This paper arises from the project Standardizing Measurement of Alcohol Related Troubles (SMART) which has received funding from the European Union, in the framework of the Public Health Programme Agreement Number: 2007308

Kim Bloomfield, Ann Hope and Ludwig Kraus

A REVIEW OF ALCOHOL SURVEY METHODOLOGY: TOWARDS A STANDARDISED MEASUREMENT INSTRUMENT FOR EUROPE

Final Version 13.12.10

Title: A review of alcohol survey methodology: Towards a standardised measurement instrument for Europe

Abstract

This paper provides an overview of alcohol survey methodologies currently used in order to help inform the development of a standardized comparative survey methodology on alcohol use, patterns of drinking and alcohol problems for Europe. It is part of an EU project "Standardized measurement of alcohol-related troubles" (SMART) co-funded by the European Commission DG-SANCO. Various instruments to measure alcohol consumption, risky drinking, abuse/dependence, negative social consequences and harm to third parties as well as mode of survey administration are reviewed.

The review of literature provides several insights for the SMART project on measurement issues, modes of survey administration and data analysis. The beverage-specific quantity-frequency (BSQF) measure is recommended for use across countries. A reference period of one year for alcohol consumption is considered important if one is to link associated problems with alcohol intake. In choosing an instrument for measuring abuse/dependence the key issue is to decide whether the instrument is to serve as a screening or diagnostic tool. In the case of screening, the AUDIT or the RAPS4 appears more appropriate. But if one desires to approximate a diagnostic category, then CIDI, AUDADIS or an operationalisation of ICD-10 criteria would be the better choice. In this instance, though, the instrument could be of some length. Due to increased cell phone ownership and its affect on sampling frame, response rate and bias, face-to-face interview as mode of survey administration has advantages over telephone interviews (as well as postal surveys). However, the face-to-face method is not without disadvantages.

1 Background

Alcohol is a major health determinant in the European Union (EU). It is estimated that 53 million EU adults (15%) do not drink alcohol at all and some 58 million (16%) are heavy drinkers, of whom some 23 million (6.5%) are dependent on alcohol (1). In the recent Eurobarometer alcohol survey, almost 133 million EU citizens reported risky drinking (≥50 grams of alcohol per occasion) at least once a month, representing 1 in 3 of the adult population (2). However, the prevalence of harmful drinking patterns depends to a large extent on the questions asked, definitions and methodologies used (1). Currently there is no standardized comparative survey methodology available on heavy drinking, episodic heavy drinking (binge drinking), drunkenness, alcohol abuse and dependency across the EU. Some work has been done through the European Community Health Indicators for Monitoring project (ECHIM) that is making, among others, the assessment of the comparability of all health indicators endorsed by the European Commission. ECHIM has to date proposed total alcohol consumption and hazardous alcohol consumption as risk factor indicators. Important efforts are being made within the European Health Interview Survey (EHIS) to elaborate standardised questions on alcohol consumption and to apply them in its large health survey. The EU Committee on alcohol data, indicators and definitions recommends three key indicators for alcohol consumption and harm, volume of consumption, pattern of consumption and alcohol-attributable health harm. The volume of alcohol consumption is defined as total consumption per capita (≥15 years), pattern of consumption as measured by harmful drinking (≥60 grams per occasion) at least monthly and alcohol-attributable years of life lost (premature mortality) (3).

In its 2007 work plan the European Commission called for the development of standardized comparative surveys on drinking patterns. The overall aim of the Standardized Measurement of Alcohol-Related Troubles project (SMART) is to develop a comprehensive comparative survey methodology on alcohol use, patterns of drinking and alcohol problems as well as public support for alcohol policy measures in the EU. The SMART project co-funded by the European

Commission Directorate General for Health & Consumers (DG-SANCO), was set up in 2008 with experts involved in alcohol survey methodology in 10 Member States. This paper provides an overview of the different alcohol survey methodologies currently used in survey research to help inform the development of a standardized comparative survey methodology for Europe. The review examines the various types of instruments used to measure alcohol consumption including measures of risky drinking and the various modes of survey administration. The review also focuses on three main areas of alcohol related harm: abuse/dependence, negative social consequences and harm to third parties.

2 Alcohol consumption measures

The work of Gmel & Rehm (4) and Dawson (5) provide excellent introductions when considering choice of various alcohol measures. Gmel & Rehm (4) first pose the question of the purpose of the measurement. Will information on alcohol consumption be used for: a) description (e.g., estimating exact intake levels), b) testing differences (e.g., comparing the consumption of different subgroups or populations, c) or establishing relationships (e.g., relating the level of consumption to a particular outcome)? For the SMART project, the goal is to develop standardized measurement instrumentation for alcohol survey research in Europe. Implicit in this charge is that such chosen measures will be used cross-nationally within Europe. Therefore, estimating alcohol consumption and testing differences is the highest priority for SMART when considering the choice of measures.

The main approaches to measuring alcohol consumption in survey research can be categorised in the following way: quantity-frequency measures; graduated frequency measures; and short-term recall measures. The simplest measure of alcohol consumption is one based purely on frequency. This has been the case in medical epidemiology studies (6) and in basic population health surveys. However, such a measure does not allow the calculation of volume of alcohol consumed. The *quantity-frequency* (QF) measure allows for calculation of volume since quantity consumed is recorded along with the frequency of drinking. Although this measure has a long history (e.g.,(7)),

it can be considered one of the most universal and practical instruments in alcohol survey research. Modifications or improvements have been added to it, such as a beverage-specific version which is asked of the three main beverage types. Also an additional question about binge or episodic heavy drinking is usually added to the basic measure to accommodate wider variability in drinking patterns. The QF measure may be asked of varying periods: usually anywhere from a week to a year, depending upon the nature of the research question and the main drinking pattern of the population under study.

The second main measure, the *graduated frequency* (GF), takes a different approach which implicitly has the goal of measuring volume. Numbers of drinks consumed on an occasion are grouped into graduated categories. In its most modern version the GF asks a respondent the maximum number of drinks he/she has had in a specified period. After it has been established which category of drinks is the highest for that respondent, he/she is then asked how often that has occurred. Then, the respondent is asked progressively in groups of fewer drinks how often he/she has drunk such amounts on an occasion. In such a way, the research attempts to cover the entire "universe" of one's drinking volume by starting with the maximum amount and working down by asking about smaller amounts. This approach has been in use since the early 1960s and is employed mainly in North America. In its modern version the GF is generally not beverage-specific and does not allow for calculation of a separate frequency dimension. The period of questioning is usually the past year. However, one of its advantages is that it can more easily identify occasions of heavy consumption.

The final main approach used in alcohol survey research is short-term recall measures, such as the *weekly recall* (WR). In this approach, respondents are basically asked to recall all alcohol they have consumed in a recent short period, such as the previous week. Due to the short period, it is assumed that respondents will be able to correctly recall all their consumption during that period. This approach is rather easy to administer and for respondents to understand. The main drawback is the

short measurement interval which is especially disadvantageous in capturing the full patterns of infrequent drinkers who may, for example, have just had a week of either no consumption or of unusually heavy consumption. Patterns and volume based on such an interval would misclassify these drinkers.

2.1 Comparison of approaches

An extensive literature exists on comparative assessments of the above-mentioned instruments. Key studies that are relevant and informative for the present study are listed in Table 1. First to be considered are the comparisons between the generic quantity-frequency (QF) measure and its extended version. The basic finding of both Kühlhorn & Leifman (8) and Williams et al (9) was that when more questions are asked, more volume is reported. In the study of Kühlhorn & Leifman (8) a QF measure that divided the days of the week into Monday-Thursday, Friday, Saturday and Sunday yielded higher volume than a measure of a "normal week's consumption". Williams et al (9) compared a generic QF measure with a beverage-specific QF (BSQF) as well as a BSQF with specific drink sizes. This resulted in increasingly higher volume with increasing elaboration of the generic QF. Additional studies illustrate the main principle that with increasing the number of questions on consumption, the recorded volume of alcohol will likewise increase (6;10;11).

-Table 1 -

Several studies have been conducted comparing the QF to the GF. An early study found no major differences between the two measures (12) in which a beverage-specific QF and general GF were compared to a 10-week prospective diary. Midanik (13) found that the GF captured a 38% higher volume than the generic QF. Later studies conducted outside the US have found complications with the GF. Although the measure yields higher volume than the generic QF, its use often results in over-counting of drinking frequencies, including substantial proportion of respondents whose estimated annual frequency surpassed 365 days (14-16). Thus it is not clear whether the resulting

higher volume is a valid result or artefact. In an international comparison across 10 countries Gmel et al (14) found problems with implementation of the GF and that BSQF performed better in such cases. In summary, the GF appears to work better with lighter drinkers and among those with more cognitive skills since the median of drinking is being asked as well as the task of dividing up a total of drinking days across levels of volume. If a wide variety of societies is to be compared with the same instrument, Gmel et al (14) find the BSQF to function better. Finally, several studies have examined the QF in relation to retrospective diaries or short recall methods. Here the results are the most consistent of all comparisons in that diaries yield larger volumes especially for those categorised as lighter drinkers (17-20). However it should be noted that such an instrument has the limitation of a short recall period and cannot be applied for larger intervals as can the summary measure of the QF.

2.2 Reference period

Very closely related to choice of alcohol intake measure is the consideration of the reference period for that intake (5) . As one can see, some of the measures have an inherent period "built in" to the instrument. For example, retrospective diary measures cover, by necessity, rather short periods in which drinking occasions can be accurately recalled. Thus, with these measures the trade-off is between a high degree of recall accuracy at the expense of a short reference period. A short reference period has the disadvantage of not adequately capturing the drinking pattern and volume of especially light and infrequent drinkers (5). On the other hand, longer periods require respondents to summarise or average their drinking quantity and frequency. In particular, in the QF approach underestimation of both measures is very likely and high volume consumption are excluded by definition. To some extent, the same is true for the GF, however, respondents are also actually asked to remember occasions on which they drank the most alcohol in a given period. Periods of one month to a year are also common for the GF. Another consideration is that a longer reference period is necessary if one is to also attempt to associate problems with alcohol intake.

Enough time needs to have occurred beforehand in order to argue that a risk exposure of alcohol is associated with problem consequences (5).

2.3 Modes of survey administration

Although recent research on survey modes in alcohol survey research have included considerations of web internet approaches (e.g., (21)), the main modes of survey administration (and data collection) used in alcohol surveys continue to be face-to-face interview, telephone interview and self-administered questionnaire (usually delivered via post). The lack of uniformity in mode of administration across countries in an international study can be seen as problematic since research has indicated differences in sample non-response, item non-response and respondent bias between modes. Therefore, it is recommended that the same mode of administration be used across countries in any international study.

Greenfield et al (22) point out that there are two primary ways in which mode of collection can affect the measurement of alcohol consumption and alcohol-related problems. The first relates to non-response rates and the second has to do with differential response bias (including item non-response). Non-response rate was found to be lowest in face-to-face surveys, higher in telephone surveys and the highest with mail surveys (23). However, since the early to mid-2000s a substantial decline in telephone survey response rates has been acknowledged (24;25). Earlier non-response associated with telephone surveys was attributed to lack of telephone ownership, something that is found to be more prevalent among those with lower education and lower income (26). But more recently the decline has been attributed to a general lack of willingness among the public to engage in survey research as well as to the difficulties to enumerate and contact potential respondents due to growing technology for call screening (25). The main concern associated with non-response in alcohol and drug survey research has been that higher non-response usually leads to lower survey estimates of consumption and problems because most non-responders are heavier

drinkers or drug abusers. However, later research has shown that this not necessarily the case (27-29) and that actually abstainers can also be overrepresented among non-responders (30).

Most research on differences in data collection mode has been devoted to investigating the second source of bias, namely, respondent bias or reporting error. The face-to-face interview has traditionally been seen as the optimal mode of questionnaire administration, due to the fact that not only are higher response rates associated with it, but higher item response as well. However, there could be problems with interview confidentiality and perceived social distance between interviewer and respondent so that the validity of the data may not always be as high as possible. This is especially true for asking sensitive or embarrassing questions (31). Telephone interviews have become popular and have increased in use mainly because the advantages of quality control throughout the interview process, cost efficiency and speed of data collection (32). Much of the literature on comparing administration modes concerns the comparison of telephone to face-to-face interview. The main conclusion of several studies is that, when well conducted, these two modes can lead to similar results although response rates for telephone surveys are lower (22;33-38). Moreover, the type of question asked should also be taken into consideration. Midanik et al (39) specifically examined alcohol-related harm and found that telephone interviews elicited higher rates of problems than did face-to-face interviews. They attribute this to the less confrontational format of the telephone interview.

Comparisons between postal or self-administered questionnaires and the other two modes of data collection indicate mixed results. In general, earlier research found that mail or self-administered questionnaires elicit better quality data, especially when sensitive questions such as alcohol and drug use are involved (33;40-42). This is because of the anonymous nature of questionnaire completion and the reduced influence of social desirability bias (e.g.,(43)). More recently, however, this has been shown to not always be the case in alcohol research. Rehm & Arminger (44) compared personal interviews to mail questionnaires and found higher reported alcohol

consumption in the personal interviews. But, Bongers and Van Oers (45) found no notable differences in alcohol consumption and alcohol-related problems in comparing the same modes of data collection. However, in comparison to telephone interviews, it appears that mail questionnaires do elicit higher reported consumption and alcohol problems (28;46). With the exception of Rehm & Arminger's (44) research, it can be summarized that, in general, mail questionnaires yield the same if not better data on alcohol consumption and alcohol-related problems than personal or telephone interviews.

But the drawbacks to mail questionnaires are that, in addition to possible higher sampling non-response already mentioned, they have been associated with higher item non-response. This is due to the absence of interviewers controlling the questionnaire completion process. Yet, despite higher item non-response, both Bongers & Van Oers (45) and Kraus & Augustin (46) still found higher levels of consumption and alcohol-related problems with mail questionnaires compared to face-to-face or telephone interviews. A common shortcoming of both, mail-questionnaires and telephone surveys is that a researcher cannot control if responses are given by a sampled individual or somebody else.

Telephone surveys appear to continue to be an increasingly popular survey approach. But the mode has definite challenges. Firstly, developing a sampling frame has become complicated, mainly due to the explosion in cell phone (47). This is because sampling frames for landline connections and cell phones are not distinct, but overlap, thus disturbing the assumptions needed for probability sampling (i.e., those who have higher number of telephones have higher likelihood to be sampled). Also cell phone ownership is not tied to a specific geographical location, but is "portable", making it difficult to enumerate potential sampling units within a region. Further, it is difficult to identify and eliminate cell phone numbers which are used for business. Thus, sampling continues to present problems. In addition, increased call screening makes it difficult for researchers to contact potential respondents. Researchers are developing new methods to address this problem such as

advance letters, interviewer training, increasing call attempts, targeting call times and conveying credentials of the research agency (47;48).

-Table 2 -

3 Measures of risk drinking

Terms to describe risky drinking include episodic heavy drinking, binge drinking, risky single occasion drinking, extreme drinking. Because average daily intake and consequently average consumption may not adequately reflect risks associated with certain outcomes, a measure of more intensive, concentrated consumption taking place within a short period has become recognized as a critical measure of alcohol drinking pattern (49). Such a measure has come to be defined generally as that which can increase blood alcohol concentration (BAC) to a level of intoxication within an occasion. Thus, this has been seen to be a dose of approximately 60 g of ethanol. In the US where a standard drink contains 13.6 g ethanol, the equivalent in drinks is roughly four to five drinks. In some countries sex-specific levels are set; for instance for some US researchers, 5 or more drinks are used for men and 4 or more for women. Table 3 provides an overview of defined quantities constituting risky drinking in surveys collected in the Gender, Alcohol and Culture: An International Study project (GENACIS).

-Table 3 -

4 Drinking limits

No systematic review of international drinking limits has been conducted. In fact, there seems to be little literature on the matter in general, perhaps reflecting the fact that such guidelines are often finally implemented through political policy decisions and not by researchers themselves. It appears that early recommendations for drinking limits, such as those in the 1960s were based more on research for risk of organic damage (e.g., liver cirrhosis) than for risk of dependence or social

harm (50). Accordingly, recommended limits were rather high, ranging from not more than 120 g per day to not more than 160 g pure alcohol per day (50). By the 1970s these limits were lowered considerably and became integrated into health education campaigns with a focus shift to the risk of alcohol dependence rather than organic damage. Dawson (51) cites that by 2000 at least a dozen countries had established their own recommended limits. These were mostly English speaking countries such as Canada, the US, Australia, New Zealand and the United Kingdom (51). Although recommended levels have declined since the 1960s, variations still exist across sex and country, as well as in the period of consumption: day or week.

Recent work undertaken for the UK Department of Health inquired from all members of the EU Commission's committee of national counterparts (government representatives) and compiled current drinking limits (Table 4). Drinking guidelines from outside of Europe were accessed from the website of the International Centre for Alcohol Policies (ICAP). ICAP is an international organisation funded by the global alcohol industry.

-Table 4 -

In Europe, of the 25 countries that provided information on drinking limits, 20 countries issue guidelines on 'safe', 'responsible', 'low risk consumption' determined by a government/official public health body. Fourteen countries give daily drinking limits, three countries give limits per drinking occasion and ten countries give weekly limits. As can be seen in Table 4 the vast majority of the countries differentiate their drinking limits by gender. Region also appears to also make a difference with higher drinking limits in some European countries, in particular for weekly limits. This obviously reflects social and cultural differences in what is perceived as risks of alcohol drinking. In countries outside of Europe, there is less variation in drinking limits, in particular between North America and Australasia.

However, in order to empirically support such levels, research, at least in the United States, has attempted to evaluate the association between drinking levels and levels of risk (49;51;52). Dawson (51) examined four different versions of American drinking limits. Two of these came from the US Department of Agriculture and another two came from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). They were tested with respect to sensitivity, specificity and positive predictive value for outcomes of alcohol dependence, impaired driving, liver disease, peptic ulcer and hypertension. Results showed that the NIAAA recommendations for both daily and weekly drinking limits possessed the optimal combination of validity indicators. These limits were: "for males, intake not to exceed 14 standard drinks per week AND not to exceed 4 standard drinks on any day; for females, intake not to exceed 7 standard drinks per week AND not to exceed 3 standard drinks on any day" ((51), p. 1821). A standard drink contains 14 g ethanol. Further Dawson et al (52) constructed risk curves to determine the relationship of consumption exceeding recommended US drinking limits ($\geq$ 5 drinks for men and $\geq$4 drinks for women) with the incidence of alcohol dependence and/or abuse. A clear, almost linear relationship between number of days exceeding the limits and the incidence of dependence with abuse could be seen when graphed. Thus, at least based on American research, drinking limits have a justification based on empirical evidence.

5 Alcohol-related problems: abuse/dependence, social consequences, third party harm

The universe of alcohol-related problems and consequences is broad and largely unmapped as an entire entity. This review focuses on three main areas of alcohol-related harm: abuse/dependence, social consequences and harm to third parties as measured in population surveys. It excludes types of alcohol-related harm that can be measured with mortality and morbidity statistics as well as police reports and economic models (53-55).

5.1 Alcohol dependence and abuse

Measures of alcohol dependence and abuse mainly have origins in psychiatric, epidemiologic and public health research. Some work has existed in the sociological survey research, but in the past 20 years this territory has become increasingly occupied by larger health authorities (e.g., WHO, NIAAA) to systematically develop valid and somewhat universal instruments. The development of diagnostic criteria and corresponding instruments to operationalise these criteria has a complicated and intertwined history that is too extensive to be related here. The reader is referred to Hasin (56) in order to gain some background.

Three main measures presently tend to dominate population survey research on alcohol: the *Composite International Diagnostic Interview Substance Abuse Module* (CIDI-SAM (57)), the *Alcohol Use Disorder and Associated Disabilities Interview Schedule* (AUDADIS (58)), and the *Alcohol Use Disorders Identification Test* (AUDIT (59)). All were developed roughly around the same time; however, only the *Composite International Diagnostic Interview* (CIDI) and the AUDIT have been meant for international, comparative use. For whatever reason, the AUDADIS appears to have been used almost solely in the United States. Shorter instruments exist such as the *CAGE* for alcohol dependence (60), the *Brief Michigan Alcoholism Screening Test* (B-MAST (61)), and the *TWEAK* for alcoholism/heavy drinking (62). Although brief, these focus to some degree on special subsets of the general population. For example, the B-MAST detects chronic alcoholics and the TWEAK was meant for pregnant women. Furthermore, their use, too, has appeared to remain largely in North America and thus may reflect their cultural specificity. Another brief screening tool, the *Rapid Alcohol Problem Screen* (RAPS4 (63)), was initially developed to quickly identify problem drinkers among emergency room patients and has subsequently been used in general population survey research (64). This tool has been applied to international research and has shown promising results.

This review will thus focus on the CIDI and the AUDIT as potential survey measurement instruments for detecting alcohol abuse and dependence in a European context. It also examines the

RAPS as a brief screening tool that could find wider use in Europe. From the outset, however, it is important to note that the three instruments have different goals in their measurement capabilities: the CIDI sets out to replicate and operationalise DSM-IV and ICD-10 criteria for alcohol abuse and dependence whereas the AUDIT was meant at the outset to be a brief instrument to screen for "persons with hazardous or harmful alcohol consumption before dependence and serious harm have occurred" ((59), p. 791). Further, the RAPS was intended to be an even briefer tool aimed at indicating dependence for screening in the emergency room setting. Despite their varying intentions, all three instruments have been used in survey and out-patient research to detect problem and/or dependent drinkers. The following table 5 (adapted from Hasin, (56)) gives an overview of the dimensions covered by each instrument.

-Table 5 -

A review of the item coverage for the instruments makes clear that the CIDI comprehensively covers both abuse/harmful use and dependence symptoms as determined by DSM-IV and ICD-10 criteria, while the AUDIT can be seen as focusing on the beginning signs of trouble. The RAPS4 (and earlier original RAPS), deriving from the brief instrument tradition, was developed to correlate highly with both DSM-IV and ICD-10 criteria (63;65) for dependence. The RAPS4 shows most similarity to items in the older CAGE screener and with some of the AUDIT items.

The aim of the CIDI (along with related but more extensive instruments such as the SCAN (66)) was to provide an instrument that accurately reflects the diagnostic criteria for alcohol abuse and dependence so that epidemiologists in particular may identify cases on an international basis within large general population mental health surveys (57). The aim of the AUDIT is similar, but its origins derive from testing 150 different items on primary health care patients in six countries (59). Its background lies more in previous screening instruments and not in reflecting DSM and ICD criteria per se. The CIDI is a lengthy instrument, covering a long list of psychiatric diagnoses

including dependence to various types of drugs (57). A major examination of the questionnaire's validity and reliability was published in 1997 (57) in which it was reported that the CIDI (as well as SCAN and AUDADIS) was tested in 10 countries. Kappa statistics were calculated to assess test-retest reliability as well as concurrent validity. The CIDI alcohol component demonstrated acceptable reliability for both DSM-IV and ICD-10 criteria as well as concurrent agreement with the SCAN and AUDADIS. Hasin (56) has reported on later assessments as acceptable. However the main obstacle in using the CIDI is its length. A group in Germany has developed the Munich composite international diagnostic interview (M-CIDI (67)). The questions assessing the diagnostic criteria are identical to those in the original CIDI but in other parts, the M-CIDI shows several improvements compared to the CIDI. For example, the screening section covering quantity and frequency measures was extended by adding circumstances of first use, effects, and problems associated with the initiation of drug use. Information about onset and recent substance use were separated into a series of questions. Moreover, more visual aids were introduced in order to simplify the identification of substances and to help in the presentation of complex questions. This instrument also demonstrated acceptable kappa values for reliability and validity. A novel feature is that the instrument was tested on a community sample and not among patients. In the last decade, CIDI has been applied on general population samples in many countries throughout the world.

Since its development and publication in 1993 (59), the AUDIT has undergone extensive testing and review (e.g., (68-70)). Interestingly, over the years the use of the AUDIT or parts of it have extended beyond primary health care settings to general population surveys (e.g., (71-79)). As mentioned, the motivation for developing the AUDIT was to have a screening tool for hazardous alcohol consumption and not to diagnose abuse or dependence via strict criteria. As such, it is a short, stand-alone questionnaire of 10 items.

Allen et al (68) and Reinert & Allen (70;80) have regularly reviewed the literature on the AUDIT's performance. These reviews maintain that the instrument has sound psychometric properties,

acceptable reliability and validity, is easy to use, relatively free of cultural bias and is obtainable without copyright fees. As of 2007 Reinert & Allen (80) report that a larger evidence base exists for the English version of the questionnaire; further psychometric work is urged on other language versions and gender-specific cut-offs should be considered. Additionally, a recent meta-analysis was conducted on 19 studies. Berner et al (69) report that results varied due to the variety of patient settings. Although sensitivities and specificities were all in the acceptable range, they urge its use to be restricted to primary care, inpatient and elderly patient settings.

Selin (78) undertook a comprehensive analysis of the instrument's validity with respect to its potential to measure various components of problematic drinking: high-volume consumption, social problems, health problems and dependence (as tested against ICD-10 criteria). From her review of previous research it is unclear what the definitive factor structure of the instrument is. Some researchers have found a two-factor structure for consumption and problems (81) while others have found a three-factor solution involving consumption, harmful use and dependence (82). Still another group has found four factors: one for each consumption item and another for problems (73). Interestingly, in reviewing the previous work neither Selin or the other authors describe any factor as representing dependence as such. Selin's work did not involve finding factors but testing the instrument against the four mentioned components. In all she determined that the AUDIT performed well against all four criteria with a cut-off of 8+ for the whole instrument (i.e. AUDIT-10). In one of the few studies to examine the AUDIT in general population surveys, Knibbe et al (76) reported on the reliability of individual items in nine European countries with a special focus on gender. The conclusions were that among the consumption items, frequency of drinking decreased Cronbach's alpha correlation in almost all countries. Of the consequence items it appeared that some questions (esp. injury and concern of others) also reduced the internal consistency leading the authors to conclude that these have varying meaning across the study countries.

As mentioned, the RAPS4 (and its antecedent, RAPS) was developed as a very brief screener for alcohol dependence in emergency room settings (63;65). The instrument has shown good psychometric properties (83) and has been successfully used in a series of general population emergency room studies in various countries such as Poland (84), Argentina (85) and Mexico (86). It also performed well in the US general population survey for alcohol dependency but less so for alcohol abuse (64). However when quantity and frequency (QF) questions (drinking five or more drinks on at least one occasion during the last year and drinking as often as once a month during the last year) were added to the RAPS4, sensitivity for alcohol abuse was improved. The instrument has been the measurement basis in large international studies encompassing countries from North and South America, Europe, Asia and Africa (e.g., (83)). In its international application, the RAPS4 has demonstrated higher sensitivity and specificity in relation to a measure of tolerance (proxy for dependence) in those countries with more problematic drinking cultures (i.e., where alcohol is not well integrated into the culture) (83). In concluding this section, a point to keep in mind when choosing an instrument for measuring "abuse/dependence" would be whether the chosen instrument is to serve as a screening tool or as a diagnostic tool. In the case of screening, the AUDIT or the RAPS4 appear more appropriate. But if one desires to approximate a diagnostic category, then CIDI, AUDADIS or an operationalisation of ICD-10 criteria would be the better choice. In this instance, though, the instrument could be of some length.

5.2 Social consequences

There is a tradition to view separately those alcohol-related problems which concern abuse and dependence (formerly alcoholism), and those which concern social problems that the drinker experiences due to his or her consumption (e.g., problems with spouse/relatives/friends, with job, with police, with finances, with aggression). These have been labelled in various ways; for example, as "intrinsic" and "extrinsic" (71), or "internal" and "external" (87), "troubles due to drinking" and "preoccupation with alcohol" and "troubles due to drinking" (88), "dependence" and "consequences" (89) and "personal consequences" and "social consequences" (90). All

dichotomies reflect the intra-personal and inter-personal dimensions of the problems. However, at least two studies have shown that when psychometrically analysing such items for evidence of their dimensionality, researchers tend to find only one main component for all items (79;91). Thus from a statistical standpoint, all alcohol-related problems appear to cohere to one main dimension even though researchers would like to separate and categorise them on theoretical or conceptual grounds.

In the case of this review the former approach of categorisation has been used to organise this section and thus its topic now concerns what can be considered those external or extrinsic problems which deal with difficulties in the social realm. As Room (92) has remarked there is little consensus on standard instruments for measuring the social consequences of one's drinking; now more commonly called "social harm". Room defines social harm as "perceived mis-performance or failure to perform in major social roles – as a family member, as a worker, as a friend or neighbour, or in terms of public demeanour." ((92), p. 94). Further, he remarks that over the last half-century, the number items measuring these areas has varied yet the number of areas has remained stable. In general these areas for which questions have been asked over the last four decades include problems with: spouse, relatives, neighbours and friends, job, police, aggression and belligerence, finances. In addition to the areas, there are levels of "intensity" at which these problems can be measured. Room (92), p. 106) observes the following levels, starting from the most intense:

(1) concrete actions reported by others in response to respondent's drinking,

(2) indications of verbal responses to or attempts to control respondent's drinking,

(3) global attribution by the respondent of harm in a life area,

(4) items describing the respondent's behaviour while drinking as problematic

According to Room's categorisations, Table 6 illustrates with selected items how questions in the various life areas at different levels of intensity manifest themselves. Note that some of the

questions are hypothetical and are not actually survey items that have been empirically or psychometrically tested.

-Table 6 -

5.3 Alcohol attribution

Another characteristic of social harm questions which should be briefly mentioned is the debate regarding attribution. It has been stated that only in alcohol epidemiology is the measurement of the risk factor (alcohol) already implicitly associated with the outcome (92;93). That is, most questions regarding social harm ask if the harm has occurred because of the presence of alcohol. This is problematic for proper epidemiological analyses where the risk factor and outcome should be measured as two separate entities. Most recently, Gmel and colleagues (94) found that self-attributed alcohol problems were underreported if compared to non-attributed problems using alcohol-attributable fractions (AAF). Despite the debate that has occurred over several years, no consensus to delete the alcohol-attribution from alcohol-related harm questions has been reached. Thus, the topic is not handled further here, but readers should be aware of the issue with respect to potential analyses and assignation of attribution or association. A further consideration is the time sequence of alcohol consumption (as the risk factor) and the occurrence of alcohol-related problems (as the outcome) (91). Alcohol consumption should, of course, precede the occurrence of problems. However, many surveys cover the same period for both variables; i.e., usually one year (5). Any period shorter would be too brief to assess sufficiently the presence of problems.

5.4 Third party harm

A relatively new area of alcohol-related problems in alcohol survey research is third party harm (92;95). This area of problems, measured at the individual level, has extremely little in the way of systematic review or documentation. The earliest item-list on third party harm in alcohol survey research appears to come from a Canadian study in 1989 (96). Around the same time a Nordic

alcohol survey contained items on social harm from others (97). From these two surveys, items are summarised in Table 8. Rossow & Hauge (95) have analysed the Nordic survey items for Norway to examine the extent of harm by others. They report that the more severe types of harm are experienced by few (ca. 3-5%) but the less severe types (e.g., kept awake at night) have a considerable prevalence; i.e., ca. 20% report that they have been kept awake at least once in the past 12 months. They also conclude that those who experience the most third party harm also are heavy drinkers, and women experience more of this kind of harm than do men (95).

The problem items considered so far are those which would be included in a general battery for a national alcohol survey. More extensive lists that concentrate on a specific dimension of third party problems such those experienced by family members (e.g., (98)) or as intimate partner violence (99) are few and little, as mentioned, is to be found in the literature.

6 Conclusions

The review of literature provides several insights for the SMART project on measurement issues, modes of survey administration and data analysis. A basic finding for the alcohol consumption measure is, when more questions are asked more volume is reported. The beverage specific quantity-frequency (BSQF) is recommended when a wide variety of societies are to be compared using the same instrument and has already been used in some Europe countries. A reference period of one year for alcohol consumption is considered important if one is to link associated problems with alcohol intake. The review focuses on three main areas for measuring alcohol related problems – abuse/dependence, social consequences and third party harm. Social consequences measures are sometimes divided into internal and external troubles, yet analyses tend to show one main dimension. The key review finding in choosing an instrument for measuring abuse/dependence is, to decide whether the chosen instrument is to serve as a screening tool or as a diagnostic tool. In the case of screening for dependence, the AUDIT or the RAPS4 appear to be appropriate instruments. But if one desires to approximate a diagnostic category, then CIDI,

AUDADIS or an operationalisation of ICD-10 criteria would be the better choice. Measuring third party harm is a relatively new area and what is known to date is that less severe types of harms are more common in comparison to more severe type of harms to others. Face-to-face interview as mode of survey administration has advantages over telephone and mail methods, due to growing cell phone ownership and its affect on sampling frame, low response rate and higher item non response. However, the face to face method is not without its disadvantages, in particular its high cost.

Table 1: Overview of key studies comparing alcohol intake measures.

Comparison	Study	Country	Results/conclusions
Generic QF vs. extended QF	Kühlhorn & Leifman, 1993	Sweden	Compared weekday-specific QF with generic QF. Day-specific yielded higher volumes
	Williams et al, 1994	USA	Compared QF and beverage-specific QF, BSQF gave higher volume estimates
GF vs. QF	Hilton, 1989	USA	No sig. differences found between GF and BSQF
	Midanik, 1994	USA	Higher volume for GF than generic QF
	Poikilainen et al, 2002	Finland	Higher volume for GF than generic QF but over-counted annual frequency
	Graham et al, 2004	Canada	GF yielded higher volume but over-counted frequency
	Gmel et al, 2006	International	BSQF performed better in international comparison. GF gave higher frequencies. Implementation problems
QF vs. retrospective diaries	Redman et al, 1987	Australia	Diaries yield more at lower volumes
	Werch, 1989	USA	Diaries yield higher frequencies and volume for 7 day recall
	Webb et al, 1990	Australia	Diaries yield more at lower volumes for 7 day recall
	Shakeshaft et al, 1999	Australia	Diaries yield higher volume for 7 day recall
Reviews	Rehm, 1998	Canada	"The more specific the questions asked about alcohol consumption, the more volume results." Does not prefer GF or QF
	Feunekes et al, 1999	Netherlands	Favor a BSQF if diary not possible
	Dawson, 2003	USA	Considers GF and QF as main measures
	Gmel & Rehm, 2004	Switzerland	Comprehensive review: conclude that "the selection of best measure will depend on purpose of the study"
	Greenfield & Kerr, 2008	USA	Made no recommendation on type of measure but stated that "GF is an efficient measure in countries where individuals drink smaller quantities per occasion"

Table 2: Summary of advantages and disadvantages of various survey administration modes.

Administration mode	Advantages	Disadvantages	Comments
Face to Face interview	• High response rates • Low item non-response	• Social desirability bias • Expensive	• Still the best but expensive and time-consuming
Telephone interview	• Low item non-response • Reduced social desirability bias • Cheaper	• Low response rates • Difficult to ascertain sampling frame	• Can be effective but requires great effort to achieve good sampling
Self-administration with postal or hand delivery	• Moderate response rates • Reduced social desirability • Cheaper	• High item non-response	• Appears to still work well enough

Table 3: Overview of risk drinking definitions from the GENACIS project (volume per drinking occasion).

Country	**Risk drinking measure definition: number of standard drinks (ethanol in grams)**	
	standard drinks /glasses	**Grams (approximately)**
Switzerland	8+	80 g for men and women
Germany	5+	70 g for men and women
Israel	5+	60 g
Mexico	5+	65 g
Sweden	6+	72 g
Finland	6+ / 5+ from GF measure	60 g
Netherlands	6+	60 g
Czech Republic	5+glasses	90 g
Hungary	Sum of frequencies drinking 3-5 and 6+ drinks	60 g
Brazil	5+	60 g
Iceland	5+	65 g
Denmark	6+	72 g
Sri Lanka	5+	60 g
Nigeria	5+	60 g
Kazakhstan	5+	60 g
Argentina	5+	60 g
Canada	5+	68 g
USA	5+	60 - 70g
Uganda	5+	60 g
Japan	Sum of frequencies for 6-9 and 10+ units	72 g
Costa Rica	5+	60 g
India	5+	50 g
Australia	6+	60 g
Ireland	6+	60 g
Uruguay	5+	60 g
Belize	5+	50 g
Peru	5+	60 g
Australia	6+	60 g
New Zealand	6+	60 g

Source: (100)

Table 4: Overview of drinking limits in European countries and in countries outside of Europe.

Country	Men		Women		Standard drink/unit
	g/day	g/week	g/day	g/week	g
Europe					
Austria	24		16		10-20
Bulgaria	16-48	224	8-32	112	8-10
Czech Republic	20		20		na
Denmark	60/po	252 (21 units)	60/po	168 (14 units)	12
Estonia	20		10		10
Finland	84/po	288	60/p.o.	192	12
France	30		20		10
Germany	24		12		14
Ireland		210 (21 SD)		140 (14 SD)	10
Italy	40		20		12
Latvia		220		120	10
Malta		168-210		112-140	8-10
Netherlands	20-30		10-20		10
Poland	40		20		10
Slovakia	20		20		na
Slovenia	20	140	10	70	10
Spain	40/day 60/po	280	20-25/day 40/po	170	10
Sweden		168		108	12
Switzerland	20		20		10-13
UK	24-32			16-24	8
Australasia					
Australia	20		20		10
Japan	1-2 units				19.75
New Zealand	30	210	20	140	10
North America					
Canada	27.2	190	27.2	121.5	13.6
US (DHHS)	14.28	196	14	98	14
US(NIAAA)	56	196	42	98	14
US(AHA)	28		14		14
Africa					
South Africa		252		168	

Notes: po: per occasion; na: not available

Sources: For European countries the list was compiled as part of work undertaken for the UK Department of Health (101). The information was supplied by representatives from National Governments of the EU Committee for National Alcohol Policy and Action; countries outside Europe (102).

Table 5: Overview of diagnostic criteria with CIDI, AUDIT and RAPS4 instruments.

DSM-IV dependence criteria (103)	ICD-10 dependence criteria (104)	CIDI	AUDIT	RAPS4
Withdrawal	Withdrawal	√		
Tolerance	Tolerance	√	√	
Drinking more or longer than intended	Drinking larger amounts or over a longer period than intended / Persistent desire to cut down or control alcohol use	√		
Persistent desire or unsuccessful efforts to cut down		√	√	
Important activities given up or reduced	Important alternative pleasures or interests reduced or given up / Great deal of time spent obtain or take alcohol or getting over its effects	√		
Great deal of time spent on drinking or getting over its effects		√		
Continued drinking despite knowledge of serious problems	Persistent use despite evidence of harmful consequences	√		
	Strong desire/compulsion to take the substance	√		
DSM-IV substance abuse criteria				
Inability to fulfil roles		√	√	√
Hazardous/dangerous use		√	√	
Recurrent legal problems		√		
Continued use despite social interpersonal problems		√		
ICD-10 harmful use criteria				
Pattern of use causing damage to health. The damage may be physical (e.g. liver cirrhosis) or mental (e.g. episodes of depressive disorder secondary to heavy consumption of alcohol).				
Other criteria				
Remorse (CAGE)			√	√
Blackout			√	√
Morning drinking (CAGE)				√
Been told to cut down (quasi-CAGE, TWEAK)			√	

Table 6: Summary of life areas for social consequences.

	Level 1	**Level 2**	**Level 3**	**Level 4**
Life area	**Concrete action**	**Verbal responses, attempts to control**	**Global attribution**	**R's behaviour while drinking seen as problem**
Spouse	My wife left me because of my drinking	My wife threatened to leave me because of my drinking	My drinking has had a harmful effect on my marriage	I almost had sex with another woman at my wife's birthday party
Relatives	A relative does not speak to me anymore because of my drinking	A relative has indicated I should cut down on my drinking	My drinking has had a harmful affect on my relatives	I offended my mother in-law after having a couple of drinks
Friends and neighbours	A friendship ended because of my drinking	A friend indicated I should cut down	My drinking has had a harmful effect on a friendship	I have said harsh or cruel things to someone/friend while drinking
Job	I have lost a job because of my drinking	People at work have indicated I should cut down	My drinking has hurt my changes for promotion or better jobs	I had to take a day of sick leave after drinking
Police	I have been arrested because of my drinking	A policeman questioned or warned me about my drinking	I have had trouble with the law because of my drinking	By chance, they failed to stop me for drunk driving being occupied with another driver
Aggression/belligerence	I have gotten into a fight while drinking; I have felt aggressive or cross while drinking	A bar tender warned me to call the police next time I misbehave	I have had trouble with violence due to my drinking	I almost hit my baby after a few drinks
Finances	I had to sell my jewellery to pay my drinking debts	Co-owner of our business told me to reduce drinking	My drinking has had a harmful effect on my finances	I drank away money intended for new tv set

Table 7: Listing of existing batteries of third party harm from alcohol surveys.

Canadian survey(96) Have you ever	**Problem area**
Been insulted or humiliated by someone who had been drinking?	Insult
Had serious agreements or quarrels as a result of someone else's drinking?	Aggression
Had friendships break up as a result of someone else's drinking?	Friendship
Had family problems or marriage difficulties due to some else's drinking?	Family/marriage
Been a passenger with a driver who had too much to drink?	Driving
Been in a motor vehicle accident because of someone else's drinking?	Driving
Had your property vandalised by someone who had been drinking?	Vandalism
Been pushed, hit or assaulted by someone who had been drinking?	Violence
Been disturbed by loud parties or the behaviour of people drinking?	Disturbance
Had financial trouble because of someone else's drinking?	Finances
Nordic survey (97) Has it ever happened that	
You have been harassed or bothered by intoxicated people on the street or in some other public place?	Insult public
You have been harassed or bothered by intoxicated people at a party or some other private setting?	Insult private
An intoxicated person has harmed you physically?	Violence
An intoxicated person has ruined your clothes or other belongings?	Vandalism
You have been called names or otherwise insulted by intoxicated people?	Insult
You have been afraid of intoxicated people you encountered on the street?	Fear
You have been kept awake at night by drunken noises?	Disturbance

References

(1) Anderson P, Baumberg B. Alcohol in Europe: A Public Health Perspective. London: Institute of Alcohol Studies; 2006.

(2) TNS Opinion & Social. EU citizens' attitudes towards alcohol. Special Eurobarometer 72.3. Brussels: European Commission; 2010.

(3) Committee on Data Collection IaD. Summary Report. http://ec europa eu/health/indicators/events/alcohol_committee_mtg_2_2009-09_minutes_final pdf 2009 September 1

(4) Gmel G, Rehm J. Measuring alcohol consumption. Contemporary Drug Problems 2004;31(3):467-540.

(5) Dawson D. Methodological issues in measuring alcohol use. Alcohol Research & Health 2003;27:18-29.

(6) Rehm J. Measuring quantity, frequency and volume of drinking. Alcohol Clin Exp Res 1998;22:4S-14S.

(7) Straus R, Bacon S. Drinking in College. New Haven: Yale University Press; 1953.

(8) Kühlhorn E, Leifman H. Alcohol surveys with high and low coverage rate: A comparative analysis of survey strategies in the alcohol field. Journal of Studies on Alcohol 1993;54:542-54.

(9) Williams G, Proudfit A, Quinn E, Campbell K. Variations in quantiy-frequency measures of alcohol consumption from a general population survey. Addiction 1994;89:413-20.

(10) Dawson D. Volume of ethanol consumption: Effects of different approaches to measurement. Journal of Studies on Alcohol 1998;59:191-7.

(11) Knibbe R, Bloomfield K. Alcohol consumption estimates in surveys in Europe: Comparability and sensitivity for gender differences. Substance Abuse 2001;22:23-38.

(12) Hilton M. A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. British Journal of Addiction 1989;84:1085-92.

(13) Midanik L. Comparing usual quantity/frequency and graduated frequency scales to assess yearly alcohol consumption: Results from the 1990 US National Alcohol Survey. Addiction 1994;89:407-12.

(14) Gmel G, Graham K, Kuendig H, Kuntsche S. Measuring alcohol consumption - should the 'graduated frequency' approach become the norm in survey research? Addiction 2006;101:16-30.

(15) Graham K, Demers A, Rehm J, Gmel G. Problems with the graduated frequency approach to measuring alcohol consumption: Results from a pilot study in Toronto, Canada. Alcohol and Alcoholism 2004;39:455-62.

(16) Poikolainen K, Podkletnova I, Alho H. Accuracy of quantity-frequency and graduated frequency questionnaires in measuring alcohol intake: Comparison with daily diary and commonly used laboratory markers. Alcohol and Alcoholism 2002;37:573-6.

(17) Redman S, Sanson-Fisher R, Wilkinson C, Fahey P, Gibberd R. Agreement between two measures of alcohol consumption. Journal of Studies on Alcohol 1987;48:104-8.

(18) Shakeshaft A, Bowman J, Sanson-Fisher R. A comparison of two retrospective measures of weekly alcohol consumption: Diary and quantity/frequency index. Alcohol & Alcoholism 1999;34:636-45.

(19) Webb G, Redman S, Sanson-Fisher R, Gibberd R. Comparison of a quantity-frequency method and a diary method of measuring alcohol consumption. Journal of Studies on Alcohol 1990;51:271-7.

(20) Werch C. Quantity-frequency and diary measures of alcohol consumption for elderly drinkers. International Journal of the Addictions 1989;24:859-65.

(21) Link M, Mokdad A. Effects of survey mode on self-reports of adult alcohol consumption: a comparison of mail, web and telephone approaches. Journal of Studies on Alcohol 2005;66:239-45.

(22) Greenfield TK, Midanik LT, Rogers JD. Effects of telephone versus fact-to-face interview modes on reports of alcohol consumption. Addiction 2000;95:277-84.

(23) Hox J, De Leeuw ED. A comparison of non-response in mail, telephone and face-to-face surveys. Quality and Quantity 1994;28:329-44.

(24) Curtin R, Presser S, Singer E. Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly 2005;69(1):87-98.

(25) Keeter S, Kennedy C, Dimock M, Best J, Craighill P. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey
21. Public Opinion Quarterly 2006;70(5):759-79.

(26) Aquilinio WS, LoSciuto LA. Effects of interview mode on self-reported drug use. Public Opinion Quarterly 1990;54:362-95.

(27) Caspar R. Survey Measurement of Drug Use. In: Turner CF, Lessler JT, Gfroerer JC, editors. Follow-up of nonrespondents in 1990.Rockville: National Institute on Drug Abuse; 1992. p. 155-73.

(28) Gmel G. The effect of mode of data collection and of non-response on reported alcohol consumption: a split-sample study in Switzerland. Addiction 2000;95:123-34.

(29) Lemmens PHHM, Tan ES, Knibbe RA. Bias due to non-response in a Dutch survey on alcohol consumption. British Journal of Addiction 1988;83:1069-77.

(30) Lahaut V, Jansen H, van de Mheen D, Garretsen H. Non-response bias in a sample survey on alcohol consumption
27. Alcohol and Alcoholism 2002;37(3):256-60.

(31) Bradburn NM. Response effects. In: Rossi P, Wright J, Anderson A, editors. Handbook of Survey Research.New York: Wiley; 1983. p. 331-40.

(32) Lavrakas P.J. Telephone Survey Methods: Sampling, Selection and Supervision. Newbury Park, CA: Sage; 1993.

(33) De Leeuw ED. Data Quality in Mail, Telephone, and Face-to-Face Surveys. Amsterdam: T.T.- Publikaties; 1992.

(34) De Leeuw ED, Collins M. Data collection methods and survey quality: an overview. In: Lyberg L, Biemer P, Collins Mea, editors. Survey Measurement and Process Quality.New York: John Wiley & Sons, Inc.; 1997. p. 199-220.

(35) Groves RM. Survey Errors and Survey Costs. New York: John Wiley & Sons, Inc.; 1989.

(36) Johnson TP, Hougland JG, Clayton RR. Obtaining reports on sensitive behavior: a comparison of substance use reports from telephone and face-to-face interviews. Social Science Quarterly 1989;70:174-83.

(37) Mangione TW, Hingson R, Barrett J. Collecting sensitive data: a comparison of three survey strategies. Sociological Methods and Research 1982;10:337-46.

(38) Sykes W, Collins M. Effects of mode of interview: experiments in the UK. In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Micholls II WL, Waksberg J, editors. Telephone Survey Methodology.New York: John Wiley & Sons, Inc.; 1988. p. 301-20.

(39) Midanik LT, Greenfield TK, Rogers JD. Reports of alcohol-related harm: Telephone vs. Face-to-face interviews. Journal of Studies on Alcohol 2001;62:74-8.

(40) De Leeuw ED, Nichols II W. Technological innovations in data collection: acceptance, data quality and costs. Sociological Research Online 1996;1:1-17.

(41) Schwartz N, Strack R, Hippler H-J, Bishop G. The impact of administration mode on response effects in survey measurement. Applied Cognitive Psychology 1991;5:193-212.

(42) Sudman S, Bradburn N. Response Effects in Surveys: a Review and Synthesis. Chicago: Aldine; 1974.

(43) Torangeau R, Smith TW. Asking sensitive questions: The impact of data collection, mode, question format and question context. Public Opinion Quarterly 1996;60:275-304.

(44) Rehm J, Arminger G. Alcohol consumption in Switzerland 1987-93: adjusting for differential effects of assessment techniques on the analysis of trends. Addiction 1996;91:1335-44.

(45) Bongers IMB, Van Oers JAM. Mode effects on self-reported alcohol use and problem drinking: mail questionnairies and personal interviewing compared. Journal of Studies of Alcohol 1998;59:280-5.

(46) Kraus L, Augustin R. Measuring alcohol consumption and alcohol related problems: Comparison of responses from self-administered questionnairies and telephone interviews. Addiction 2001;96:459-71.

(47) Kempf A, Remington P. New challenges for telephone survey research in the twentieth century. Annual Review of Public Health 2007;28:113-26.

(48) O'Toole J, Sinclair M, Leder K. Maximising reponse rates in household telephone surveys. BMC Medical Reserach Methodology 2008;8(71).

(49) Dawson D, Li T, Grant B. A prospective study of risk drinking: at risk for what? Drug and Alcohol Dependence 2008;95:62-72.

(50) Bühringer G, Augustin R, Bergmann E, Bloomfield K, Funk W, Junge B, et al. Alcohol Consumption and Alcohol-related Problems in Germany
31. Göttingen: Hogrefe & Huber; 2002.

(51) Dawson D. US low-risk drinking guidelines: an examination of four alternatives. Alc Clin Exp Res 2000;24:1820-9.

(52) Dawson D, Grant B, Li T. Quantifying the risks associated with exceeding recommended drinking limits. Alc Clin Exp Res 2005;29:902-8.

(53) Department of Mental Health and Substance Dependence WHO. International Guide for Monitoring Alcohol Consumption and Related Harm. 2000. Report No.: WHO/MSD/MESB/00.4.

(54) Rehm J, Mathers C, Popova S, Thavornacharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009;373(9682):2223-33.

(55) World Health Organization. Global Health Risks: Mortality and burden of disease attributable to selected major risks. Geneva: World Health Organization; 2009.

(56) Hasin D. Classification of alcohol use disorders. Alcohol Research & Health 2003;27:5-17.

(57) Üstün B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji Seal. WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview and methods. Drug and Alcohol Dependence 1997;47:161-9.

(58) Grant B, Harford T, Dawson D, Chou P, Pickering R. The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug and Alcohol Dependence 1995;39:37-44.

(59) Saunders J, Aasland O, Babor T, de la Fuente J, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption - II. Addiction 1993;88:791-804.

(60) Ewing J. Detecting alcoholism: the CAGE questionnaire. Journal of the American Medical Association 1984;252:1905-7.

(61) Porkorny A, Miller B, Kaplan H. The brief MAST: A shortened version of the Michigan Alcoholism Screening Test. American Journal of Psychiatry 1972;129:342.

(62) Russell M, Martier S, Sokol R, Mudar P, Bottoms S, Jacobson S, et al. Screening for pregnancy risk-drinking. Alc Clin Exp Res 1994;18:1156-61.

(63) Cherpitel C. A brief screening instrument for problem drinking in the emergency room: The RAPS4. Journal of Studies on Alcohol 2000;61:447-9.

(64) Cherpitel C. Screening for alcohol problems in the US general population: Comparison of the CAGE RAPS4 and RAPS4-QF by gender, ethnicity and service utilization. Alcoholism-Clinical and Experimental Research 2002;26:1686-91.

(65) Cherpitel C. Screening for alcohol problems in the emergency room: a rapid alcohol problems screen. Drug and Alcohol Dependence 1995;40:133-7.

(66) WHO. Schedules for Clinal Assessment in Neuropsychiatry. Washington, DC.: American Psychiatric Press; 1993.

(67) Lachner G, Wittchen H-U, Perkonigg A, Holly A, Schuster P, Wunderlich Uea. Structure content and reliability of the Munich-composite international diagnostic interview (M-CIDI) substance use sections. European Addiction Research 1998;4:28-41.

(68) Allen J, Litten R, Fertig J, Barbor T. A review of research on the alcohol use disorders identification test (AUDIT). Alcohol Clin Exp Res 1997;21:613-9.

(69) Berner M, Kriston L, Bentele M, Härter M. The alcohol use disorders identification test for detecting at-risk drinking: a systematic review and meta-analysis. Journal on Studies of Alcohol and Drugs 2007;68:461-73.

(70) Reinert D, Allen J. The alcohol use disorders identification test (AUDIT): a review of recent research. Alcohol Clin Exp Res 2002;26:272-9.

(71) Bloomfield K, Wicki M, Gustafsson N-K, Mäkelä P, Room R. Changes in alcohol-related problems after alcohol policy changes in Denmark, Finland and Sweden. Journal on Studies of Alcohol and Drugs 2010;71:32-40.

(72) Fleming M, Barry K, McDonald R. The alcohol use disorders identification test (AUDIT) in a college sample. International Journal of the Addictions 1991;26:1173-85.

(73) Gmel G, Heeb J, Rehm J. Is frequency of drinking an indicator of problem drinking? A psychometric analysis of a modified version of the alcohol use disorders identification test in Switzerland. Drug and Alcohol Dependence 2001;64:151-61.

(74) Homila M. Intoxication and hazardous use of alcohol: results from the 1992 Finnish drinking habits study. Helsinki: Social Research Institute on Alcohol Studies; 1995.

(75) Ivis F, Adlaf E, Rehm J. Incorporating the AUDIT into a general population telephone survey: a methodological experiment. Drug and Alcohol Dependence 2000;60:97-104.

(76) Knibbe R, Derickx M, Kuntsche S, Grittner U, Bloomfield K. A comparison of the Alcohol Use Disorder Identification Test (AUDIT) in general population surveys in nine European countries. Alcohol and Alcoholism 2006;41:i19-i25.

(77) Medina-Mora E, Carreño S, de la Fuente J. Experience with the alcohol use disorders identification test (AUDIT) in Mexico. In: Marc Galanter, Lee Ann Kaskutas, editors. Recent Developments in Alcoholism.New York: Plenum Press; 1998.

(78) Selin KH. Alcohol Use Disorder Identification Test (AUDIT): What does it screen? Performance of the AUDIT against four different criteria in a Swedish population sample. Substance Use & Misuse 2006;41:1881-99.

(79) Wicki M, Gustafsson N-K, Mäkelä P, Gmel G. Dimensionality of drinking consequences - cross-cultural comparability and stability over time. Addiction Research & Theory 2009;17:2-16.

(80) Reinert D, Allen J. The alcohol use disorders identification test: An update of research findings. Alcohol Clin Exp Res 2007;31:185-99.

(81) Kelly T, Donovan J. Confirmatory factor analysis of the alcohol use disorder identification test (AUDIT) among adolescents treated in emergency departments. Journal of Studies on Alcohol 2001;62:838-42.

(82) Rist F, Glöckner-Rist A, Demmel R. The alcohol use disorders identification test revisited: Establishing its structure using nonlinear factor analysis and identifying asubgroups of respondents using latent class factor analysis. Drug and Alcohol Dependence 2009;100:71-82.

(83) Cherpitel C, Ye U, Bond J, Borges G, Cremonte Mea. Cross-national performance of the RAPS4/RAPS-QF for tolerance and heavy drinking: data from 13 countries. Journal of Studies on Alcohol 2005;66:428-32.

(84) Cherpitel C, Ye U, Bond J, Borges G, Cremonte M, et al. Screening, Brief Intervention and Referral to treatment (SBIRT) in a Polish emergency Department: Three-month outcomes for a randomized, controlled clinical trial. Journal of Studies on Alcohol and Drugs 2009;70:982-90.

(85) Cremonte M, Cherpitel CJ. Performance of screening instruments for alcohol use disorders in emergency department patients in Argentina. Subst Use Misuse 2008;43(1):125-38.

(86) Cherpitel CJ, Borges G. Screening instruments for alcohol problems: a comparison of cut points between Mexican American and Mexican patients in the emergency room. Subst Use Misuse 2000 Aug;35(10):1419-30.

(87) Plant M, Miller P, Thornton C, Plant M, Bloomfield K. Life stage, alcohol consumption patterns, alcohol-related consequences and gender. Substance Abuse 2000;21:265-81.

(88) Mulford H, Miller D. Drinking in Iowa: IV. Preoccupation with alcohol and definitions of alcohol heavy drinking and trouble due to drinking. Quartely Studies on Alcohol 1960;60:203-8.

(89) Hilton M. Demographic characteristics and the frequency of heavy drinking as predictors of self-reported drinking problems. In: W.Clark, M.Hilton, editors. Alcohol in America: drinking practices and problems.Albany: State University of New York Press; 1991. p. 194-212.

(90) Rootman I, Moser J. Community response to alcohol-related problems. Washington, DC: USGPO; 1985.

(91) Gmel G, Rehm J, Room R, Greenfield T. Dimensions of alcohol-related social and health consequences in survey research. Journal of Substance Abuse 2000;12:113-38.

(92) Room R. Concepts and items in measuring social harm from drinking. Journal of Substance Abuse 2000;12:93-111.

(93) Rehm J, Gmel G. Patterns of alcohol consumption and social consequences. Results from an 8-year follow-up study in Switzerland. Addiction 1999;94:899-912.

(94) Gmel G, Kuntsche E, Wicki M, Labhart F. Measuring alcohol-related consequences in school surveys: alcohol-attributable consequences or consequences with students' alcohol attribution. Am J Epidemiol 2010;171(1):93-104.

(95) Rossow I, Hauge R. Who pays for the drinking? Characteristics of the extent and distribution of social harms from others' drinking. Addiction 2004;99:1094-102.

(96) Eliany M, Giesbrecht N, Nelson M, Wellman B, Wortley S. Alcohol and other drug use by Canadians: a national alcohol and other drug survey. Ottawa: Health Canada; 1992.

(97) Makela P, Fonager K, Hibell B, Nordlund S, Sabroe S, Simpura J. Drinking Habits in the Nordic Countries. Oslo: National Institute for Alcohol and Drug Research; 1999. Report No.: 2/99.

(98) Orford J, Templeton L, Vellemann R, Copello A. Family members of relatives with alcohol, drug and gambling problems: a set of standardized questionnairies for assessing stress, coping and strain. Addiction 2005;100:1611-24.

(99) Bergmark KH, Graham K, Nordvik M. A comparison of alcohol-related aggression in six European countries. In: Bloomfield K, et al, editors. Gender, Culture and Alcohol Problems: A Multi-national Study.Berlin: Charité University Medicine; 2005. p. 103-15.

(100) GENACIS. Recommendations for the use of drinking indicators. http://www genacis org/index php 10 A.D. October 26 [cited 10 A.D. Nov 24];

(101) Anderson P, Gordon R. Recommendations about sensible drinking amounts. Results of a survey. http://ec europa eu/health/ph_determinants/life_style/alcohol/documents/ev_20090217_co02_en pdf 2009 [cited 10 A.D. Nov 10];

(102) International Center for Alcohol Policies. International drinking guidelines. http://www icap org/table/InternationalDrinkingGuidelines 2010 [cited 10 A.D. Nov 10];Available from: URL: http://www.icap.org/table/InternationalDrinkingGuidelines

(103) American Psychiatric Association. DSM-IV: Diagnostic and statistical manual of mental disorders. Washington, D.C.: American Psychiatric Association; 1994.

(104) World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization; 1992.