International Journal of Drug Policy xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

International Journal of Drug Policy

journal homepage: www.elsevier.com/locate/drugpo

Research paper

Are cannabis prevalence estimates comparable across countries and regions? A cross-cultural validation using search engine query data

Martin Steppan[a,*], Ludwig Kraus[a,b], Daniela Piontek[a], Valeria Siciliano[c]

[a] IFT Institut für Therapieforschung, München, Germany
[b] Centre for Social Research on Alcohol and Drugs, SoRAD, Stockholm University, Stockholm, Sweden
[c] Consiglio Nazionale delle Ricerce, Institute of Clinical Physiology, Pisa, Italy

ARTICLE INFO

Article history:
Received 16 August 2011
Received in revised form 25 April 2012
Accepted 10 May 2012

Keywords:
Cannabis use prevalence
Cultural effects
Validation
School survey

ABSTRACT

Background: Prevalence estimation of cannabis use is usually based on self-report data. Although there is evidence on the reliability of this data source, its cross-cultural validity is still a major concern. External objective criteria are needed for this purpose. In this study, cannabis-related search engine query data are used as an external criterion.
Methods: Data on cannabis use were taken from the 2007 European School Survey Project on Alcohol and Other Drugs (ESPAD). Provincial data came from three Italian nation-wide studies using the same methodology (2006–2008; ESPAD-Italia). Information on cannabis-related search engine query data was based on Google search volume indices (GSI). (1) Reliability analysis was conducted for GSI. (2) Latent measurement models of "true" cannabis prevalence were tested using perceived availability, web-based cannabis searches and self-reported prevalence as indicators. (3) Structure models were set up to test the influences of response tendencies and geographical position (latitude, longitude). In order to test the stability of the models, analyses were conducted on country level (Europe, US) and on provincial level in Italy.
Results: Cannabis-related GSI were found to be highly reliable and constant over time. The overall measurement model was highly significant in both data sets. On country level, no significant effects of response bias indicators and geographical position on perceived availability, web-based cannabis searches and self-reported prevalence were found. On provincial level, latitude had a significant positive effect on availability indicating that perceived availability of cannabis in northern Italy was higher than expected from the other indicators.
Conclusion: Although GSI showed weaker associations with cannabis use than perceived availability, the findings underline the external validity and usefulness of search engine query data as external criteria. The findings suggest an acceptable relative comparability of national (provincial) prevalence estimates of cannabis use that are based on a common survey methodology. Search engine query data are a too weak indicator to base prevalence estimations on this source only, but in combination with other sources (waste water analysis, sales of cigarette paper) they may provide satisfactory estimates.

Introduction

Prevalence estimation of cannabis use is usually based on survey data. The resulting estimates are regularly compared across countries without having detailed knowledge about their validity (Hibell et al., 2009; Smart & Ogborne, 2000). Although reliability of drug prevalence studies has proved to be satisfactory by means of retest, parallel test, and analysis of inconsistent answers, the "validity of answers is a major concern in survey research, particularly in surveys of sensitive behaviours such as substance use" (Hibell et al., 2009; Molinaro, Siciliano, Curzio, Denoth, & Mariani, 2012).

In general, self-reports of drug use may be distorted by denial or exaggeration. Such response biases may be due to an individual's tendency to give social desirable answers (Dillman, 2000). Students admitting more risk behaviour at home than in school or reporting use of a fantasy drug are empirical examples of such response tendencies (Brener et al., 2010; Hibell et al., 2009). In addition, reporting of drug use may be influenced by the individual's normative assessment of legal regulations, law enforcement measures, social norms as well as attitudes and perceived risks related to drugs and drug use behaviour. According to Harrison, more stigmatised drugs are less validly reported than less stigmatised ones (Harrison, 1997). Assuming variation in the normative

* Corresponding author at: IFT Institut für Therapieforschung, Parzivalstrasse 25, 80804 München, Germany. Tel.: +49 089 360804 62; fax: +49 089 360804 69.
E-mail address: steppan@ift.de (M. Steppan).

0955-3959/$ – see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.drugpo.2012.05.002

rating of drugs across countries, prevalence estimates of drug use in countries with stricter drug norms may be less valid.

Self-reports have been validated using bogus pipeline methods (a "pseudo lie detector" that increases the probability of honest response behaviour) or biological tests based on saliva, urine or hair samples (Campanelli, Dielman, & Shope, 1987; Kokkevi & Stefanis, 1991; Wish, Hoffman, & Nemes, 1997). A meta-analysis summarizing the results of 24 studies using biological measures found substantial evidence for biased measurements in terms of underreported drug use (Magura & Kang, 1995). Obviously, individual level biases will affect country level prevalence rates, over- or underestimating the true prevalence. If the bias is constant across countries, the ranking of countries in terms of national drug prevalence rates will not be affected. If the bias varies from country to country, national prevalence estimates will be differentially affected, and consequently not be comparable.

In order to test whether these individual level response biases seriously affect prevalence estimates indirect measures that do not rely on self-reported information are needed. Person et al., for instance, estimated the prevalence of heroin use on regional level for the US by combining external drug-related indicators such as price and treatment demand (Person, Retka, & Woodward, 1978). Others applying the same method used drug-related, social and demographic indicators such as drug-related deaths, drug seizures, AIDS related to drug use, residential mobility, unemployment rate, population size or housing density (Brugal et al., 1999; Comiskey, 2001; Hser, Prendergast, Anglin, Chen, & Hsieh, 1998; Rhodes, 1993; Smit, Toet, Oers, & Wiessing, 2003; van Nuijs et al., 2009; Wickens, 1993). While these indicators may work within a single country, once applied across countries they are subject to country-specific influencing mechanisms and cannot be used for cross-national validation. For instance, autopsy practice, treatment demand or drug seizures strongly depend on legal regulations, health care budgeting, law enforcement and police activities. More recently, "sewage epidemiology" indirectly estimated drug consumption by analysing river water on the occurrence of human metabolites of drugs (van Nuijs et al., 2009). Although, detecting consumption-peeks of cocaine at weekends, this method was not comprehensively applied and seems to measure the total amount of drugs consumed rather than to estimate national or regional prevalence.

Consequently, for an unbiased prediction of national cannabis use prevalence indicators need to be independent of (a) drug users' potential response bias and (b) country-specific influencing factors. One potential indicator that has recently been described by Hibell and Anderson (2008) measures perceived availability based on self-reports of the likelihood of access to cannabis within 24 hours. Predicting national cannabis prevalence rates by perceived availability the authors were able to explain about 70% of the variance. Another approach using search engine query data of corresponding search terms (brand names, flu symptoms) has been successfully used for predicting sales in economics or influenza epidemics (Choi & Varian, 2009; Ginsberg et al., 2008). One may argue that neither web-based cannabis searches nor reports on the likelihood of getting access to cannabis are disclosing an individual's attitude about drugs or potential drug using behaviour. Moreover, on national level they are not subject to influences other than the interest to search the Internet for information or the perceived availability of cannabis.

Considering the lack of information about the validity of self-report prevalence rates, there is need for validation of aggregate survey data on country and regional level, Cannabis-related search engine query data are used as an external criterion of the "true cannabis prevalence". Moreover, self-reported prevalence data and perceived availability are used as internal (survey based) criteria. This paper aims at (1) testing whether these three criteria measure the same latent construct "true cannabis prevalence", (2) assessing the impact of response tendencies and geographical variation on the criteria and the true prevalence, and (3) testing these associations at regional level. In the latter case, the variability of external country-specific influences on the indicators (e.g. language, legislation) is assumed to be reduced in a single country.

Methods

Data

For this study, (a) international data from the European School Survey Project on Alcohol and other Drugs (ESPAD) and (b) national data based on the same survey in Italy were used. In the 2007 ESPAD study (Hibell et al., 2009), students in 36 European countries were surveyed in a class room setting by filling in self-administered questionnaires (total $N = 104{,}828$). Countries participating in ESPAD were Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Faroe Island, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Isle of Man, Italy, Latvia, Lithuania, Malta, Monaco, Netherlands, Norway, Poland, Portugal, Romania, Russia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Ukraine, and the United Kingdom. Six countries (Cyprus, Faroe Islands, Isle of Man, Iceland, Malta, Monaco) had to be excluded because no information on cannabis-related search engine query data was available. The study aimed at a random sample of students of the 1991 birth cohort, i.e. students aged 15–16 years. Sample sizes varied between 877 and 9981 in Denmark and Italy, respectively. Non-response rates varied between .1% (Isle of Man, Poland) and 1.8% (Norway). In the US, the same questionnaire was used in a sample of $N = 16{,}000$ students (Johnston, O'Malley, Bachman, & Schulenberg, 2008). To reduce the influence of cultural heterogeneity (legislation, language, Internet distribution), an additional analysis was conducted for Italy in which data could be broken down at provincial level. To this end, data from three consecutive surveys (2006–2008) for 103 Italian provinces were used. These surveys applied the ESPAD methodology. Sample sizes were 7742 (2006), 9981 (2007) and 7737 (2008) (Molinaro & Siciliano, 2009).

Measures

Cannabis prevalence. The question assessing cannabis use read "On how many occasions (if any) have you used marihuana or hashish (cannabis)?". This question was asked for (a) in your lifetime, (b) during the last 12 months, and (c) during the last 30 days. Response categories were 0, 1–2 times, 3–5 times, 6–9 times, 10–19 times, 20–39 times, 40 or more times. Lifetime, 12-month and 30-day prevalence rates were calculated based on individuals' responses of having used cannabis at least once in the given time period.

Cannabis availability. The level of perceived access to cannabis was measured by asking students "How easy do you think it would be for you to get marihuana or hashish (cannabis) if you wanted?". Response options were impossible, very difficult, fairly difficult, fairly easy, very easy, don't know. Aggregated means of responses for the ESPAD countries and the US were taken from the international report (Hibell et al., 2009). Data for Italian provinces was provided by ESPAD-Italia (Molinaro & Siciliano, 2009).

Web-based searches for cannabis-related terms were measured using Google search volume indices (GSI) (Google Labs, 2010). This source provides data on the number of requests for a search term in a given area, divided by the amount of all Google searches in this area. The index reflects the probability of a term to be searched in a particular area. GSI were extracted for the years 2004–2011 in order to test the temporal stability. GSI were available for more

Table 1
Analysis of reliability of cannabis web traffic (GSI).

Period	Country level			Provincial level		
	Search term	Correlation	Alpha	Search term	Correlation	Alpha
Thematic consistency						
2004–2011	cannabis	.4114	.8119	canapa	.9126	.9463
2004–2011				canna	.5794	.9537
2004–2011				cartine[a]	.7819	.9490
2004–2011				erba[a]	.0422	.9579
2004–2011	ganja	.1048	.8356	ganja	.7634	.9497
2004–2011	grass	.7508	.7816	grass	.7606	.9505
2004–2011	hashish	.2287	.8249	hashish	.8429	.9478
2004–2011				il fumo[a]	.8708	.9473
2004–2011	legalize	.4717	.8062	legalize	.4800	.9541
2004–2011	marijuana	.7800	.7796	marijuana	.8621	.9474
2004–2011				paranoia	.8108	.9485
2004–2011	purple haze[a]	.6472	.7906	purple haze[a]	.3631	.9554
2004–2011	spice	.3672	.8150	spice	.9134	.9466
2004–2011				spinello[a]	.6703	.9516
2004–2011	THC	.3230	.8234	THC	.8177	.9491
2004–2011	weed	.8655	.7689	weed	.8131	.9484
2004–2011	joint	.5367	.8012			
Scale			.8194			.9530
Temporal stability						
2004	All[b]	.5618	.9641	All[b]	.9840	.9931
2005	All[b]	.7637	.9552	All[b]	.9427	.9973
2006	All[b]	.7667	.9537	All[b]	.9644	.9964
2007	All[b]	.9262	.9436	All[b]	.9912	.9953
2008	All[b]	.9425	.9425	All[b]	.9900	.9954
2009	All[b]	.9352	.9431	All[b]	.9826	.9953
2010	All[b]	.9013	.9461	All[b]	.9849	.9948
2011	All[b]	.9013	.9461	All[b]	.9727	1.0110
Scale			9556			.9948

THC: tetrahydrocannabinol; correlation: item-rest correlation (discriminatory power of item); Alpha: alpha of scale after exclusion of the item. Country level: N = 81 countries; Italy: N = 103 provinces.

[a] Translations: il fumo: the smoke; cartine: papers; spinello: joint; erba: herb; purple haze: famous sort of cannabis.

[b] Summative GSI for all of the above mentioned search terms was used.

countries (N = 81) than ESPAD prevalence data. For Cyprus, Faroe Islands, Iceland, Isle of Man, Malta and Monaco, no valid figures could be determined. For Armenia, the search volume was too low and GSI were set "0". Cannabis-related search terms were *cannabis, ganja, grass, hashish, marijuana, purple haze, THC* [Tetrahydrocannabinol], *weed* and *joint*. The terms *legalize* and *spice* were also used as common interests in this field. For Italy, further local expressions were used: *canna, cartine, erba, il fumo* and *spinello* (for translations see footnote in Table 1).

Indicators of response biases and geographical position. To measure the tendency to overreport drug experience, a fantasy drug was introduced in the ESPAD questionnaire. Different fantasy names were used (e.g. "Relevin" in Germany, "Netalin" in Italy) and the prevalence of its use was assessed in the same way as for cannabis described above. Any response except "0 times" was taken as an indicator of overreporting. The proportion varied between .1% (Romania) and 1.9% (Isle of Man) (Hibell et al., 2009). To measure the tendency to underreport drug use, the proportion of individuals answering "definitely not" to the question "If you had ever used marijuana or hashish (cannabis) in your life, do you think you would have said so in this questionnaire?" was used. The percentage varied between 3% in Slovenia and 17% in Lithuania. Data on both indicators were missing for the US and Spain (Hibell et al., 2009). Geographical position (latitude and longitude) was retrieved from http://www.wikipedia.org.

Statistical analyses

Reliability analysis. Reliability was measured to test whether GSI are temporally stable and thematically consistent. Internal consistency (Cronbach's α), discriminatory power (correlation of each annual GSI with the average across all other GSI) and squared multiple correlation were tested. Analysis was performed across time and search terms. Cronbach's α was also calculated without each annual GSI to determine the effect of a GSI's elimination. Global reliability analysis was performed using data from all countries where GSI were available.

Structural equation modelling (SEM) represents a statistical technique that aims at estimating causal effects based on unobserved (latent) dimensions. Testing the relationship between indicators and one or more latent dimension(s) denotes a measurement model that is equivalent to confirmatory factor analysis (CFA; Kline, 2010). Including also statistical effects on the latent dimensions (e.g. by covariates) turns a measurement model into a structure model (Geiser, 2010). Covariates may be reasonably related to the latent dimension (e.g. education and intelligence) not impairing the quality of the measurement model. However, effects of covariates on single indicators of a latent factor suggest an impaired measurement model in terms of non-invariance of the indicator. Compared to regression analysis, SEM does not require one indicator to be predicted by the other indicators. On the contrary, SEM estimates a latent dimension that explains each indicator. If none of the indicators can be *a anteriori* considered the "true" criterion, SEM takes this uncertainty into account. To this end, SEM has often been applied for the identification of dimensionality and impaired measurement (e.g. Gillespie, Neale, Prescott, Aggen, & Kendler, 2007).

In a first step, latent measurement models of "true" cannabis prevalence were tested using perceived availability, web-based cannabis searches and self-reported prevalence as independent manifest indicators. On country level, self-reported prevalence was modelled as latent dimension using measures on lifetime, 12 month and 30 day use. Availability and web-based searches based on the search term "THC" were included as manifest observations. The term THC was used because of its ubiquitous use in all languages. All variables were based on 2007 data. Due to follow-up availability (2006–2008) of the Italian data, perceived availability and web-based cannabis searches could also be modelled as latent variables. Information on self-reported prevalence was not available for the last 30 days.

In a second step, structure models were set up including response bias indicators and geographical position (covariates). The model with the best fit was identified using a stepwise approach: (1) the effect of response tendencies on each independent indicator and the "true" cannabis prevalence was tested and (2) the effect of geographical position (latitude, longitude) on the independent indicators, response tendencies and the "true" prevalence was tested. Finally, a full model was tested. Significant paths from covariates to indicators suggest non-invariant measurement. If response tendencies vary across countries or regions, significant effects of the covariates on the indicators are expected. Notwithstanding, significant effects of the covariates on the "true prevalence" may reflect reasonable associations, they do, however, not question measurement invariance.

Each measurement model's goodness of fit was tested separately. For all structure models only those paths with a significance probability below $p < .15$ were considered. Model fits for all models were examined using the comparative fit index (CFI), the standardized root mean square residual (SRMR), and the Tucker–Lewis index (TLI). Recommended cut-off points for these measures are: CFI > .96, SRMR < .05 and TLI > .95 (Bühner, 2006; Geiser, 2010).

Results

Reliability of cannabis GSI

Table 1 shows the results of reliability analyses in terms of thematic consistency and temporal stability. Thematic consistency was found lower on country level ($\alpha = .8194$) than on provincial level ($\alpha = .9530$) considering the overall reliability of all search terms. However, some of the search terms had low discriminatory power and impaired the scale's reliability, e.g. ganja" on country level ($r = .10$) or "erba" in Italy ($r = .04$). Based on discriminatory power, "marijuana" and "weed" most representatively assessed the scale on country level. "Spice" and "canapa" were the most representative search terms in Italy. Temporal stability was very high on both levels. The summative GSI was highly constant over time on country level ($\alpha = .9556$) and in Italy ($\alpha = .9948$). The highest item-rest correlation was found in the years 2007–2009.

Structural equation models

Country level

The full structure model is shown in Fig. 1 while fit indices of all measurement and structure models are depicted in Table 2. Although assessing different time spans, lifetime- ($\lambda = .97$, $p < .001$), 12-month ($\lambda = 1.01$, $p < .001$) and 30-day prevalence ($\lambda = .93$, $p < .001$) constituted a reliable sub-measurement model, which itself showed high model fit (CFI = 1.000; TLI = 1.000; SRMR = .000). Self-reported prevalence ($\lambda = .89$, $p < .001$) and perceived availability ($\lambda = .91$, $p < .001$) displayed about equal factor loadings on the latent "true" cannabis prevalence. Web-based searches was a weaker, but significant indicator of the latent dimension ($\lambda = .55$, $p < .001$). The overall measurement model was significant (CFI = .981; TLI = .952; SRMR = .039).

Table 2
Hierarchical (stepwise) tabulation of goodness of fit.

	CFI	TLI	SRMR
Country level			
Self-reported prevalence	**1.000**	**1.000**	**.000**
True prevalence	**.981**	**.952**	**.039**
True prevalence + geographic position	**.986**	.927	.061
True prevalence + response bias indicators	.941	.918	.123
Full model	.926	.901	.120
Italian provinces			
Self-reported prevalence	**1.000**	**1.003**	**.009**
Self-reported perceived availability	**1.000**	**1.000**	**.000**
Web-based searches	.921	.881	**.020**
True prevalence	.908	.889	.048
Underreporting	**1.000**	**1.000**	**.000**
Overreporting		No convergence	
True prevalence + geographic position	.907	.891	**.049**
True prevalence + response bias indicators	.908	.889	**.048**
Full model	.908	.893	.060

CFI: comparative fit index; TLI: Tucker–Lewis index; SRMR: standardized root mean square residual. Recommended cut-off points (bold values): CFI > .96, SRMR < .05 and TLI > .95.

The strongest geographical effect on "true" cannabis prevalence was found for *longitude* ($\beta = -.66$, $p < .001$) indicating a west-east decline in cannabis use. This structure model showed an adequate fit according to CFI (CFI = .956). No significant effects of the response bias indicators on availability, self-reported prevalence, web-based THC searches and the "true" prevalence were found. However, underreporting was affected by a *longitude* indicating an increase from west to east ($\beta = .46$, $p < .085$). Whereas the measurement models showed excellent model fit, the fit indices decreased with increasing complexity of the structure models.

Provincial level

Fig. 2 depicts the full structure model for the data on provincial level in Italy. Perceived availability, self-reported prevalence and underreporting (all CFI = 1.000; TLI ≥ 1.000; SRMR = .000) represent highly significant measurement models on their own. However, factor loadings of consecutive years on underreporting were rather low ($.39 \leq \lambda = .49$). Despite high factor loadings ($.97 \leq \lambda = 1.000$), the measurement model of cannabis-related web searches had lower goodness of fit (CFI = .921; TLI = .881; SRMR = .02). The "true" prevalence measurement model indicates that web-based query data are a modest, but significant ($\lambda = .23$, $p < .05$) indicator of the "true" prevalence, compared to perceived availability ($\lambda = .74$, $p < .001$) and self-report prevalence ($\lambda = 1.02$, $p < .001$). The latent model for overreporting did not converge. Combining all models with the exception of overreporting within a full structure model lead to a reduced model fit (CFI = .908; TLI ≥ .893; SRMR = .060). Again, longitude ("easterness") affected underreporting ($\beta = .34$; $p < .05$) and the "true" prevalence ($\beta = -.79$; $p < .001$). Moreover, latitude ("northness") had a significant positive effect on availability ($\beta = .27$; $p < .001$) indicating that perceived availability of cannabis in northern Italy was higher than expected from the other indicators. Contrary, no effect of underreporting on either of the indicators was found.

Discussion

Using structural equation modelling this study examined whether direct and indirect indicators of cannabis use prevalence on country and regional level are affected by individual response biases and country-specific influences. The proposed measurement model of "true" cannabis prevalence using the indirect indicators perceived availability and search engine query data and the direct measure of self-reported cannabis use showed excellent model fit. When indicators of response tendencies (over- and

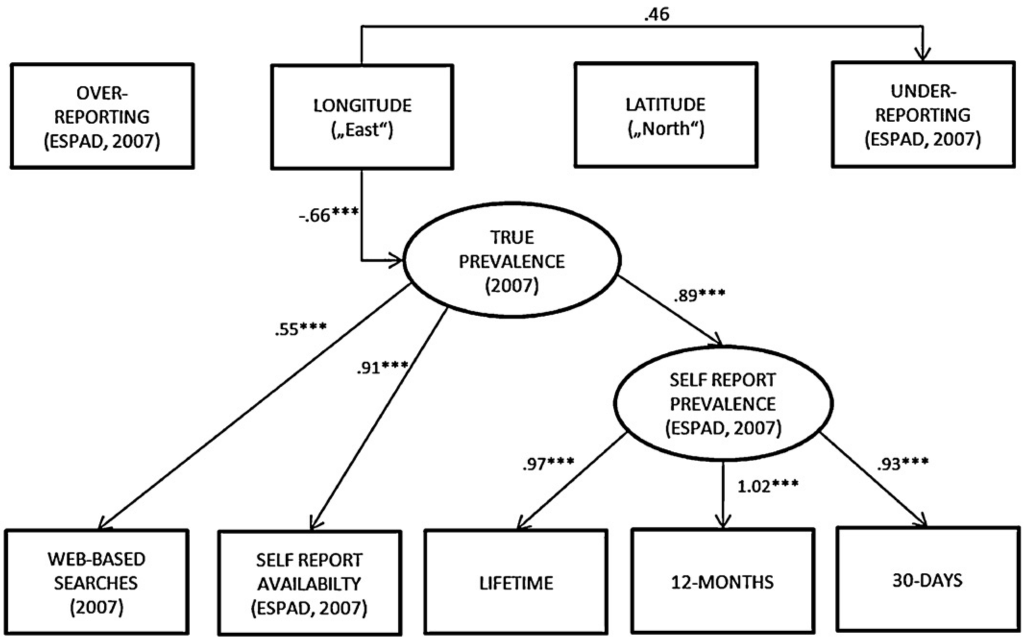

Fig. 1. Measurement and structure models based on indicators of cannabis use, response bias and geographical position on national level.

underreporting) and country-specific effects (geographical position) were included in the model, no significant effects on the cannabis use indicators were found at country level. Although general underreporting of cannabis use cannot be excluded, this indicates that prevalence estimates for adolescents obtained from studies using standardized survey methodology such as ESPAD can be considered comparable across countries on a relative level.

The analysis on provincial level in Italy revealed a significant effect of latitude on perceived availability indicating that cannabis is perceived easier accessible in Northern than in Southern Italy. Furthermore, the country-specific results point at a west-east decline in "true" cannabis prevalence and an increase in underreporting from west to east. On provincial level, latitude also affected underreporting with higher rates of underreporting in the Eastern Italian provinces. In other words, based on

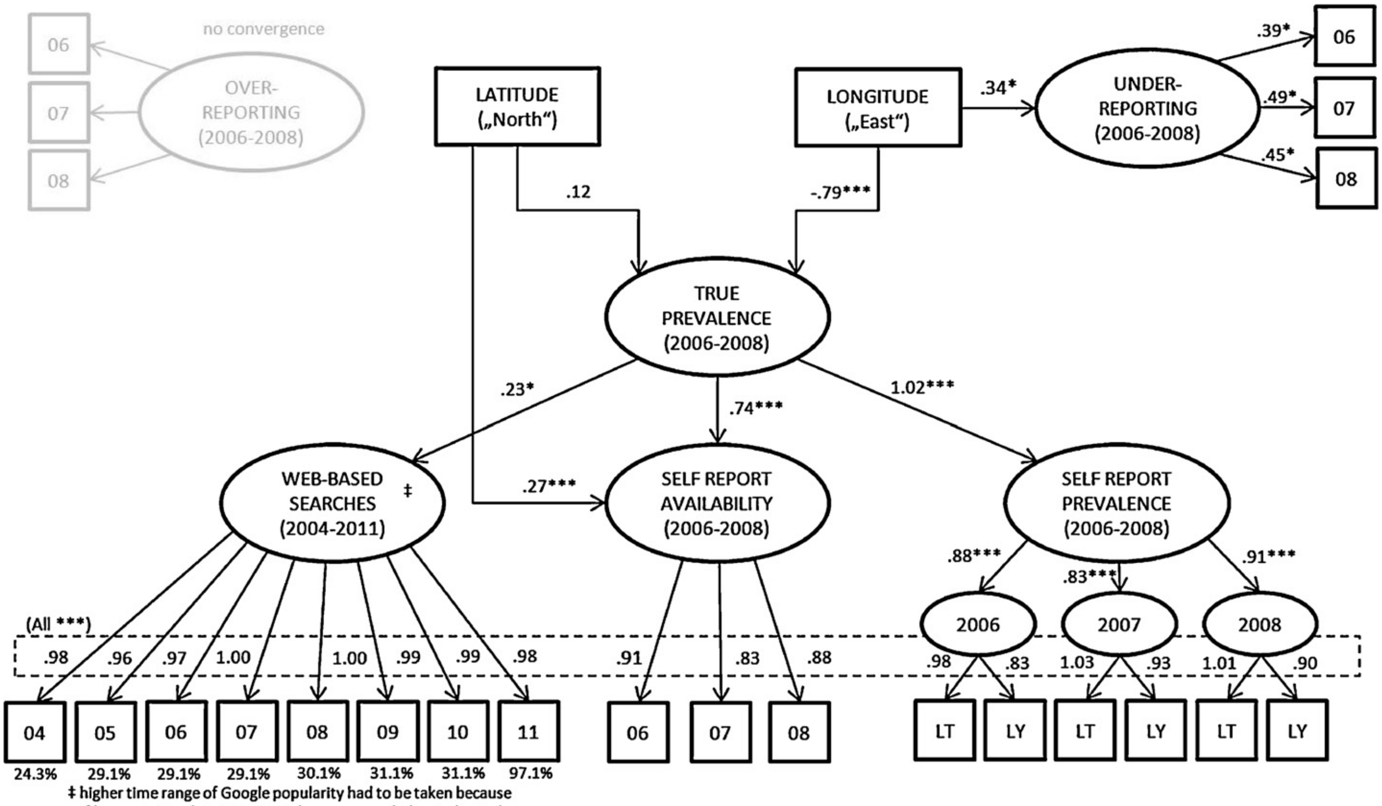

Fig. 2. Measurement and structure models based on indicators of cannabis use, response bias and geographical position on provincial level in Italy.

the measure of availability the "true" prevalence in north Italy would be overestimated indicating that responses to perceived availability seem not to be independent of province-specific influences.

The reliability of Google search volume indices (GSI) was found to be satisfactory and preceding studies indicate the same for availability (Hibell et al., 2009). Internal consistency of GSI at all levels of aggregation could be classified as 'good' according to Fisseni's scale (Fisseni, 1997), which indicates that all annual GSI largely measure one latent dimension. Reliability in terms of contextual consistency was higher on provincial than on country level. This might be due to higher lingual, demographic and cultural homogeneity across regions in a single country. The temporal stability of GSI was also higher on provincial level, but must be considered very high on both levels of the analysis. Due to the reliability-validity-dilemma (Lienert & Raatz, 1994), a well-balanced, but not perfect reliability enhances the ecological validity. This could be one reason for higher factor loading, despite lower reliability of GSI on country than on provincial level.

The validity of cannabis availability and GSI is difficult to test. However, compared to other indicators such as cannabis seizures or cannabis-related police arrests, individuals' perceptions of the likelihood to acquire cannabis and Internet behaviour may be considered largely unbiased by external conditions. Due to their non-self-relatedness, perceptions of access to cannabis seem rather independent of social norms. Furthermore, reporting the likelihood of getting access to cannabis is not disclosing an individual's drug using behaviour. Analogously, Internet behaviour may be considered unbiased by legal frames or social norms simply reflecting peoples interests, of whatever kind, in cannabis. The Internet is highly associated with anonymity in a safe environment; users and non-users alike can search for information on cannabis without danger of being identified. Although access to the Internet may vary across countries, coverage in Europe and the US is rather high suggesting that differences may be negligible (International Telecommunications Union, 2011).

The observed association between perceived availability and prevalence of cannabis use is in line with previous research confirming the role of cannabis availability as a predictor of use (Hibell and Anderson, 2008). Although the link between GSI and cannabis use was less strong, the findings corroborate the usefulness of search engine query data as an external validation source (Choi & Varian, 2009; Ginsberg et al., 2008; Pelat, Turbelin, Bar-Hen, Flahault, & Valleron, 2009; Valdivia & Monge-Corella, 2010). Considering the manifold reasons causing cannabis-related search engine queries, its function as an indicator of prevalence is not self-evident. Cannabis-related web queries could be traced back to many sources, e.g. consumers, families, therapists and scientists, and thus, the link between web-based searches and prevalence remains unclear. It may be assumed however, that the more a society is exposed to cannabis the more cannabis-related web searches will be executed. Our findings corroborate the assumption of cultural independence of these indicators. Neither individual response tendencies nor geographical position did affect perceived availability and search engine query data on country level. Consequently, cannabis use prevalence data may be considered unbiased by country-specific response biases and thus comparable across countries. In practice, these findings indicate that perceptions of access to cannabis positively correspond with cannabis prevalence. Similarly, possible external influences on search engine counts such as internet access, media awareness, or parental control of adolescents' internet use do not negatively affect the association between counts and reported prevalence. However, since the "true" prevalence is not known, comparisons can only be made on relative level, i.e. the estimates preserve the order of countries along the dimension "extent of cannabis use". In addition, the indicators "dummy drug use" and "truthful reporting of cannabis use if taken" are assumed valid measures of exaggeration and denial, respectively. Confidence in these indicators however, is primarily based on face validity. Future research needs to test these measures. Moreover, external country-specific effects were represented by geographical position. Specific measures of law enforcement or cultural norms might be used in further analyses.

Despite similar model results when external county-specific variability was controlled for by using data at regional level in a single country, there were also differences encountered. The analysis on provincial level in Italy showed a significant effect of latitude on perceived availability indicating a north-south gradient with cannabis being perceived easier accessible in Northern than in Southern Italy. Since Italian's north is rather associated with a more liberal attitude, while the south is considered more conservative with stronger family bounds and stricter social control, one may argue that these social norms differentially affected the perceptions of access to cannabis. Consequently, this finding seems to contradict the general assumption of perceived availability constituting a "culture" independent measure of cannabis use prevalence. Even more importantly, the latent model for the indicator overreporting did not converge, challenging the validity of this particular measure.

This research is not without limitations. (1) The high accuracy of the models was obtained for industrialized regions, i.e. Europe and the US. Latent modelling might not be as accurate, if other countries with lower Internet coverage or lower privacy protection standards had been included. (2) While Internet search terms such as "cannabis" and "THC" can be used across countries and cultures without needing translation, the use of related terms in different languages may negatively affect the model. (3) Using cannabis GSI over the years 2004–2011 increased the coverage of the measure in Italy but resulted in a reference period that deviates from the prevalence measures. (4) Geographical position might to some extent reflect cultural aspects, but cannot be considered a sufficient indicator of cultural differences. Finally, the fit statistics for the full models in both data sets were below the recommended cut-off points. This suggests that there is a substantial part of unexplained variance among the covariates. However, since there were no significant paths between the covariates and the indicators at country level, this does not impair the validity (invariance) of the indicators. Only at regional level, perceived availability was affected by latitude indicating non-invariance of this measure.

Apart from the important finding suggesting acceptable relative comparability of national cannabis use prevalence estimates that are based on a common survey methodology, this approach may provide additional value for epidemiologic research. Doubtlessly, cannabis-related search engine query data are a too weak indicator of cannabis prevalence in order to base prevalence estimations only on this source. However, search engine query data may be combined with other external sources such as waste water analysis or sales figures of cigarette paper or other cannabis smoking utilities. Future research needs to test estimation models without the use of self-report data in order to get valid estimates also in countries where surveys on sensitive behaviour are not valid, not possible or not welcome. Moreover, search engine query data may also be considered for estimating the prevalence of rare diseases, sexual dysfunctions, extreme political opinions or stigmatized behaviours.

Acknowledgement

Google Inc. agrees with the use of Google Search Volume Indices (GSI) for research purposes (http://www.google.com/intl/en/insights/terms.html).

Conflicts of interest

None.

References

Brener, N. D., Eaton, D. K., Kann, L., Grunbaum, J. A., Gross, L. A., Kyle, T. M., et al. (2010). The association of survey setting and mode with self-reported health risk behaviours among high school students. *Public Opinion Quarterly*, *70*, 354–374.

Brugal, M. T., Domingo-Salvani, A., Maguire, A., Cayla, J. A., Villalbi, J. A., & Hartnoll, R. (1999). A small area analysis estimating the prevalence of addiction to opioids in Barcelona. *Journal of Epidemiology and Community Health*, *53*, 488–494.

Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.

Campanelli, P. C., Dielman, T. E., & Shope, J. T. (1987). Validity of adolescents' self reports of alcohol use and misuse using a bogus pipeline procedure. *Adolescence*, *22*, 7–22.

Choi, H., & Varian, H. (2009). *Predicting the present with Google Trends*. Google Inc.

Comiskey, C. (2001). Methods for estimating prevalence of opiate use as an aid to policy and planning. *Substance Use and Misuse*, *36*, 131–150.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method*. New York: Wiley.

Fisseni, H. J. (1997). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.

Geiser, C. (2010). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag.

Gillespie, N. A., Neale, M. C., Prescott, C. A., Aggen, S. H., & Kendler, K. S. (2007). Factor and item-response analysis of DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. *Addiction*, *102*, 920–930.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2008). Detecting influenza epidemics using search engine query trends. *Nature*, *457*, 1012–1014.

Google Labs. (2010). *Google trends*. Available online: http://www.google.com/insights/search.

Harrison, L. (1997). *The validity of self-reported drug use: Improving the accuracy of survey estimates*. Rockville: National Institute of Drug Abuse.

Hibell, B., & Anderson, B. (2008). Patterns of cannabis use among students in Europe. In European Monitoring Centre for Drugs and Drug Addiction (Ed.), *A cannabis reader: Global issues and local experiences*. Lisbon: European Monitoring Centre for Drugs and Drug Addiction.

Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., et al. (2009). *The 2007 ESPAD report. Alcohol and other drug use among students in 35 European countries*. Stockholm: The Swedish Council for Information on Alcohol and Other Drugs.

Hser, Y.-I., Prendergast, M., Anglin, M. D., Chen, J. K., & Hsieh, S.-C. (1998). A regression analysis estimating the number of drug-using arrestees in 185 US cities. *American Journal of Public Health*, *88*, 487–490.

International Telecommunications Union. (2011). *The world in 2011. Facts and figures*. Retrieved from http://www.itu.int/ITU-/ict/facts/2011/material/ICTFactsFigures2011.pdf

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2008). *Monitoring the future national results on adolescent drug use: Overview of key findings, 2007 (NIH Publication No. 08-6418)*. Bethesda, MD: National Institute on Drug Abuse.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Kokkevi, A., & Stefanis, C. (1991). The epidemiology of licit and illicit substance use among high school students in Greece. *Journal of Public Health*, *81*, 48–52.

Lienert, G. A., & Raatz, U. (1994). *Testaufbau und Testanalyse*. Weinheim: Beltz.

Magura, S., & Kang, S. Y. (1995). *Validity of self-reported drug use in high risk populations: A meta-analytic review*. New York: National Development and Research Institute.

Molinaro, S., & Siciliano, V. (2009). *ESPAD-Italia. Online-Database*.

Molinaro, S., Siciliano, V., Curzio, O., Denoth, F., & Mariani, F. (2012). Concordance and consistency of answers to the self-delivered ESPAD questionnaire on use of psychoactive substances. *International Journal of Methods in Psychiatric Research* (published online).

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., & Valleron, A.-J. (2009). More diseases tracked by using Google trends. *Emerging Infectious Diseases*, *15*, 1327–1328.

Person, P. H., Retka, J. A., & Woodward, J. A. (1978). *A method for estimating heroin use prevalence*. Rockville: National Institute on Drug Abuse.

Rhodes, W. (1993). Synthetic estimation applied to the prevalence of drug use. *Journal of Drug Issues*, *23*, 297–322.

Smart, R. G., & Ogborne, A. C. (2000). Drug use and drinking among students in 36 countries. *Addictive Behaviours*, *25*, 455–460.

Smit, F., Toet, J., van Oers, H., & Wiessing, L. (2003). Estimating local and national problem drug use prevalence from demographics. *Addiction Research and Theory*, *11*, 401–413.

Valdivia, A., & Monge-Corella, S. (2010). Diseases tracked by using Google trends. *Spain. Emerging Infectious Diseases*, *16*, 168.

van Nuijs, A. L. N., Pecceu, B., Theunis, L., Dubois, N., Charlier, C., Jorens, P. G., et al. (2009). Spatial and temporal variations in the occurrence of cocaine and benzoylecgonine in waste- and surface water from Belgium and removal during wastewater treatment. *Water Research*, *43*, 1341–1349.

Wickens, T. D. (1993). Quantitative methods for estimating the size of a drug-using population. *Journal of Drug Issues*, *23*, 185–216.

Wish, E. D., Hoffman, S., & Nemes, S. (1997). *The validity of self-reports of drug use at treatment admission and at followup: Comparisons with urine analysis and hair assays*. Rockville: National Institute on Drug Abuse.

Please cite this article in press as: Steppan, M., et al. Are cannabis prevalence estimates comparable across countries and regions? A cross-cultural validation using search engine query data. *International Journal of Drug Policy* (2012), http://dx.doi.org/10.1016/j.drugpo.2012.05.002