Addiction

REVIEW

doi:10.1111/j.1360-0443.2009.02842.x

Brief screening questionnaires to identify problem drinking during pregnancy: a systematic review

Ethel Burns[1], Ron Gray[2] & Lesley A. Smith[1]

School of Health and Social Care, Oxford Brookes University, Jack Straws Lane, Marston, Oxford, UK[1] and National Perinatal Epidemiology Unit, University of Oxford, Old Road Campus, Headington, Oxford, UK[2]

ABSTRACT

Aims Although prenatal screening for problem drinking during pregnancy has been recommended, guidance on screening instruments is lacking. We investigated the sensitivity, specificity and predictive value of brief alcohol screening questionnaires to identify problem drinking in pregnant women. **Methods** Electronic databases from their inception to June 2008 were searched, as well as reference lists of eligible papers and related review papers. We sought cohort or cross-sectional studies that compared one or more brief alcohol screening questionnaire(s) with reference criteria obtained using structured interviews to detect 'at-risk' drinking, alcohol abuse or dependency in pregnant women receiving prenatal care. **Results** Five studies (6724 participants) were included. In total, seven instruments were evaluated: TWEAK (Tolerance, Worried, Eye-opener, Amnesia, Kut down), T-ACE [Take (number of drinks), Annoyed, Cut down, Eye-opener], CAGE (Cut down, Annoyed, Guilt, Eye-opener], NET (Normal drinker, Eye-opener, Tolerance), AUDIT (Alcohol Use Disorder Identification Test), AUDIT-C (AUDIT-consumption) and SMAST (Short Michigan Alcohol Screening Test). Study quality was generally good, but lack of blinding was a common weakness. For risk drinking sensitivity was highest for T-ACE (69-88%), TWEAK (71–91%) and AUDIT-C (95%), with high specificity (71–89%, 73–83% and 85%, respectively). CAGE and SMAST performed poorly. Sensitivity of AUDIT-C at score ≥ 3 was high for past year alcohol dependence (100%) or alcohol use disorder (96%) with moderate specificity (71% each). For life-time alcohol dependency the AUDIT at score ≥ 8 performed poorly. **Conclusion** T-ACE, TWEAK and AUDIT-C show promise for screening for risk drinking, and AUDIT-C may also be useful for identifying alcohol dependency or abuse. However, their performance as stand-alone tools is uncertain, and further evaluation of questionnaires for prenatal alcohol use is warranted.

Keywords Alcohol, prenatal, questionnaires, screening, sensitivity, specificity, systematic review.

Correspondence to: Lesley Smith, School of Health and Social Care, Oxford Brookes University, Marston Campus, Jack Straws Lane, Oxford OX3 OFL, UK.
E-mail: lesleysmith@brookes.ac.uk

Submitted 22 January 2009; initial review completed 24 March 2009; final version accepted 16 October 2009

INTRODUCTION

Heavy alcohol consumption during pregnancy is associated with a range of effects, including fetal alcohol syndrome (FAS) and milder variants encompassed by fetal alcohol spectrum disorder (FASD) [1,2]. FAS is a complex disorder characterized by multiple birth defects, including growth and facial anomalies and neurodevelopmental problems, with a prevalence of 0.5–2 per 1000 live births [1–3]. FASD reflects a wider range of disorders including physical, cognitive and behavioural problems, and is more common than FAS, with a prevalence of nine to 10 per 1000 live births [4]. While the adverse effects of binge drinking [5,6] or low to moderate consumption [7–9] on pregnancy outcomes are inconclusive, the potential for harm cannot be ruled out.

Lack of an international consensus on what constitutes a safe level of drinking during pregnancy has led to different recommendations in national guidelines [10]; some advocate alcohol abstinence, while others advise that pregnant women drink no more than 1 or 2 units of alcohol once or twice a week [11]. Whereas many women either stop or reduce their alcohol consumption once they are aware that they are pregnant, some continue to drink. Estimates for prenatal alcohol consumption vary widely: around 10–12% in the United States [12] to

higher rates in Europe (25–54%), particularly in the United Kingdom [13,14]. Furthermore, a significant minority of women continue to drink at problematic levels when pregnant, although accurate estimates are unavailable [7,13]. Evidence suggests that pre-pregnancy drinking patterns predict alcohol consumption during pregnancy: the more a woman drinks before she is aware that she is pregnant, the more she is likely to continue drinking during her pregnancy [14–18].

Clinical guidelines recommend detailed screening for alcohol consumption be undertaken prenatally to identify high-risk or problem drinking [19]. However, there is no clear guidance on how to screen women effectively, such as recommending which screening test to use. Obtaining accurate estimates of alcohol consumption is difficult: direct questions on drinking habits may inhibit accurate reporting, and is complicated further by differences in glass sizes and strength of alcoholic beverages [20,21]. This is confusing for caregivers and women. Studies have revealed the limitations of routine prenatal care as a means of identifying women with problem drinking [17,22–24].

Standard tests for excessive alcohol consumption include in-depth interviews, biochemical tests and self-reported questionnaires. Several screening questionnaires designed to elicit information on alcohol consumption are available. Most were designed for detecting alcohol dependency in men; few are specifically for use with pregnant women. None are in widespread use in prenatal care settings, despite evidence to suggest that screening in itself can reduce alcohol consumption [25,26]. It is essential to use the best screening tool to identify accurately women who drink at a level which is potentially harmful to the embryo, fetus or child in order to administer an effective intervention. Studies have shown that, combined with a brief intervention (BI), screening for prenatal alcohol use results in reduced alcohol consumption [27]. BIs typically takes 5–10 minutes, and consist of advice aiming to motivate individuals to modify their behaviour [28].

Several reviews have highlighted the importance of screening for alcohol use in pregnancy; however, none have used systematic methodology in their appraisal of available screening tools [12,29–33].

The aim of this systematic review was to investigate the sensitivity, specificity and predictive value of brief alcohol screening questionnaires to identify 'risk' drinking in pregnant women.

METHODS

We conducted this review in accordance with guidelines for performing systematic reviews of diagnostic accuracy studies [34].

Search methods for identification of studies

We searched five electronic databases via OVID from their inception to June 2008: MEDLINE, EMBASE, PsycINFO, CINAHL and British Nursing Index.

The search strategy was adapted for each database using a combination of text words and MeSH terms. Three sets of search terms were combined. The first set comprised terms for named tests, general terms for screening questionnaires and statistical terms associated with tests such as sensitivity and specificity [35]. The second set of terms encompassed alcohol drinking behaviour, and the third set comprised terms to limit the screening to pregnant women. Terms to serve as a methodological filter were not used to increase the specificity of the search, as this has been shown to lack sensitivity and thus miss potentially relevant studies [36,37]. To identify other potentially eligible studies, we screened reviews of related research and reference lists of eligible studies, and carried out forward citation searching using Science Citation Index and PubMed related articles functions.

Criteria for including studies in the review

We included cohort or cross-sectional studies that compared one or more brief alcohol screening questionnaire(s) with an appropriate reference standard for risk drinking, alcohol abuse or dependency in pregnant women of any age or ethnic origin. We excluded case–control studies due to their potential for introducing spectrum bias where milder, more difficult to diagnose cases are excluded from case–control studies resulting in overestimation of sensitivity and specificity. A review of 218 studies demonstrated that case–control studies tended to overestimate the diagnostic odds ratio (DOR) threefold compared with studies on a clinical cohort [38].

Index (screening) tests

We limited the review to brief alcohol screening questionnaires, as they are more feasible to administer. Longer questionnaires may be less acceptable to practitioners undertaking comprehensive prenatal health assessments. Index tests sought were: Alcohol Use Disorders Identification Test (AUDIT) [39], AUDIT-C (AUDIT consumption questions) [40], AUDIT-3 (third AUDIT question) [41], CAGE (Cut-down, Annoyed, Guilt, Eye-opener) [42], SMAST (Short Michigan Alcoholism Screening Test) [43], T-ACE [44] [Take (number of drinks), Annoyed, Cut down, Eye-opener], TWEAK [45] [T: tolerance: how many drinks can you hold ('hold' version >5 indicates tolerance) or how many drinks can take before you begin to feel the effects ('high' version >2 indicates tolerance), W: have close friends or relatives worried or complained about your drinking in the last year, E: eye-

openers: do you sometimes take a drink in the morning when you first get up, A: amnesia: has a friend or family member ever told you about things you said or did while you were drinking that you could not remember, K: kut down: do you sometimes feel the need to cut down on your drinking] and NET (Normal, Eye-opener, Tolerance) [46] (see Table 1). We also included any other brief questionnaires not listed above but subsequently identified by the search.

Reference standards

As there is no recognized gold standard method for ascertaining alcohol consumption, we chose what are generally considered to be more reliable methods of ascertainment as reference standards. First, for 'risk' drinking, we accepted quantity and frequency of alcohol use obtained using a structured interview, such as Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS) [47] or a time-line follow-back (TLFB) procedure [48,49]. Secondly, for alcohol dependency or abuse we accepted clinical diagnoses made according to (a) the Diagnostic and Statistical Manual (DSM) III-R or IV-R or (b) International Classification of Disease (ICD)-10 criteria obtained during standardized clinical interview. We excluded studies that used methods other than a structured interview as a reference test such as biomarkers or self-completed questionnaires.

Data extraction

Data were extracted by two reviewers independently, with discrepancies resolved by discussion. We extracted details of study design, population characteristics and results.

Quality assessment

We assessed the studies for methodological quality against the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria [50]. We evaluated the reporting of patient selection, blinding, completeness of descriptions of reference and screening tests and data collection, and likelihood of verification bias.

Statistical analysis

For each study, the sensitivity (proportion of true positives identified by the screening test), specificity (proportion of true negatives identified by the screening test), positive predictive value (PPV: proportion of those who screen positive who are true positives) and area under the receiver operator curve (AUROC) were either extracted from the paper or, where this was not possible, calculated from the data extracted from the paper by one author and checked by the other.

RESULTS

Searches

We identified 1374 unique citations, and after screening titles and abstracts we considered 32 as potentially relevant for which full text papers were obtained. After screening the full text papers, six papers reporting on five studies met our inclusion criteria [32,44,51–54]. Of the 26 excluded studies: 10 were reviews, guidelines or editorials and seven did not validate a screening tool against a reference standard. One each did not evaluate a brief questionnaire, prenatal population or problem drinking, one was a duplicate publication of an included study and one was unavailable from the British Library. Four were excluded on the basis of the reference standard: one used a biomarker [55] and one a self-completed questionnaire [45]. Two used brief interviews, the content of which were not described [56,57] (Fig. 1).

Study characteristics

The five included studies involved 6724 participants. Participant characteristics, screening questionnaires and details of reference standard tests for each study are described in Table 2. The study by Russell was published in two papers [32,53]. In Russell (1994), the tolerance question asked 'how many drinks does it take to make you feel "high" '; in Russell (1996) the question was 'how many drinks can you hold?'. All five studies were conducted in the United States. Three included ethnic minority women of low socio-economic status [32,44,53,54], and two included a more diverse group representative of the general population [51,52]. Women were assessed at the first prenatal visit and were in the first or second trimesters of pregnancy in four of the studies [32,44,51,53,54]; in the other, the population comprised women who responded that they were pregnant at the time the survey was conducted rather than women seeking care at antenatal clinics [52]. Brief questionnaires assessed were: a derived version of AUDIT-C [52], CAGE [32,44], NET [32], SMAST [51], T-ACE [32,44,51,53] and TWEAK [32,53,54]. In four studies [32,44,52–54] the questionnaires were not administered as independent instruments, but were derived from other questionnaires: T-ACE, TWEAK and NET were derived from a composite of MAST, CAGE and a tolerance question [32,44,53]. This may have desensitized participants, making them feel more comfortable about disclosing their alcohol consumption and increasing the sensitivity of the questionnaires. The order in which questions are delivered has also been shown to affect responses: when the TLFB interview was administered before brief questionnaires, alcohol intake was under-reported [32].

Table 1 Brief alcohol screening questionnaires.

Name	*Components*	*Strengths and limitations*
AUDIT (Saunders 1993) [39]	10 multiple-choice questions; range 0–40 How often do you have a drink containing alcohol? How many drinks containing alcohol do you have on a typical day when you are drinking? How often do you have 6 or more drinks on one occasion? How often during the last year have you found that you were not able to stop drinking once you had started? How often during the last year have you failed to do what was normally expected of you because of drinking? How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session? How often during the last year have you had a feeling of guilt or remorse after drinking? How often during the last year have you been unable to remember what happened the night before because you were drinking? Have you or someone else been injured as a result of your drinking? Has a relative or friend, or a doctor or other health care worker been concerned about your drinking or suggested you cut down? A cut-off score of 8 (men) or 6 (women) is used to identify hazardous drinkers	Takes about 2 minutes to complete Evaluated in a variety of settings, populations and cultural groups Developed by WHO to detect alcohol dependency in men
AUDIT-C (Bush 1998) [40]	How often have you had a drink containing alcohol in the past year? Never, monthly or less, 2–4 times a month, 2 or 3 times a week, ≥4 times a week How many drinks did you have on a typical day when you were drinking in the past year? 1 or 2, 3 or 4, 5 or 6, 7–9, ≥10 How often did you have six* or more drinks on one occasion during the past year? Never, < monthly, monthly, weekly, daily or almost daily Score 0–4 for each question, maximum score 12. Threshold for positive score ≥3. *Adapted for NESARC study [52] to reflect larger drink size in United States—how often do you drink five or more drinks on one occasion?	Developed from AUDIT Shorter and therefore quicker to administer
CAGE (Mayfield 1974) [42]	Four short questions C: have you ever felt you should cut down on your drinking? A: have people annoyed you by criticizing your drinking? G: have you ever felt bad or guilty about your drinking? E: eye-opener: have you ever had a drink first thing in the morning to steady your nerves or to get rid of a hangover? Score 1 point for each 'yes' response; range 0–4. Threshold for positive score ≥2	Takes less than 1 minute to complete Developed for use in men Not considered to be accurate for identifying alcohol abuse in women, and for use in different ethnic groups
T-ACE (Sokol 1989) [44]	Four-item derivative of MAST and CAGE T: tolerance: how many drinks does it take to make you feel high? (>2 indicates tolerance) A: annoyed question from CAGE C: cut-down question from CAGE E –eye-opener question from CAGE Score 2 points for tolerance; 1 point for others; range 0–5; threshold for positive score ≥2	Takes about 1 minute to complete Developed for detecting excessive alcohol use in pregnant women
TWEAK (Russell 1979) [45]	Five-item derivative of CAGE T: tolerance: how many drinks can you hold ('hold' version >5 indicates tolerance) or how many drinks can take before you begin to feel the effects ('high' version >2 indicates tolerance) W: have close friends or relatives worried or complained about your drinking in the last year? E: eye-openers: do you sometimes take a drink in the morning when you first get up? A: amnesia: has a friend or family member ever told you about things you said or did while you were drinking that you could not remember? K: kut down: do you sometimes feel the need to cut down on your drinking? Score 2 points each for first 2 items and 1 point each for last 3; range 0–7; threshold for positive score ≥2	Takes about 1 minute to complete Developed for detecting excessive alcohol use in pregnant women
NET (Bottoms 1989) [46]	N: normal drinker: do you feel you are a normal drinker? E: eye-opener question from CAGE T: tolerance: how many drinks does it take to make you feel high? (>2 indicates tolerance) Score 1 point each for not normal or eye openers and 2 points for tolerance; range 0–4	Takes about 1 minute to complete Developed for use in an obstetric population

Table 1 *Cont.*

Name	*Components*	*Strengths and limitations*
SMAST (Selzer 1975) [43]	13 'yes' or 'no' questions (past 12 months only) 1. Do you feel you are a normal drinker? 2. Do your spouse, parents or other close relative worry or complain about your drinking? 3. Do you ever feel guilty about your drinking? 4. Do friends or relatives think you are a normal drinker? 5. Are you able to stop drinking when you want to? 6. Have you ever attended a meeting of Alcoholics Anonymous? 7. Has your drinking ever caused problem between you, a spouse, parents or close relative? 8. Have you ever got into trouble at work because of drinking? 9. Have you ever neglected your obligations your family or your work for 2 or more days in a row because you were drinking? 10. Have you ever gone to anyone for help about your drinking? 11. Have you ever been in a hospital because of drinking? 12. Have you ever been arrested for drunk driving or driving after drinking? 13. Have you ever been arrested, however short a time, because of drinking? Score: each 'yes' response = 1 point	Shortened version of MAST, quicker to administer

AUDIT: Alcohol Use Disorders Identification Test; AUDIT-C: AUDIT consumption questions; NESARC: National Epidemiologic Survey on Alcohol and Related Conditions; MAST: Michigan Alcohol Screening Test; SMAST: Short Michigan Alcohol Screening Test; WHO: World Health Organization.

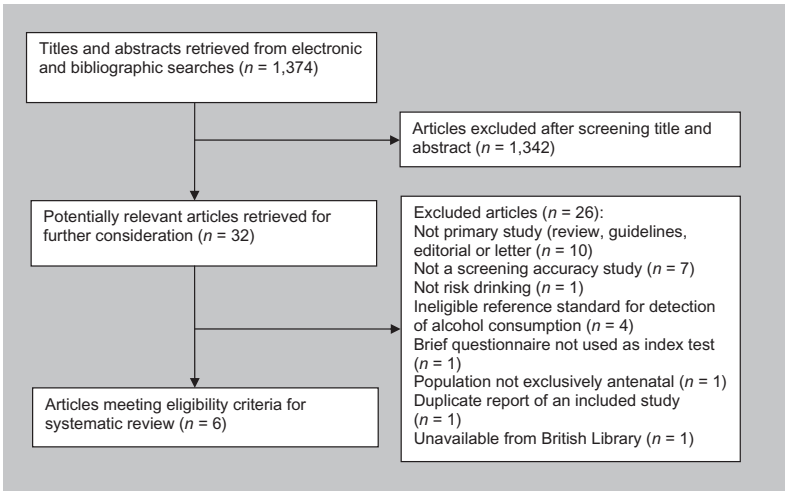

Figure 1 Results of search

AUDIT-C was derived from the questionnaire used in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), and asked about drinking five or more drinks on a single occasion rather than the usual six to account for the larger drink size in the United States. The study by Chang [51] administered T-ACE embedded in a health and habits survey.

Quality assessment

In two studies [32,44] the reference standard was risk drinking (at least 1 oz absolute alcohol per day in the periconception period) derived from quantity–frequency information assessed using TLFB. One study defined risk drinking as more than two drinks per day before pregnancy derived from all available information on drinking including interview, and life-time alcohol abuse or dependence using a structured interview [51]. Another defined risk drinking as more than seven drinks per week or more than three drinks on a single day at least once a month in the past year, and also evaluated past year alcohol dependence or alcohol use disorder according to DSM criteria [52]. One study reported results for high-risk drinking in pregnancy defined as more than one drink or at least three drinks monthly during pregnancy [54].

Table 3 shows the results for the quality assessment. Study quality was generally good, with only two QUADAS items not met by any of the studies (interpretation of screening test result without knowledge of reference

Table 2 Study characteristics

Study	*Participants*	*Sample size*	*Screening tools*	*Gold standard*
Sokol 1989 [44] USA	Consecutive pregnant women attending first visit at inner-city antenatal clinic Age: mean (SD) 23.9 (6.0); range 13–32 years Ethnicity: black 100% Parity: mean (SD): 1.1 (1.4) Gestational age: NR Smoking: mean (SD) cigs/day 5.1 (5.1)	Eligible: 1065 Recruited: 971 (having drunk alcohol at some time) Analysed: 971 Excluded: 94 (non-drinkers)	MAST, CAGE, tolerance, T-ACE (latter constructed from preceding tools) tolerance = > 2 drinks to feel 'high'	One-week recall of average periconceptual drinking and 2-week drinking history (standardized clinical interview) Risk drinking defined as at least 1 oz absolute alcohol per day (periconception)
Russell 1994 [32] Detroit, USA	Consecutive pregnant women attending first visit at inner-city antenatal clinic, low socio-economic status Age: mean 24.2 Ethnicity: black 100% Parity: 1.2 Gestational age: 23 weeks Smoking: 6.6 per day	Eligible: Recruited: 4743 (having drunk alcohol at some time) Analysed: 4743 Excluded:	MAST, CAGE, tolerance, T-ACE, TWEAK, NET (latter 3 constructed from preceding 3 tools)	Time-line follow-back procedure periconceptual drinking (clinical interview) Risk drinking defined as at least 1 oz absolute alcohol per day (periconception)
Russell 1996 [53] Detroit, USA	Subset of women recruited in Russell 1994. Purpose of study to evaluate 'Hold' version of TWEAK and T-ACE Age: mean 24.4 years Ethnicity: black 100% Parity: 1.2 Gestational age: 23 weeks Smoking: 6.7 per day	Eligible: 3077 Recruited: 3056 Analysed: 2717 (having drunk alcohol at some time) Excluded: 339 (non-drinkers)	MAST, CAGE, tolerance (how many drinks can you hold?), TWEAK (latter constructed from previous tools), T-ACE administered alone to (1420) and constructed from preceding tools	Time-line follow-back procedure periconceptual drinking (clinical interview) Risk drinking defined as at least 1 oz absolute alcohol per day (periconception)

Chang 1998 [51] Boston, USA	Pregnant women attending first visit at inner-city antenatal clinic. Exclusion criteria: substance abuse/dependence, gestational age >28 weeks, intention to terminate, alcohol abstinence >6 months in T-ACE pos, non-English speakers Age: mean 30.1 years Ethnicity: Caucasian 71%, black 18%, Hispanic 8%, Asian 3% Parity: nulliparae 53% Gestational age: NR Smoking: 8% smokers	Eligible: Recruited: 1165 returned initial survey Analysed: 250 consecutive screen positive, 100 consecutive screen negative (total 350) Excluded: 175 as surplus to recruitment need	T-ACE embedded in a health and habits survey (tolerance question scored positive if women reported felt intoxicated with 2 drinks rather than more than 2 drinks, AUDIT, SMAST	Life-time alcohol abuse or dependence according to DSM-III-R criteria obtained during structured clinical interview Risk drinking defined as >2 drinks per drinking day before pregnancy derived from all available information: time-line follow-back procedure, AUDIT question (how many drinks do you have on a typical day and health and habits self-reported questionnaire)
Dawson 2001 [54] Washington DC, USA	Pregnant women attending first visit at antenatal clinics in metropolitan area. Non-English speakers excluded. Age: mean 26.6 years Ethnicity: Hispanic African American 92% Parity: NR Gestational age: 24 weeks Smoking: NR	Eligible: 631 Recruited: 404 life-time drinkers Analysed: 404 Excluded: NR	TWEAK	Audio, computer-assisted self-interview High-risk drinking defined as >1 drink or 3+ drinks at least once a month during pregnancy Any risk drinking defined as any drinking below threshold for high risk during pregnancy
Dawson 2005 [52] USA	Subsample of pregnant women from a large population survey (NESARC), random sample of adults aged at least 18 years Age: At least 18 years Ethnicity: NR Parity: NR Gestational age: NR Smoking: NR	Eligible: Recruited: $n = 43\,093$ (NESARC) Analysed: 256 past year drinkers who answered 'yes' to being pregnant Excluded: 197 past year abstainers	AUDIT-C (derived from individual questions in NESARC questionnaire) Note: Q3 frequency of drinking ≥5 drinks rather than ≥6 to account for larger standard drink size in USA (12 to 14 g ethanol rather than the usual 10)	Past year alcohol dependence or AUD, according to DSM-IV criteria by clinical interview (AUDADIS-IV) Risk drinking defined as >7 drinks per week or >3 drinks in a single day at least once a month self-reported from questionnaire

AUDIT: Alcohol Use Disorders Identification Test; AUDIT-C: AUDIT consumption; NESARC: National Epidemiologic Survey on Alcohol and Related Conditions; MAST: Michigan Alcohol Screening Test; SMAST: Short Michigan Alcohol Screening Test; TWEAK: T: tolerance: how many drinks can you hold ('hold' version ≥6 indicates tolerance) or how many drinks can take before you begin to feel the effects ('high' version >2 indicates tolerance), W: have close friends or relatives worried or complained about your drinking in the last year, E: eye-openers: do you sometimes take a drink in the morning when you first get up, A: amnesia: has a friend or family member ever told you about things you said or did while you were drinking that you could not remember, K: kut down: do you sometimes feel the need to cut down on your drinking; T-ACE: T: tolerance: how many drinks does it take to make you feel high? (>2 indicates tolerance), A: annoyed question from CAGE, C: cut-down question from CAGE, E: eye-opener question from CAGE; CAGE: C: cut-down, annoyed, guilt, eye-opener; SD: standard deviation; NR: not recorded.

Table 3 Quality appraisal.

Quality criterion	*Sokol 1989 [44]*	*Russell 1994 [32]*	*Russell 1996 [51]*	*Chang 1998 [51]*	*Dawson 2001 [54]*	*Dawson 2005 [52]*
Participants representative of who will receive the test in practice (consecutive series or random selection)	a	a		a		a
Selection criteria clearly described				a	a	a
Gold standard likely to correctly classify target condition	a	a	a	a	a	a
Time span between screening tests and gold standard short enough to ensure no change in target condition	a	a	a		a	
All participants or a random selection received gold standard	a	a	a		a	a
Participants received the same gold standard regardless of screening test result	a	a	a	a	a	a
Gold standard and screening test were independent of one another	a	a	a	a	a	
Screening test described in sufficient detail to permit replication	a	a	a	a	a	a
Gold standard described in sufficient detail to permit replication	a	a	a	a		a
Screening test results interpreted without knowledge of gold standard result				a		
Gold standard results interpreted without knowledge of screening test result						
Participant baseline demographic characteristics reported	a	a	a	a	a	
Screening tests reported for all participants including unclear results	a	a		a		
Participant exclusions explained	a		a			

[a]Indicates that the measure was addressed adequately in the study.

standard result and the converse). One study [51] was susceptible to partial verification bias—all women were screened initially with T-ACE, then only 250 T-ACE-positive and 150 T-ACE-negative women were verified with a reference standard. All studies were cross-sectional in design. Selection criteria were not described in two studies [32,44], and the time lag between screening and clinical interview was unclear in two [51,52].

Screening test results

The sensitivity, specificity and PPV values of the screening tests are given in Table 4. Risk drinking prevalence ranged from 4.3% [44] to 30% [51], reflecting the different risk drinker criteria. Prevalence of life-time alcohol abuse or dependency was 33% [51], and past year alcohol dependency or any past year alcohol use disorder were 3.5% and 5.5%, respectively [52].

T-ACE and TWEAK had higher AUROCs compared with CAGE and NET for risk drinking (at least 1 oz alcohol daily periconception) (Table 4). This was true for both versions of the questionnaires (tolerance question 'hold' or 'high'). Because the questionnaire requires a cut-off score for meeting the definition of risk drinker, the results of questionnaires at traditional cut-off scores were considered (Table 1). Sensitivity (69% and 70%) and specificity (89% and 85%) of T-ACE (high) score ≥2 were similar in two studies [32,44]. The 'hold' version showed higher sensitivity (88%) with specificity 79% [53]. Sensitivity (74%) and specificity (71%) were lower for identification of risk drinking defined as >two drinks per drinking day before pregnancy [51].

Two studies evaluated TWEAK [32,53,54]. For identification of periconception risk drinking and score of ≥2, the 'hold' version had greater sensitivity (91%) than the 'high version (79%) with specificity of 77% and 83%, respectively [32,53]. For high-risk drinking during pregnancy, sensitivity and specificity were 71% and 73%, respectively [54]. AUDIT-C score ≥3 showed good sensitivity (95%) and specificity (85%) for past year risk drinking (more than seven drinks per week or more than three drinks on a single day at least once per month) [52].

Sensitivity (38% and 49%) of CAGE ≥2 was lower than for T-ACE and TWEAK in two studies, although with higher specificity (92% and 93%) [32,44] (Table 4). NET was evaluated in one study and showed best sensitivity (71%) and specificity (86%) at a cut-off score of ≥1, and AUROC higher than CAGE, but lower than T-ACE and TWEAK [32]. SMAST showed high specificity (96%) but low sensitivity (11%) [51].

For identification of life-time alcohol abuse or dependence, AUDIT had higher AUROC than T-ACE or SMAST (Table 4). However, sensitivity of AUDIT at cut-off scores

Table 4 Screening test results.

Study	*Reference standard*	*Prevalence*	*Test*	*Cut-off scores*	*Sensitivity (%)*	*Specificity (%)*	*PPV (%)*	*AUROC*
Sokol 1989 [44]	Risk drinking	42/971 (4.3%)	CAGE	≥1	59	82	13	NR
				≥2	38	92	18	
			T-ACE	≥1	76	79	14	NR
				≥2	69	89	23	
				≥3	38	97	40	
Russell 1994 [32]	Risk drinking	270/4743 (5.7%)	CAGE	≥1	68	82	18	0.776
				≥2	49	93	30	
				≥3	30	98	52	
			T-ACE[a]	≥1	83	75	17	0.840
				≥2	70	85	22	
				≥3	45	97	46	
			TWEAK[a]	≥1	87	72	16	0.865
				≥2	79	83	22	
				≥3	59	94	39	
			NET	≥1	71	86	23	0.793
				≥2	61	87	22	
				≥3	24	99	58	
Russell 1996 [53]	Risk drinking	181/2717 (6.7%)	T-ACE[b]	≥1	91	70	18	0.887
				≥2	88	79	23	
				≥3	61	95	47	
			TWEAK[b]	≥1	92	67	17	0.894
				≥2	91	77	22	
				≥3	67	92	37	
Chang 1998 [51]	Life-time alcohol abuse or dependence	114/350 (33%)	T-ACE (tolerance ≥2)[c]	≥2	88	37	40	0.644
			T-ACE (tolerance >2)[d]	≥2	60	66	46	
			SMAST	≥3	15	98	77	0.624
			AUDIT	≥8	23	97	79	0.763
			AUDIT	≥10	11	99	87	
			AUDIT	≥11	7	100	100	
	Risk drinking	105/350 (30%)	T-ACE (tolerance ≥2)[c]	≥2	92	38	39	0.687
			T-ACE (tolerance >2)[d]	≥2	74	71	53	
			SMAST	≥3	11	96	55	0.551
Dawson 2005 [52]	Past year alcohol dependence	9/256 (3.5%)	AUDIT-C	≥3	100	71	11	0.939
				≥4	98	86	21	
				≥5	96	90	27	
				≥6	57	94	25	
				≥7	22	97	22	
	Past year alcohol use disorder	14/256 (5.5%)	AUDIT-C	≥3	96	71	17	0.888
				≥4	92	87	30	
				≥5	73	91	31	
				≥6	46	94	29	
				≥7	21	97	30	
	Risk drinking	45/256 (17.6%)	AUDIT-C	≥3	95	85	57	0.965
				≥4	69	97	84	
				≥5	54	98	83	
				≥6	34	99	94	
				≥7	18	100	100	
Dawson 2001 [54]	Any risk drinking	119/404 (29.5%)	TWEAK	≥1	66	64	43	NR
	High-risk drinking	34/404 (8.4%)		≥2	71	73	19	

PPV: positive predictive value; AUROC: area under receiver operator curve; [a]'High' version tolerance question; [b]'Hold' version tolerance question; [c]original tolerance score (see Table 1); [d]Modified tolerance score; AUDIT: Alcohol Use Disorders Identification Test; AUDIT-C: AUDIT consumption questions; NESARC: National Epidemiologic Survey on Alcohol and Related Conditions; SMAST: Short Michigan Alcohol Screening Test; TWEAK: T: tolerance: how many drinks can you hold ('hold' version >5 indicates tolerance) or how many drinks can take before you begin to feel the effects ('high' version >2 indicates tolerance), W: have close friends or relatives worried or complained about your drinking in the last year, E: eye-openers: do you sometimes take a drink in the morning when you first get up, A: amnesia: has a friend or family member ever told you about things you said or did while you were drinking that you could not remember, K: kut down: do you sometimes feel the need to cut down on your drinking; T-ACE: T: tolerance: how many drinks does it take to make you feel high? (>2 indicates tolerance), A: annoyed question from CAGE, C: cut-down question from CAGE, E: eye-opener question from CAGE; CAGE: C: cut-down, annoyed, guilt, eye-opener; SD: standard deviation; NR: not recorded.

8, 10 and 11 was much lower than for T-ACE. When the tolerance question of T-ACE was modified to include '≥two drinks to feel high', sensitivity was higher than the original requirement for >two drinks [51].

Sensitivity of AUDIT-C at cut-off score ≥3 was high for past year alcohol dependence (100%) or alcohol use disorder (96%) with specificity of 71% (Table 4).

DISCUSSION

Our review of five studies comparing a screening questionnaire with a structured interview found that TWEAK, T-ACE and AUDIT-C had the highest sensitivity for identifying prenatal risk drinking. Sensitivity values indicate that about seven to nine of 10 risk drinkers would be identified correctly using one of these brief questionnaires. This higher sensitivity comes at the cost of decreased specificity, and also affects the PPV. The PPV of T-ACE and TWEAK was low, indicating that for every woman identified correctly with the questionnaire, as many as three women could be identified falsely as risk drinkers. This has resource implications, as all women with a positive screen would require further verification and/or intervention. However, the cost of delivering a screening questionnaire has been estimated to be as little as £1.70 per participant [58], and BIs have been demonstrated to be cost-effective for reducing alcohol consumption and alcohol-related harm in primary care [59]. PPV of AUDIT C was higher than for TWEAK or T-ACE; however, PPV comparison across populations with different prevalence of the condition of interest cannot be made, because the PPV is affected by the prevalence. The higher prevalence of risk drinking in the study that evaluated AUDIT-C [52] may be due to a risk drinking definition which reflected a lower level of drinking than in the other studies. It might also be indicative of heavier drinking in the time-period before the pregnancy was known, or it might be that women were more comfortable disclosing drinking behaviour in the context of a general population survey conducted at home rather than in a prenatal care setting.

TWEAK had greater sensitivity, although slightly lower specificity, than T-ACE for periconception risk drinking; however, TWEAK sensitivity was lower when used to identify any risk drinking during pregnancy. CAGE and SMAST were poor at identifying risk drinking. NET, while slightly better than CAGE, had lower sensitivity than T-ACE or TWEAK.

AUDIT-C was derived from consumption questions embedded in a larger survey and identified correctly nine in 10 women for past year risk drinking, alcohol dependency or alcohol use disorder. However, the lack of independence between AUDIT-C and the risk drinking reference standard (also constructed from the questions relating to alcohol consumption) probably led to the higher sensitivity and specificity.

In contrast with AUDIT-C, which comprises direct questions about alcohol consumption, T-ACE and TWEAK were developed to ascertain prenatal risk drinking indirectly, as it was thought that this would facilitate more honest reporting. If women under-report their drinking in response to direct questions, then they may also under-report drinking in structured interviews, and the 'gold standard' method of ascertaining alcohol consumption for risk drinking would have its own limitations. If direct questions do not lead to under-reporting, a brief questionnaire such as AUDIT-C would be an effective way of identifying those at risk. Irrespective of these constraints, asking women either directly or indirectly about their drinking presents an opportunity to identify problems and advise women about prenatal alcohol consumption.

CAGE did not perform well in identifying prenatal risk drinking, neither at the suggested cut-off score of ≥2 nor at the lower cut-off score of 1. CAGE has been shown to be relatively insensitive in identifying alcohol abuse or dependence in women in general [60]. CAGE was developed originally to identify alcohol dependency and has demonstrated better accuracy in in-patient populations, but is of limited value in primary care populations at the suggested cut-off score of ≥2 [61]. This underscores the need for tools to be re-tested and validated in different population groups, and for different alcohol use constructs and diagnoses before they are adopted for widespread screening.

Overall, the included studies were of good quality, improving the confidence in our conclusions. We included only studies that used a structured interview to ascertain drinking status as a reference standard. This could be perceived either as a strength, because it minimized the effect of bias associated with using less robust reference standards, or a weakness, by applying a restrictive inclusion criterion leading to exclusion of some studies. Application of this criterion resulted in exclusion of two studies. One used another self-completed questionnaire as a reference standard [45]. Comparing two self-completed questionnaires may overestimate the sensitivity and specificity of the screening test if the pattern of reporting is biased in a similar direction [54]. The other study validated data gathered from self-report by using biomarkers [55]. Although there are a number of promising approaches using biomarkers of exposure [62], self-report is still the main method of exposure ascertainment and is likely to remain so. The issue of how to accurately measure and summarize exposure [63,64] to alcohol remains problematic, and errors in self-report may lead to misclassification of consumption level or problem drinking.

The potential for results to be affected by selection bias seems small, as either a random selection or a consecutive series of participants was selected in all the included studies. While lack of blinding was a common methodological weakness, this has been shown to have a small effect on results [38] although, with subjective outcomes such as those assessed by clinical interview, we cannot rule out the possibility that these results are overestimated. It has been suggested that good reporting of study methods is a surrogate for the overall rigour of the study, and if described well it is likely that the study was methodologically sound [38]. While it was not stated explicitly that studies were reported in accordance with STARD guidelines [65] most did, in fact, conform to these guidelines, and methods were well described.

Two other studies were excluded because little information was provided on either the content or conduct of the interview. One study evaluated T-ACE in 150 pregnant women and found that T-ACE was in 100% agreement with identification of alcohol abuse obtained by interview conducted over the telephone [56]. The interview, conducted 120 days after T-ACE was administered, consisted of only two questions about alcohol consumption: what drinks were taken with food, and in what quantity, and did not report how alcohol abuse was defined. One other study compared detection of prenatal high-risk drinking in 70 Native American women by T-ACE and CAGE questionnaires and a 'brief' interview which was not defined [57]. T-ACE was shown to have higher sensitivity (93%) than CAGE (79%) for detecting high-risk drinking defined as five or more drinks on any one occasion during pregnancy.

A limitation of the included studies lies in uncertainty as to how the tools would perform in different populations of women, as they were all conducted in the United States, and in two the participants were socially disadvantaged. Also, as the questionnaires were not administered independently, it is uncertain if their performance would be as good when administered alone. In a sample of women given T-ACE alone, authors reported that the sensitivity and PPV were lower than when it was administered after MAST and CAGE [53].

We conducted a comprehensive literature search and obtained papers published in any language. Publication bias is always a threat to the validity of a review, and we may have missed studies. However, previous research has suggested that unpublished studies generally have small sample sizes and lower sensitivity and specificity than published studies [66].

Information gathered from the prenatal alcohol screening tests described in this review relates mainly to periconception alcohol consumption. As many women are some weeks into their pregnancy before realizing, one could contest the benefit of prenatal screening, as alcohol-related harm may have already occurred. However, alcohol may affect the fetus when consumed at any point in the pregnancy: it crosses the placenta readily but takes longer to be eliminated from the amniotic fluid than from the maternal circulation [67,68]. The risk of any potential harm is increased if the mother smokes as well as drinks: many women do both [69,70]. An alcohol-exposed newborn can experience similar withdrawal phenomena as a baby born to an opiate user [71]. Screening alone has been shown to reduce alcohol consumption during pregnancy [44]. The benefits of providing information about risks of drinking in pregnancy or delivery of a BI have been shown to persist beyond the prenatal period [72,73]. This is especially important if the mother is breastfeeding, as alcohol is transmitted readily to breast milk.

Pregnancy offers a unique health education opportunity: all women want a healthy baby, so they may be more receptive to behavioural modification. Midwives and nurses are usually the lead professional caring for women prenatally, hence are in an ideal position to screen women for alcohol consumption. Transforming research-based strategies into routine prenatal care can present a major challenge. One investigation has demonstrated that midwives can be educated to provide effective alcohol screening within the constraints of service provision resources [74]. A 1-day training session combined with continuous expert support was sufficient to implement a screening strategy with a brief questionnaire within the resources of regular prenatal care.

Screening identifies pregnant women who are problem drinkers. Depending upon the severity of the problem, women should be offered advice and/or a BI or referred to an alcohol treatment service. BIs developed for pregnant women comprise typically between one and four short counselling sessions, using a motivational interviewing technique with a trained professional (e.g. midwife, general practitioner, social worker) followed by personalized feedback. Randomized controlled trials have shown BIs to be consistently effective in reducing alcohol consumption among pregnant women [75–79].

CONCLUSIONS

Although the prevalence of risk drinking may be low (around 5%) the actual number of women involved is large, and the potential consequences for mother and baby are serious enough to warrant screening. At present there is no clear guidance about how to question women about alcohol consumption. T-ACE, AUDIT-C and TWEAK show promise as screening tools for identifying risk drinking in pregnant women, and could be administered by practitioners as part of their prenatal care. However, there is insufficient evidence on the sensitivity

and specificity of these tools outside the United States in representative samples of women and when administered as independent instruments. It is necessary to evaluate these screening tools in different settings and populations in order to select the optimal instruments. Then it would be useful to test them in conjunction with BIs in order to assess the effectiveness and cost-effectiveness of a programme of screening and treating pregnant women for 'risk' drinking and alcohol use disorders.

Declarations of interest

The authors have no conflicts of interest. No external funding was received.

Acknowledgements

We thank Nancy Bower for her translation of the Italian study by Bellati 1998, and Marcia Russell for clarifying her methodology. We are also grateful for the helpful comments from one of our peer reviewers. L.S. conceived the idea and wrote the protocol. All authors participated equally in all aspects of the review and writing the manuscript. L.S. and E.B. are funded by the School of Health and Social Care, Oxford Brookes University.

References

1. Abel E. L. *Fetal Alcohol Abuse Syndrome*. New York: Plenum Press; 1998.
2. Sokol R. J., Delaney-Black V., Nordstrom B. Fetal alcohol spectrum disorder. *JAMA* 2003; **290**: 2996–9.
3. May P. A., Gossage J. P. Estimating the prevalence of fetal alcohol syndrome. A summary. *Alcohol Res Health* 2001; **25**: 159–67.
4. Sampson P. D., Streissguth A. P., Bookstein F. L., Little R. E., Clarren S. K., Dehaene P. *et al.* Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. *Teratology* 1997; **56**: 317–26.
5. Henderson J., Kesmodel U., Gray R. Systematic review of the fetal effects of prenatal binge-drinking. *J Epidemiol Commun Health* 2007; **61**: 1069–73.
6. Strandberg-Larsen K., Nielsen N. R., Gronbaek M., Andersen P. K., Olsen J., Andersen A. M. Binge drinking in pregnancy and risk of fetal death. *Obstet Gynecol* 2008; **111**: 602–9.
7. Gray R., Henderson J. Review of the fetal effects of prenatal alcohol exposure. Report to the Department of Health; 2006. Available at: http://www.npeu.ox.ac.uk/downloads/reports/alcohol-report.pdf (accessed 5 June 2007).
8. Henderson J., Gray R., Brocklehurst P. Systematic review of effects of low-moderate prenatal alcohol exposure on pregnancy outcome. *Br J Obstet Gynaecol* 2007; **114**: 243–52.
9. Kelly Y., Sacker A., Gray R., Kelly J., Wolke D., Quigley M. A. Light drinking in pregnancy, a risk for behavioural problems and cognitive deficits at 3 years of age? *Int J Epidemiol* 2009; **38**: 129–40.
10. International Centre for Alcohol Policies (ICAP). *Government Policies on Alcohol and Pregnancy*. ICAP Reports; 1999. Available at: http://www.icap.org/portals/0/download/all_pdfs/ICAP_Reports_English/report6.pdf (accessed 10 January 2009).
11. Department of Health. *Drinking in Pregnancy Guidance*. 2008. Available at: http://www.nhs.uk/news/2008/03March/Pages/Pregnancydrinkinglimits.aspx (accessed 30 June 2009).
12. Krulewitch C. J. Alcohol consumption during pregnancy. *Annu Rev Nurs Res* 2005; **23**: 101–34.
13. Anderson P., Baumberg B. *Alcohol in Europe: A Public Health Perspective*. Report for the European Commission; 2006. Available at: http://ec.europa.eu/health-eu/doc/alcoholineu_content_en.pdf (accessed 10 January 2009).
14. Bolling K. G. C., Hamlyn B., Thornton A. *Infant Feeding Survey 2005*. Survey conducted on behalf of The Information Centre for Health and Social Care and the UK Health Departments by BMRB Social Research; 2007. Available at: http://www.ic.nhs.uk/pubs/ifs2005 (accessed 10 January 2009).
15. de Chazeron I. L. P., Ughetto S., Vendittelli F., Boussiron D., Sapin V., Coudore F. *et al.* Is pregnancy the time to change alcohol consumption habits in France?*Alcohol Clin Exp Res* 2008; **32**: 868–73.
16. Gladstone J. L. M., Nulman I., Koren G. Characteristics of pregnant women who engage in binge alcohol consumption. *Can Med Assoc J* 1997; **156**: 789–94.
17. Goransson M., Magnusson A., Bergman H., Rydberg U., Heilig M. Fetus at risk: prevalence of alcohol consumption during pregnancy estimated with a simple screening method in Swedish antenatal clinics. *Addiction* 2003; **98**: 1513–20.
18. Harrison P., Sidebottom A. Alcohol and drug use before and during pregnancy: an examination of use patterns and predictors of cessation. *Matern Child Health* 2009; **13**: 386–94.
19. National Institute for Health and Clinical Excellence (NICE). *Antenatal care routine care for the healthy pregnant woman*. 2008. Available at: http://www.nice.org.uk/nicemedia/pdf/CG62FullGuidelineCorrectedJune2008.pdf (accessed 24 April 2009).
20. Clift-Matthews V. Alcohol in pregnancy: what is a safe amount? *Br J Midwifery* 2008; **16**: 208.
21. Gill J. S., Donaghy M. Variation in the alcohol content of a 'drink' of wine and spirit poured by a sample of the Scottish population. *Health Educ Res* 2004; **19**: 485–91.
22. Elliott E. J., Payne J. M., Morris A., Haan E., Bower C. A. Fetal alcohol syndrome: a prospective national surveillance study. *Arch Dis Child* 2008; **93**: 732–7.
23. Leggate J. Improving pregnancy outcomes: mothers and substance misuse. *Br J Midwifery* 2008; **16**: 160–5.
24. Magnusson A., Goransson M., Heilig M. Unexpectedly high prevalence of alcohol use among pregnant Swedish women: failed detection by antenatal care and simple tools that improve detection. *J Stud Alcohol* 2005; **66**: 157–64.
25. Kypri K., Langley J. D., Saunders J. B., Cashell-Smith M. L. Assessment may conceal therapeutic benefit: findings from a randomized controlled trial for hazardous drinking. *Addiction* 2007; **102**: 62–70.
26. McCambridge J., Day M. Randomized controlled trial of the effects of completing the Alcohol Use Disorders Identification Test questionnaire on self-reported hazardous drinking. *Addiction* 2008; **103**: 241–8.
27. Chang G. Screening and brief intervention in prenatal care settings. *Alcohol Res Health* 2004; **28**: 80–4.
28. Babour T. F., Higgins-Biddle J. C., Saunders J. B., Monteiro M. G. *The Alcohol Use Disorders Identification Test: Guidelines for*

Use in Primary. Geneva: WHO Department of Mental Health and Substance Dependence; 2001.
29. Bailey B. A., Sokol R. J. Pregnancy and alcohol use: evidence and recommendations for prenatal care. *Clin Obstet Gynecol* 2008; **51**: 436–44.
30. Chang G. Alcohol-screening instruments for pregnant women. *Alcohol Res Health* 2001; **25**: 204–9.
31. Morse B. A., Hutchins E. Reducing complications from alcohol use during pregnancy through screening. *J Am Med Womens Assoc* 2000; **55**: 225–7.
32. Russell M., Martier S. S., Sokol R. J., Mudar P., Bottoms S., Jacobson S. *et al.* Screening for pregnancy risk-drinking. *Alcohol Clin Exp Res* 1994; **18**: 1156–61.
33. Savage C., Wray J., Ritchey P. N., Sommers M., Dyehouse J., Fulmer M. Current screening instruments related to alcohol consumption in pregnancy and a proposed alternative method. *J Obstet Gynecol Neonatal Nurs* 2003; **32**: 437–46.
34. Leeflang M. M., Deeks J. J., Gatsonis C., Bossuyt P. M. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; **149**: 889–97.
35. Mallett S., Deeks J. J., Halligan S., Hopewell S., Cornelius V., Altman D. G. Systematic reviews of diagnostic tests in cancer: review of methods and reporting. *BMJ* 2006; **333**: 413.
36. Doust J. A., Pietrzak E., Sanders S., Glasziou P. P. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *J Clin Epidemiol* 2005; **58**: 444–9.
37. Leeflang M. M., Scholten R. J., Rutjes A. W., Reitsma J. B., Bossuyt P. M. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol* 2006; **59**: 234–40.
38. Lijmer J. G., Mol B. W., Heisterkamp S., Bonsel G. J., Prins M. H., Van Der Meulen J. H. *et al.* Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; **282**: 1061–6.
39. Saunders J. B., Aasland O. G., Babor T. F., de la Fuente J. R., 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.
40. Bush K., Kivlahan D. R., McDonell M. B., Fihn S. D., Bradley K. A. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998; **158**: 1789–95.
41. Gual A., Segura L., Contel M., Heather N., Colom J. Audit-3 and audit-4: effectiveness of two short forms of the alcohol use disorders identification test. *Alcohol Alcohol* 2002; **37**: 591–6.
42. Mayfield D., McLeod G., Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. *Am J Psychiatry* 1974; **131**: 1121–3.
43. Selzer M. L., Vinokur A., van Rooijen L. A self-administered Short Michigan Alcoholism Screening Test (SMAST). *J Stud Alcohol* 1975; **36**: 117–26.
44. Sokol R. J., Martier S. S., Ager J. W. The T-ACE questions: practical prenatal detection of risk-drinking. *Am J Obstet Gynecol* 1989; **160**: 863–8, discussion 868–70.
45. Russell M., Bigler L. Screening for alcohol-related problems in an outpatient obstetric–gynecologic clinic. *Am J Obstet Gynecol* 1979; **134**: 4–12.
46. Bottoms S., Martier S., Sokol R. Refinements in screening for risk drinking in reproductive-aged women: the 'NET' results. *Alcohol Clin Exp Res* 1989; **13**: 339.
47. Grant B. F., Dawson D. A., Stinson F. S., Chou P. S., Kay W., Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend* 2003; **71**: 7–16.
48. Sobell L. C., Sobell N. B. Timeline followback: a technique for assessing self-reported ethanol consumption. In: Litten R. Z., Allen J. P., editors. *Measuring Alcohol Consumption: Psychological and Biochemical Method*. New Jersey: Humana Press Inc; 1992, p. 41–72.
49. Sobell L. C., Sobell M. B., Riley D. M., Schuller R., Pavan D. S., Cancilla A. *et al.* The reliability of alcohol abusers' self-reports of drinking and life events that occurred in the distant past. *J Stud Alcohol* 1988; **49**: 225–32.
50. Whiting P., Rutjes A. W., Reitsma J. B., Bossuyt P. M., Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25.
51. Chang G., Wilkins-Haug L., Berman S., Goetz M. A., Behr H., Hiley A. Alcohol use and pregnancy: improving identification. *Obstet Gynecol* 1998; **91**: 892–8.
52. Dawson D. A., Grant B. F., Stinson F. S., Zhou Y. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. *Alcohol Clin Exp Res* 2005; **29**: 844–54.
53. Russell M., Martier S. S., Sokol R. J., Mudar P., Jacobson S., Jacobson J. Detecting risk drinking during pregnancy: a comparison of four screening questionnaires. *Am J Public Health* 1996; **86**: 1435–9.
54. Dawson D. A., Das A., Faden V. B., Bhaskar B., Krulewitch C. J., Wesley B. Screening for high- and moderate-risk drinking during pregnancy: a comparison of several TWEAK-based screeners. *Alcohol Clin Exp Res* 2001; **25**: 1342–9.
55. Budd K. W., Ross-Alaolmolki K., Zeller R. A. Two prenatal alcohol use screening instruments compared with a physiologic measure. *J Obstet Gynecol Neonatal Nurs* 2000; **29**: 129–36.
56. Bellati U., Gentile C., Tiboni G. M., Liberati M., Rosati V. [Alcoholism and pregnancy]. *Minerva Ginecol* 1994; **46**: 409–12.
57. Gale T. C., White J. A., Welty T. K. Differences in detection of alcohol use in a prenatal population (on a northern plains Indian reservation) using various methods of ascertainment. *South Dakota J Med* 1998; **51970**: 235–40.
58. Coulton S., Drummond C., James D., Godfrey C., Bland J. M., Parrott S. *et al.* Opportunistic screening for alcohol use disorders in primary care: comparative study. *BMJ* 2006; **332**: 511–7.
59. Fleming M. F., Mundt M. P., French M. T., Manwell L. B., Stauffacher E. A., Barry K. L. Brief physician advice for problem drinkers: long-term efficacy and benefit–cost analysis. *Alcohol Clin Exp Res* 2002; **26**: 36–43.
60. Bradley K. A., Boyd-Wickizer J., Powell S. H., Burman M. L. Alcohol screening questionnaires in women: a critical review. *JAMA* 1998; **280**: 166–71.
61. Aertgeerts B., Buntinx F., Kester A. The value of the CAGE in screening for alcohol abuse and alcohol dependence in

general clinical populations: a diagnostic meta-analysis. *J Clin Epidemiol* 2004; **57**: 30–9.
62. Koren G., Hutson J., Gareri J. Novel methods for the detection of drug and alcohol exposure during pregnancy: implications for maternal and child health. *Clin Pharmacol Ther* 2008; **83**: 631–4.
63. Dawson D. A. Methodological issues in measuring alcohol use. *Alcohol Res Health* 2003; **27**: 18–29.
64. Greenfield T. K., Kerr W. C. Alcohol measurement methodology in epidemiology: recent advances and opportunities. *Addiction* 2008; **103**: 1082–99.
65. Bossuyt P. M., Reitsma J. B., Bruns D. E., Gatsonis C. A., Glasziou P. P., Irwig L. M. *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Biochem* 2003; **36**: 2–7.
66. Song F., Khan K. S., Dinnes J., Sutton A. J. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002; **31**: 88–95.
67. Nava-Ocampo A. A., Velazquez-Armenta Y., Brien J. F., Koren G. Elimination kinetics of ethanol in pregnant women. *Reprod Toxicol* 2004; **18**: 613–7.
68. Ng P. K., Cottle M. K., Baker J. M., Johnson B., Van Muyden P., Van Petten G. R. Ethanol kinetics during pregnancy. Study in ewes and their fetuses. *Prog Neuropsychopharmacol Biol Psychiatry* 1982; **6**: 37–42.
69. Alvik A., Heyerdahl S., Haldorsen T., Lindemann R. Alcohol use before and during pregnancy: a population-based study. *Acta Obstet Gynecol Scand* 2006; **85**: 1292–8.
70. Bobo J. K., Klepinger D. H., Dong F. B. Changes in the prevalence of alcohol use during pregnancy among recent and at-risk drinkers in the NLSY cohort. *J Womens Health (Larchmt)* 2006; **15**: 1061–70.
71. Lall A. Neonatal abstinence syndrome. *Br J Midwifery* 2008; **16**: 220–6.
72. Manwell L. B., Fleming M. F., Mundt M. P., Stauffacher E. A., Barry K. L. Treatment of problem alcohol use in women of childbearing age: results of a brief intervention trial. *Alcohol Clin Exp Res* 2000; **24**: 1517–24.
73. Waterson E. J., Murray-Lyon I. A. Preventing fetal alcohol effects: a trial of three methods of giving information in the antenatal clinic. *Health Educ Res* 1990; **5**: 53–61.
74. Goransson M., Magnusson A., Heilig M. Identifying hazardous alcohol consumption during pregnancy: implementing a research-based model in real life. *Acta Obstet Gynecol Scand* 2006; **85**: 657–62.
75. Chang G., McNamara T. K., Orav E. J., Koby D., Lavigne A., Ludman B. *et al.* Brief intervention for prenatal alcohol use: a randomized trial. *Obstet Gynecol* 2005; **105**: 991–8.
76. Chang G., Wilkins-Haug L., Berman S., Goetz M. A. Brief intervention for alcohol use in pregnancy: a randomized trial. *Addiction* 1999; **94**: 1499–508.
77. Handmaker N. S., Miller W. R., Manicke M. Findings of a pilot study of motivational interviewing with pregnant drinkers. *J Stud Alcohol* 1999; **60**: 285–7.
78. O'Connor M. J., Whaley S. E. Brief intervention for alcohol use by pregnant women. *Am J Public Health* 2007; **97**: 252–8.
79. Reynolds K. D., Coombs D. W., Lowe J. B., Peterson P. L., Gayoso E. Evaluation of a self-help program to reduce alcohol consumption among pregnant women. *Int J Addict* 1995; **30**: 427–43.

