (past 12-month) and current (past 30-day) use of cannabis among 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 time frame. In this article, we conduct this analysis with cannabis users from Australia, the United States, and Switzerland and report results from a similar analysis on alcohol use in the supplementary appendix.

## Method

*Global Drug Survey.* Global Drug Survey 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 article, we employ data from the GDS2014 (see Table 1 column three for more information). In this study, the sample was restricted to respondents living in Australia (N=5789), the United States (N=6419), and Switzerland (N=4971). All respondents reported being 16 years or older. Of the total 16 828 respondents who reported their age from the selected countries, 10 708 (63.6%) were men and 6120 (36.4%) were women. The mean reported age was 34.2 (SD=13.9) years. 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 whether they had ever used alcohol and whether they had used alcohol in the past 12 months. All GDS participants were also asked whether they had ever used or used in the past year: cannabis/marijuana – herbal-high potency (hydro), cannabis/marijuana – herbal-normal weed (eg, bush weed/pressed), cannabis (resin/hash), and Cannabis oil. Global Drug Survey 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 (Australia).* The National Drug Strategy Household Survey (NDSHS) is a national survey conducted by the Australian Institute of Health and Welfare every three years.<sup>60</sup> 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 (United States).* 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-institutionalised individuals in the 50 US states and in the District of Columbia.<sup>61</sup> The 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-level and individual-level non-response and to ensure that estimates are consistent with estimates provided by the US Census Bureau; however, we did not use sampling weights for the following analyses. The (unweighted) interview response rate in 2013 was 76.4%.<sup>62</sup>

Participants were asked their sex (ie, male, female) and age. The NSDUH provided pre-coded categorical variables indicating participant age, which we recoded into the following: age 16 to 20, 21 to 25, 26 to 34, 35 to 49, and 50 years 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 whether 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 regarding 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 (Switzerland).* The Addiction Monitoring in Switzerland (AMIS) is a cross-sectional national survey that has been conducted by the Swiss Federal of Public Health every year from 2011 to 2016.<sup>63</sup> 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 12008 Swiss residents aged 15 years or older. Addiction Monitoring in Switzerland 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. However, 8.5% of the calls used mobile phone numbers that were generated by random digit dialing. The survey was conducted via a self-completion, drop and collect method. In 2013, people aged 15 to 29 years were intentionally over-represented. 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 to 30 minutes long, whereas 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-level 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 German, French, or Italian. To assess lifetime and recent alcohol use, all participants in the AMIS 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 (ie, male, female) and age. The AMIS provided pre-coded categorical variables indicating participant age: age 15 to 19, 20 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, 65 to 74, and 75 years 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 data set comprised respondents who recorded responses to questions about cannabis use (yes/no), sex (male/female), and age. (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.) 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 were used for comparison with GDS respondents. Although 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 actually complete the household survey with those who complete the non-probability survey (GDS). Predicted probability analyses were conducted using Stata 14.<sup>64</sup> 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 to 3.

## Results

*Australia.* As shown in Tables 2 and 3, the analytic sample from the NDSHS (N = 23 855) comprised 10 624 (44.5%) men and 13 231 (55.5%) women with a mean age of 47.5 (SD: 16.7) years. With respect to cannabis use, 8101 (34.4%) indicated that they had used cannabis in their lifetime, 2174 (26.8% of lifetime users) reported using cannabis in the past 12 months, and 1143 (52.7% of past-year users) reported use in the past month.

A total of 5789 Australians participated in GDS2014. This sample comprised 3479 (61.0%) men and 2223 (39.0%) women with a mean age of 37.13 (SD: 13.1) years. With respect to cannabis use, 4063 (70.2%) reported using cannabis in their lifetime, 2037 (50.1% of lifetime users) reported using cannabis in the past 12 months, and 1379 (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 data sets. For both the GDS and the NDSHS data, regardless of age, men are typically more likely to report using cannabis than women: this can be observed for lifetime use and past 12-month use in both data sets. However, the bottom figures show some crossover between the sexes for reported cannabis use in the past month with women at particular ages more likely to report cannabis use compared with men. 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 and 64 years of age.

For respondents 40 years and above, overall, men and women 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 with 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 women was similar, especially for those between the ages of 25 and 30 years and 55 and 60 years of age. For both data sets, as the age of female respondents increased, so did the probability of using cannabis within the past month. For women more than 55 to 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 men in the GDS sample (those 30 years of age or less), the patterns were more similar.

Although the probability of ever using cannabis is higher in the GDS sample, once the data sets 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 is similar across the GDS and the NDSHS samples.

*United States.* As presented in Tables 2 and 3, the analytic sample from the NSDUH (N = 43 465) comprised 20 302 (46.7%) men and 23 163 (53.3%) women. The greatest proportion of respondents was aged 16 to 20 years (29.3%) and 21 to 25 years (26.3%). With respect to cannabis, 21 075 (48.5%) reported lifetime use, 9498 (45.1% of lifetime users) reported using cannabis in the past 12 months, and 5664 (59.6% of past-year users) reported cannabis use in the past month.

A total of 6419 US residents participated in GDS2014. The US sample of GDS respondents comprised 3846 (61.2%) men and 2439 (38.8%) women. The distribution of respondents was similar across age categories with 18.9% aged 16 to 20 years, 19.8% aged 21 to 25 years, 20.4% aged 26 to 34 years, 20.0% aged 35 to 49 years, and 20.9% aged more than 50 years. With respect to cannabis, 5774 (90.0%) reported that they had used it sometime during their lifetime, 4519 (78.3% of lifetime users) indicated using in the past 12 months, and 3879 (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 data sets. For both the GDS and the NSDUH data, regardless of age, men were typically more likely to report using cannabis than women. 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 to 49 and 50+ years. This is observed in both the GDS and the NSDUH data. In both cases, women were more likely to report cannabis use compared with men.

In both the GDS and the NSDUH data, men and women typically show similar trends of a decreasing probability of lifetime and past-year cannabis use with age. Alternately, the probability of

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

	GDS (NO. %) <sup>a</sup>	NATIONAL SURVEY (NO. %) <sup>a</sup>
<b>Australia</b>	N=5789	N=23855
<i>Sex</i>		
Male	3479 (61.0)	10624 (44.5)
Female	2223 (39.0)	13231 (55.5)
Missing	87	0
<i>Age</i>		
Mean (SD)	37.1 (13.1)	47.5 (16.7)
Missing	75	1157
<i>Cannabis use</i>		
Cannabis ever	4063 (70.2)	8101 (34.4)
Missing	0	330
Cannabis past 12 months	2037 (50.1)	2174 (26.8)
Missing	0	0
Cannabis past month <sup>c</sup>	1379 (67.7)	1143 (52.7)
Missing	0	4
<i>Alcohol use</i>		
Alcohol ever	5711 (98.7)	20881 (87.6)
Missing	0	11
Alcohol past 12 months	5403 (94.6)	19161 (91.8)
Missing	0	0
<b>US</b>	N=6419	N=43465
<i>Sex</i>		
Male	3846 (61.2)	20302 (46.7)
Female	2439 (38.8)	23163 (53.3)
Missing	134	0
<i>Age, y</i>		
16-20	1194 (18.9)	12750 (29.3)
21-25	1251 (19.8)	11433 (26.3)
26-35	1287 (20.4)	5446 (12.5)
36-49	1263 (20.0)	7511 (17.2)
50+	1323 (20.9)	6325 (14.5)
Missing	101	0
<i>Cannabis use</i>		
Cannabis ever	5774 (90.0)	21075 (48.5)
Missing	0	0
Cannabis past 12 months	4519 (78.3)	9498 (45.1)
Missing	0	0

Table 2. (Continued)

	GDS (NO. %) <sup>a</sup>	NATIONAL SURVEY (NO. %) <sup>a</sup>
Cannabis past month <sup>c</sup>	3879 (85.8)	5664 (59.6)
Missing	0	0
<i>Alcohol use</i>		
Alcohol ever	6292 (98.0)	35811 (82.4)
Missing	0	0
Alcohol past 12 months	5621 (89.3)	30758 (85.9)
Missing	0	0
<b>Switzerland</b>	<b>N=4971</b>	<b>N=12008</b>
<i>Sex</i>		
Male	3381 (69.9)	5542 (46.2)
Female	1456 (30.1)	6466 (53.8)
Missing	134	0
<i>Age, y</i>		
15-19	775 (16.0)	2306 (19.2)
20-29	2108 (43.5)	2438 (20.3)
30-39	1067 (22.0)	1043 (8.7)
40-49	533 (11.1)	1625 (13.5)
50+	358 (7.4)	4596 (38.3)
Missing	130	0
<i>Cannabis use</i>		
Cannabis ever	3236 (65.1)	3465 (28.9)
Missing	0	40
Cannabis past 12 months	2056 (63.5)	1057 (30.5)
Missing	0	2
Cannabis past month <sup>c</sup>	1599 (77.8)	481 (45.5)
Missing	0	0
<i>Alcohol use</i>		
Alcohol ever	4913 (98.8)	11136 (92.7)
Missing	0	0
Alcohol past 12 months	4679 (95.2)	10353 (93.5)
Missing	0	58

Abbreviation: GDS, Global Drug Survey.

<sup>a</sup>Percentages are of valid cases (excluding missings) and may sum to greater than 100 due to rounding.

<sup>b</sup>Past year only includes respondents who indicated that they had used cannabis/alcohol in their lifetime.

<sup>c</sup>Past month only includes respondents who indicated that they had used cannabis in the past 12 months.

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 women 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 women 26 to 34 years and older).

Although the probability of ever using cannabis is higher in the GDS sample, the probability of using cannabis in the past year

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

	GDS (NO. %) <sup>a</sup>		NATIONAL SURVEY (NO. %) <sup>a</sup>	
	MALE	FEMALE	MALE	FEMALE
<b>Australia</b>	N=3479	N=2223	N=10357	N=12976
<i>Age, y</i>				
16-20	249 (7.2)	135 (6.1)	622 (5.9)	686 (5.2)
21-25	465 (13.4)	380 (17.1)	616 (5.8)	763 (5.8)
26-34	1029 (29.6)	716 (32.2)	1494 (14.1)	2183 (16.5)
35-49	1011 (29.1)	590 (26.5)	2374 (22.4)	3228 (24.4)
50+	725 (20.8)	402 (18.1)	5251 (49.4)	6116 (42.2)
<b>US</b>	N=3837	N=2430	N=20302	N=23163
<i>Age, y</i>				
16-20	921 (24.0)	267 (11.0)	6380 (31.4)	6370 (27.5)
21-25	797 (20.8)	445 (18.3)	5233 (25.8)	6200 (26.8)
26-34	736 (19.2)	547 (22.5)	2448 (12.1)	2998 (12.9)
35-49	679 (17.7)	575 (23.7)	3426 (16.9)	4085 (17.6)
50+	704 (18.5)	596 (24.5)	2815 (13.9)	3510 (15.2)
<b>Switzerland</b>	N=3361	N=1451	N=5542	N=6466
<i>Age, y</i>				
15-19	537 (15.9)	237 (16.3)	1177 (21.2)	1129 (17.5)
20-29	1374 (40.6)	728 (50.1)	1186 (21.4)	1252 (19.4)
30-39	767 (22.7)	293 (20.2)	440 (7.9)	603 (9.3)
40-49	408 (12.1)	119 (8.2)	708 (12.8)	917 (14.2)
50+	275 (8.2)	74 (5.1)	2031 (36.7)	2565 (39.7)

Abbreviation: GDS, Global Drug Survey.

<sup>a</sup>Percentages are of valid cases (excluding missings).

among lifetime users, and using within the past month among past-year users, is similar across the GDS and NSDUH data sets.

*Switzerland.* As presented in Tables 2 and 3, the analytic sample from the AMIS (N = 12008) comprised 5542 (46.2%) men and 6466 (53.8%) women. The greatest proportion of respondents was aged more than 50 years (38.3%) and 20 to 29 years (20.3%). With respect to cannabis, 3465 (28.9%) reported lifetime use, 1057 (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 4972 Swiss residents participated in GDS2014. The Swiss sample of GDS respondents consisted of 3381 (69.9%) men and 1456 (30.1%) women. The greatest proportion of respondents was aged 20 to 29 years (43.5%) and 30 to 39 years (22.0%). With respect to cannabis use, 3236 (65.1%) reported lifetime use, 2056 (63.5% of lifetime users) reported

using cannabis in the past year, and 1599 (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 data sets. For both the GDS and the AMIS data, regardless of age, men were typically more likely to report using cannabis than women. Only in the GDS, did women report higher prevalence of cannabis use than men for use in the past month. As shown in the GDS figure for past-month use, there is a constant, slightly increasing probability of men reporting use of cannabis within the past month across the age groups. However, for women, the probability of using cannabis within the past month increases substantially between the ages of 20 to 29 years and 40 to 49 years at which point the probability of using cannabis in the past month is greater for women than for men.

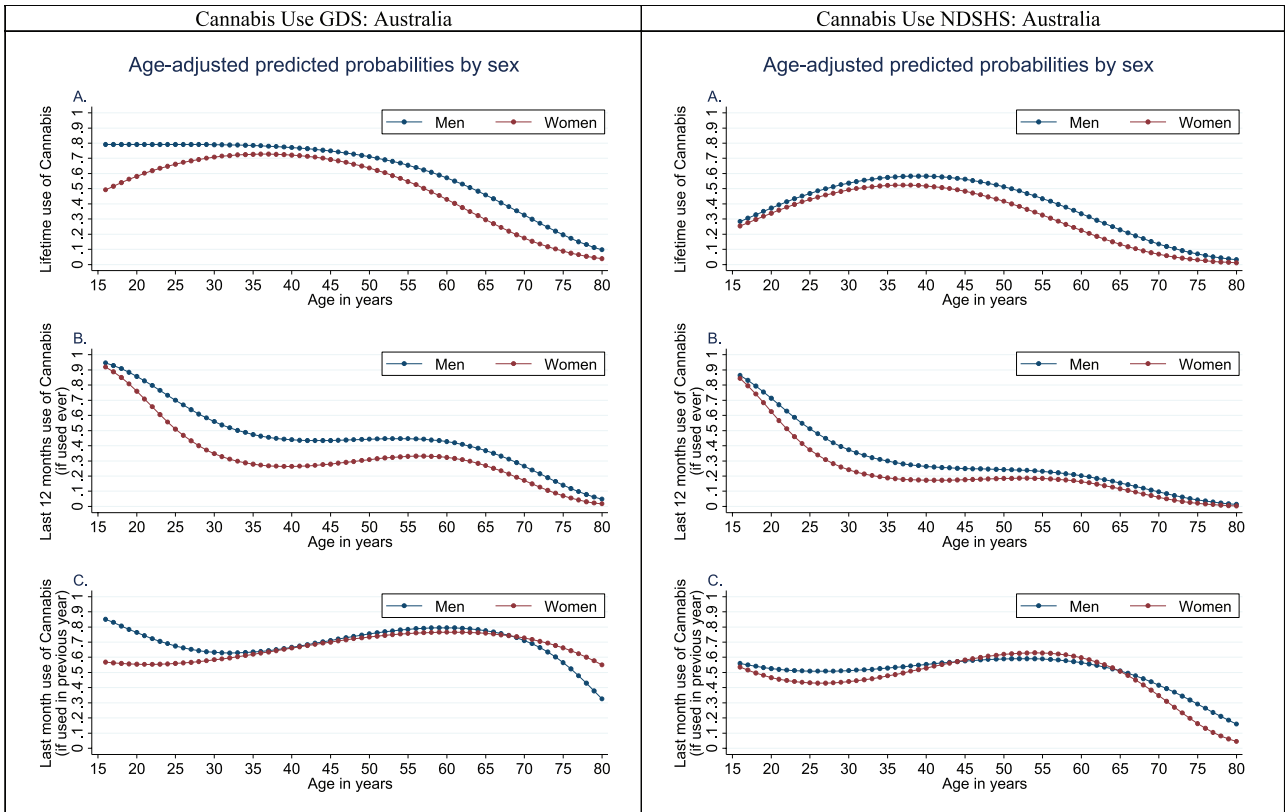


Figure 1. Cannabis use in Australia – GDS and NDSHS. GDS, Global Drug Survey; NDSHS, National Drug Strategy Household Survey.

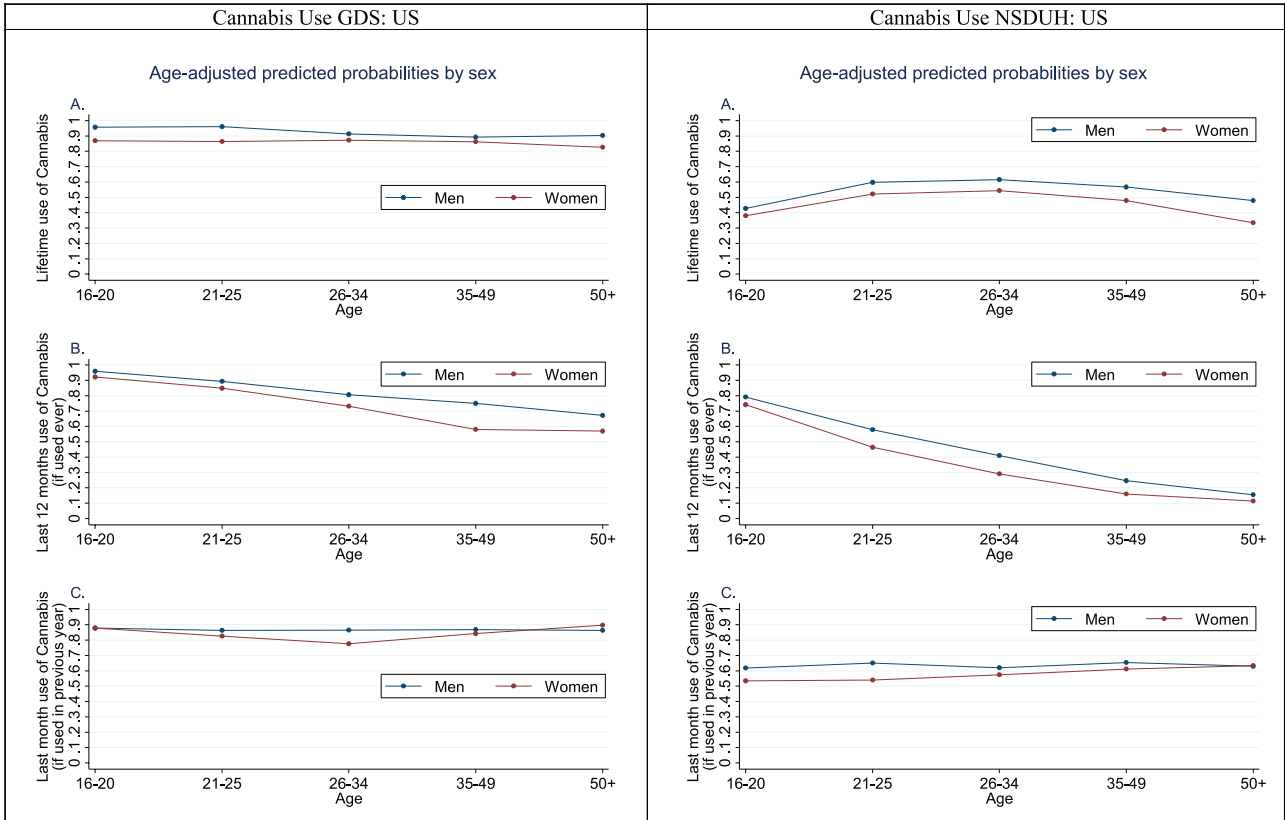


Figure 2. Cannabis use in United States – GDS and NSDUH. GDS, indicates Global Drug Survey; NSDUH, National Survey on Drug Use and Health.

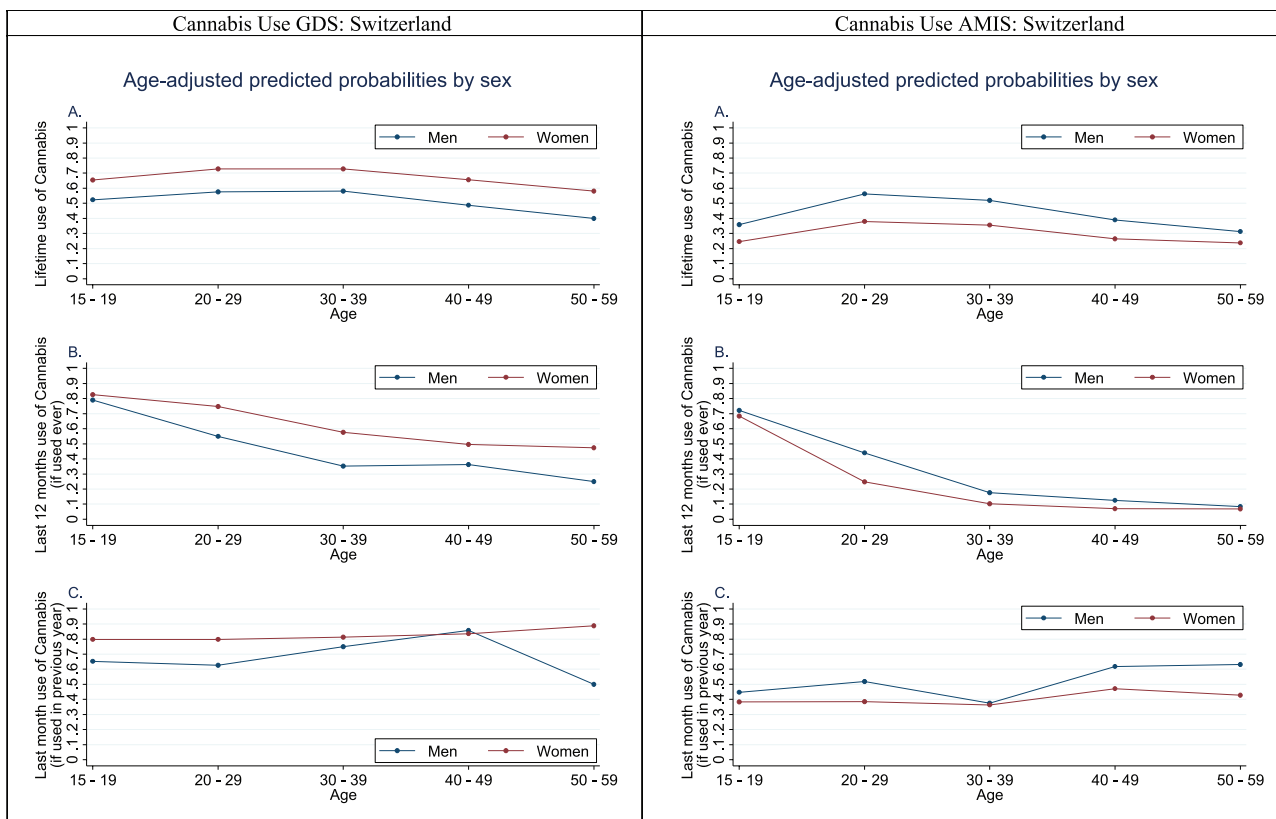


Figure 3. Cannabis use in Switzerland – GDS vs AMIS. GDS, indicates Global Drug Survey; AMIS, Addiction Monitoring in Switzerland.

Overall, men and women 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 data sets, particularly for those aged more than 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 more than 50 years and reported cannabis use in the past year.

**Discussion**

Although the probability of ever using cannabis is higher in the GDS sample, as would be expected, once the data sets 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 is similar across the GDS and the equivalent probability samples, across the three countries. That is, we can see in Figures 1 to 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 – Australia, the United States, 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 1379 past-month cannabis users from surveying 5789, compared with 1143 past-month cannabis users from surveying 23 855 via the household survey. In the United States, a greater number of past-month cannabis users was recruited through the probability survey (5664) compared with the GDS (3879), the total sample surveyed by the US probability survey (43 465) was almost seven times the total number surveyed by GDS (6419). In Switzerland, the GDS recruited 1599 past-month cannabis users compared with 481 recruited by the probability sampling. Although it is beyond the scope of this article 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 are 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?'.<sup>14,15</sup> The analysis presented in this article 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, whereas 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 such as GDS. Further work of this kind should be conducted, for example, looking at characteristics of populations of novel substance 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 article. First, the use of age categories in the NSDUH and AMIS prevented a more nuanced probability analysis as was conducted with the Australian data. Furthermore, 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 (eg, GDS uses a web interface, NDSHS uses drop and collect, NSDUH uses computer-assisted personal interviewing, 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. Furthermore, 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 focused 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.<sup>65</sup> 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 synthesise these diverse data sources into the most robust and valid picture of drug use within our communities. Global Drug Survey 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 that GDS can make in its procedures to be more inclusive. For example, we could recruit booster samples from groups less well represented (eg, women, 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 (eg, translation from English, to Spanish, and 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.

### Conclusions

In this article, 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 with 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 article 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,<sup>22,25,49,66,67</sup> this article 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 stigmatised

behaviours, and are appropriate, as long as they are not used to estimate drug use prevalence of the general population.

### Acknowledgements

The authors would like to thank the participants who gave so generously of their time to complete the GDS. They are grateful for the promotion of GDS by a long list of world media partners. They 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. They thank the Australian Institute of Health and Welfare (AIHW) and the Australian Data Archive (ADA) for providing access to the NDSHS data, the Swiss Federal Office of Public Health for providing access to the Addiction Monitoring in Switzerland (AMIS) 2013 to 2014 data set (<http://suchtmonitoring.ch/>), and the Interuniversity Consortium for Political and Social Research for providing access to NSDUH data (<http://www.icpsr.umich.edu/icpsr-web/landing.jsp>). They 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.

### Author Contributions

MJB, JAF, LJM, and ARW conceived of and developed the paper and together form the GDS Core Research Team. MJB wrote the first draft of the manuscript. RZ, JAF, LJM, and JJP conducted statistical analyses. All authors reviewed, made critical revisions, and approved the final manuscript.

### 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

- 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.
- Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42:1012–1014.
- Gobo G, Mauceri S. *Constructing Survey Data: An Interactional Approach*. London, England: SAGE; 2014.
- Winstock AR, Mitcheson LR, Deluca P, Davey Z, Corazza O, Schifano F. Mephedrone, new kid for the chop? *Addiction*. 2011;106:154–161.
- Winstock AR, Mitcheson L, Gillatt DA, Cottrell AM. The prevalence and natural history of urinary symptoms among recreational ketamine users. *BJU Int*. 2012;110:1762–1766.
- 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.
- Barratt MJ, Ferris JA, Winstock AR. Use of Silk Road, the online drug marketplace, in the UK, Australia and the USA. *Addiction*. 2014;109:774–783.
- 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:e010112.
- Rothman KJ, Gallacher JE, Hatch EE. Rebuttal: when it comes to scientific inference, sometimes a cigar is just a cigar. *Int J Epidemiol*. 2013;42:1026–1028.
- Richiardi L, Pizzi C, Pearce N. Commentary: representativeness is usually not necessary and often should be avoided. *Int J Epidemiol*. 2013;42:1018–1022.
- 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;42:1016–1017.
- Elwood JM. Commentary: on representativeness. *Int J Epidemiol*. 2013;42:1014–1015.
- Ebrahim S, Davey Smith G. Commentary: should we always deliberately be non-representative? *Int J Epidemiol*. 2013;42:1022–1026.
- Pasek J. When will nonprobability surveys mirror probability surveys? considering types of inference and weighting strategies as criteria for correspondence. *Int J Public Opin R*. 2016;28:269–291.
- Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J Roy Stat Soc A Sta*. 2016;179:319–376.
- Kreuter F. Facing the nonresponse challenge. *Ann Am Acad Polit SS*. 2013;645:23–35.
- Zhao J, Stockwell TIM, Macdonald S. Non-response bias in alcohol and drug population surveys. *Drug Alcohol Rev*. 2009;28:648–657.
- 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.
- Skalland B, Khare M. Geographic inaccuracy of cell phone samples and the effect on telephone survey bias, variance, and cost. *J Surv Stat Methodol*. 2013;1:45–65.
- 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;12:45.
- 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 Med Res Methodol*. 2013;13:41.
- Barratt MJ, Ferris JA, Lenton S. Hidden populations, online purposive sampling, and external validity: taking off the blindfold. *Field Methods*. 2015;27:3–21.
- Groves RM. Three eras of survey research. *Public Opin Q*. 2011;75:861–871.
- 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, England: SAGE; 2008:195–216.
- Barratt MJ, Lenton S. Beyond recruitment? Participatory online research with people who use drugs. *Int J Internet Res Ethics*. 2010;3:69–86.
- Miller PG, Sønderlund AL. Using the internet to research hidden populations of illicit drug users: a review. *Addiction*. 2010;105:1557–1567.
- 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.
- van Gelder MM, Bretveld RW, Roeleveld N. web-based questionnaires: the future in epidemiology? *Am J Epidemiol*. 2010;172:1292–1298.
- 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.
- Kruskal W, Mosteller F. Representative sampling, I: non-scientific literature. *Int Stat Rev*. 1979;47:13–24.
- Kruskal W, Mosteller F. Representative sampling, II: scientific literature, excluding statistics. *Int Stat Rev*. 1979;47:111–127.
- Kruskal W, Mosteller F. Representative sampling, III: the current statistical literature. *Int Stat Rev*. 1979;47:245–265.
- Kruskal W, Mosteller F. Representative sampling, IV: the history of the concept in statistics, 1895–1939. *Int Stat Rev*. 1980;48:169–195.
- Gobo G. Sampling, representativeness and generalizability. In: Seale C, Gobo G, Gubrium JF, Silverman D, eds. *Qualitative Research Practice*. London, England: SAGE; 2007:405–426.
- McCreech N, Copas A, Seeley J, et al. Respondent driven sampling: determinants of recruitment and a method to improve point estimation. *PLoS ONE*. 2013;8:e78402.
- Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. *J Offic Stat*. 2008;24:79–97.



37. Global Drug Survey. Global Drug Survey Brochure. [https://www.globaldrug-survey.com/wp-content/uploads/2016/09/GDS\\_Brochure\\_2016.pdf](https://www.globaldrug-survey.com/wp-content/uploads/2016/09/GDS_Brochure_2016.pdf). Published 2016.
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:e47.
39. Barratt MJ, Ferris JA, Winstock AR. Safer scoring? Cryptomarkets, social supply and drug market violence. *Int J 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;29:698–703.
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. *Hum Psychopharmacol*. 2013;28:390–393.
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. *J Subst Abuse*. 2016;21:115–120.
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 Alcohol Rev*. 2016;35:212–217.
47. Winstock A, Bell J, Borschmann R. Friends, doctors, and tramadol: we might have a problem. *BMJ*. 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;68:1147–1151.
49. Barratt MJ, Potter GR, Wouters M, et al. Lessons from conducting trans-national internet-mediated participatory research with hidden populations of cannabis cultivators. *Int J 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. *The Sydney Morning Herald*. November 16, 2016. <http://www.smh.com.au/national/health/drug-users-swap-stories-share-warnings-online-in-search-for-safety-20161111-gsnnt6.html>.
52. Aubusson K, Spooner R. Global Drug Survey set to peel back curtain on professionals microdosing on LSD and magic mushrooms. *The Sydney Morning Herald*. December 2, 2016. <http://www.smh.com.au/national/health/global-drug-survey-set-to-peel-back-curtain-on-professionals-microdosing-on-bsd-and-magic-mushrooms-20161116-gsqi74.html>.
53. Winstock A, Barratt M. Dark web markets: the good, the bad & the ugly. *High Times*. December 12, 2016. <http://hightimes.com/business/dark-web-markets-the-good-the-bad-the-ugly/>.
54. Winstock A, Barratt M. Zero tolerance for drugs doesn't work. Let's embrace zero harm. *The Sydney Morning Herald*. December 29, 2016. <http://www.smh.com.au/comment/zero-tolerance-for-drugs-doesnt-work-lets-embrace-zero-harm-20161228-gtiukr.html>.
55. Bowen A, Daniel C, Williams M, Baird G. Identifying multiple submissions in internet research: preserving data integrity. *AIDS Behav*. 2008;12:964–973.
56. von Lüsi M. Etwa jeder 10. Konsument hat ein Drogenproblem [About every 10th consumer has a drug problem]. *20 Minuten*. June 14, 2016. <http://www.20min.ch/schweiz/news/story/-Etwa-jeder-10-Konsument-hat-ein-Drogenproblem-29541451>
57. Eysenbach G, Wyatt J. Using the internet for surveys and health research. *J Med Internet Res*. 2002;4: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. *Int J 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;126:887–893.
60. Australian Institute of Health and Welfare. *2013 National Drug Strategy Household Survey* (Detailed report). Canberra, ACT: Australian Institute of Health and Welfare; 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 (SAMHSA). *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*. Rockville, MD: SAMHSA; 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;111:34–48.
66. Barratt MJ. The efficacy of interviewing young drug users through online chat. *Drug Alcohol Rev*. 2012;31:566–572.
67. Barratt MJ, Maddox A. Active engagement with stigmatised communities through digital ethnography. *Qual Res*. 2016;16:701–719.