Evidence Table 5. Data for KQ 2 from systematic reviews

| Author, Year  Funding Source  Aim of Review  Studies included in Review | Inclusion/Exclusion Criteria | Screening Instruments | Outcomes | Conclusions  Limitations |
| --- | --- | --- | --- | --- |
| Berks, 200839  Other or NR  Not explicit: to determine appropriate alcohol screening tests in older adult (60+) population  Number of Studies  9 (8 analyzed together with 1 separate)  Number of patients  6,353 | Inclusion   * English studies focusing on screening in 60+ year olds * Patients presenting to primary care   Exclusion   * Excluded if gave average age but no cutoff, no gold-standard comparator, allowed test result to influence decision to perform gold-standard, if included data insufficient for calculation of sensitivity and specificity | CAGE  MAST  MAST-G  SMAST  AUDIT  AUDIT  AUDIT-C  ARPS  shARPS  SMAST-G | CAGE for abuse/dependence: >=1 sens: 79-88%, spec: 56-88%  CAGE for hazardous/excessive: >=1 sens: 31-60%, spec: 92-100%  >=2 sens: 14-39%, spec: 97-97.1%  MAST for abuse/dependence: >=4: sens 91%, spec 84%  >=3 sens: 64-97%, spec: 67-79%  MAST-G for abuse/dependence: cutoff>=5: sens 70-91%, spec 81-84%  2 studies compared MAST with CAGE: one showed MAST slightly better, other showed CAGE was better  SMAST for heavy drinking: cutoff >=2: sens 48%, spec 100%  AUDIT for abuse/dependence: >=8: sens 33%, spec 91%  AUDIT for hazardous: >=8: sens 67%, spec 95%  AUDIT-C for hazardous: >=3: sens 100%, spec 81%  Moore 2002:  ARPS for hazardous: unclear cutoff: sens 93%, spec 63%  shARPS for hazardous: unclear cutoff: sens 92%, spec 51%  AUDIT for hazardous: >=8 sens 28%, spec 100%  SMAST-G for hazardous: >=2 sens 52%, spec 96% | **Conclusions**   * AUDIT appears superior to others for hazardous (AUDIT-C as good or better than AUDIT), CAGE appears better for abuse/dependence screening * If age-specific definitions of hazardous/harmful needed then ARPS and variations are superior.   **Limitations**   * Narrative synthesis of included studies. No meta-analysis conducted. |

Evidence Table 5. Data for KQ 2 from systematic reviews (continued)

| Author, Year  Funding Source  Aim of Review  Studies included in Review | Inclusion/Exclusion Criteria | Screening Instruments | Outcomes | Conclusions  Limitations |
| --- | --- | --- | --- | --- |
| Berner, 200740  Government  Assess diagnostic accuracy of AUDIT for detection of at risk drinking  Number of Studies  23 (27 articles) included in review, 19 for meta-analysis  Number of patients  25,940 total, 23,190 in meta-analysis | Inclusion   * AUDIT compared with reference standard of at-risk consumption assessed by quantity/frequency and/or heavy episodic drinking frequency * Used 10 item AUDIT * Compared with same reference in all subjects regardless of result * AUDIT not used as reference standard * Rreference test performed within 1 month * AUDIT performed by >50% of participants   Exclusion   * NA | AUDIT | AUDIT cutoff 8 points:   * Primary care (8 studies): sens 0.31-0.89, spec 0.83-0.96, pooled LR+: 6.78, LR-: 0.40, OR: 18.3 * Inpatient: se 0.93, sp 0.94, LR+: 15.07, LR-: 0.08, OR: 198.0   ED:   * SE: 0.72 * SP: 0.88 * LR+: 6.09 * LR-: 0.32 * OR: 19.1   University:   * SE: 0.82 * Spec: 0.88 * LR+: 3.73 * LR-: 0.23 * OR: 15.99   Older adults:   * SE: 0.55-0.83 * SP: 0.96 (pooled) * LR+: 20.11 * LR-: 0.33 * OR: 59.8 | **Conclusions**   * AUDIT use restricted to primary care, inpatients, older adults   **Limitations**   * Large heterogeneity in studies partly explained by setting, thus could not pool 17 studies together |

Evidence Table 5. Data for KQ 2 from systematic reviews (continued)

| Author, Year  Funding Source  Aim of Review  Studies included in Review | Inclusion/Exclusion Criteria | Screening Instruments | Outcomes | Conclusions  Limitations |
| --- | --- | --- | --- | --- |
| Bradley, 199841  Government  Describe performance of alcohol screening questionnaires for heavy drinking/abuse/dependence in females in general clinical populations in the U.S.  Number of Studies  9 (13 articles)  Number of patients  12,407 total (includes females and males)  About 10,883 women | Inclusion   * Studies with women comparing brief alcohol screening with valid standard for heavy drinking/abuse/ dependence in U.S. general clinical population * Screening questionnaires with 10 or less items * Limited to studies in U.S.   Exclusion   * Studies outside of U.S. or not published in English * Excluded nonclinical and special clinical populations * Studies without valid comparison group * Excluded data regarding screening for ICD harmful use * Excluded studies using self-administered questions for estimates of typical quantity/frequency as reference standard | CAGE  TWEAK  AUDIT  T-ACE  BMAST  NET | CAGE for abuse/dependence: >=2: auROC 0.84-0.92 in mainly black populations, se 0.38-0.50 in mainly white populations  TWEAK and AUDIT for abuse/dependence: se: <0.80, auROC 0.87-0.93  AUDIT for heavy drinking: auROC 0.87  TWEAK and T-ACE heavy drinking before pregnancy: auROC 0.84-0.87 in black OB patients  No pooling of data due to subjective heterogeneity (but not statistically assessed)  Primary care only:  CAGE >=2 for abuse/dependence in 80% black population: se 0.74, sp 0.93  CAGE >=2 for abuse/dependence in 93% white population: se 0.38, sp 0.92  AUDIT for abuse/dependence: auROC 0.87-0.93  AUDIT for heavy drinking: auROC 0.86-0.87 | **Conclusions**   * CAGE, AUDIT, TWEAK performed best for identifying dependence in black women (TWEAK best for white women) and that AUDIT was the only screening test assessed for identifying heavy drinking in nonobstetric population but was effective * Brief screens may be less sensitive for abuse/dependence among women because consumption questions based on male drinking * Appears no statistical differences in performance based on auROC for females vs males * Alcohol screening performance may vary by ethnicity   **Limitations**   * Mentions heterogeneity but does not quantify |

Evidence Table 5. Data for KQ 2 from systematic reviews (continued)

| Author, Year  Funding Source  Aim of Review  Studies included in Review | Inclusion/Exclusion Criteria | Screening Instruments | Outcomes | Conclusions  Limitations |
| --- | --- | --- | --- | --- |
| Burns, 201042  Academic  Investigate performance of brief alcohol screening questionnaires to identify problem drinking in pregnant women  Number of Studies  5  Number of patients  6,724 | Inclusion   * Cohort/cross sectional studies comparing brief alcohol screening instruments with reference criteria using structured interviews to detect at-risk drinking/abuse/dependency in pregnant women receiving prenatal care * Included only brief screening questionnaires * Reference standard based on quantity/frequency from structured interview (AUDADIS or timeline follow-back) or clnical diagnoses from DSM or ICD-10   Exclusion   * Excluded case-control studies * Excluded studies that used methods other than structured interview as referent (biomarkers, self-administered questionnaires) | TWEAK  T-ACE  CAGE  NET  AUDIT  AUDIT-C  SMAST | At-risk drinking:   * T-ACE: se 0.69-0.88, sp 0.71-0.89 * TWEAK: se 0.71-0.91, sp 0.73-0.83 * AUDIT-C se 0.95, sp 0.85 * CAGE >=2: se 0.38-0.49, sp 0.92-0.93 * NET >=1: se 0.71, sp 0.86 * SMAST: se 0.11, sp 0.96 * T-ACE and TWEAK higher auROC vs CAGE and NET * TWEAK, T-ACE, AUDIT-C highest sensitivities for at-risk * T-ACE, TWEAK lower PPVs than AUDIT-C * CAGE and SMAST performed poorly vs. others for identifying at-risk   Abuse/dependence:   * AUDIT-C >=3: dependece: se 1, sp 0.71. AUD: se 0.96, sp 0.71 * AUDIT >=8: lifetime dependency performed poorly * AUDIT had higher auROC than T-ACE, SMAST | **Conclusions**   * T-ACE, TWEAK, AUDIT-C have promise for screening for prenatal at risk drinking and AUDIT-C may be helpful to identify dependency/abuse. * CAGE did not perform well. |

Evidence Table 5. Data for KQ 2 from systematic reviews (continued)

| Author, Year  Funding Source  Aim of Review  Studies included in Review | Inclusion/Exclusion Criteria | Screening Instruments | Outcomes | Conclusions  Limitations |
| --- | --- | --- | --- | --- |
| Fiellin, 200043  Multiple  Evaluate accuracy of screening methods for alcohol problems in primary care  Number of Studies  38  11 for at-risk/hazardous/harmful drinking  27 for abuse/dependence  Number of patients  NR | Inclusion   * Published in peer-reviewed journal * Studies in English * Primary care setting * Reported performance (sens/spec) of screening methods compared with a criterion standard (structured interview)   Exclusion   * Studies not in English or were performed outside of primary care * Studies that did not report performance of screening methods * Excluded reviews, letters, editorials * Excluded studies that did not have comparators | AUDIT and AUDIT variations  CAGE  MAST  2-question QF  General health screen  Clinical/lab indicators | At-risk/hazardous/harmful:   * AUDIT >=8 most effective for at-risk/hazardous/harmful: se 0.51-0.97, sp 0.78-0.96 * CAGE >=2 for at-risk/hazardous/harmful: se 0.14 - 0.84, sp 0.74-0.97 * SMAST >=2: se 0.68, sp 0.92 * Single question screen for problem drinking: se 0.62, sp 0.93 * CDT for heavy drinking: se 0.39-0.69, sp 0.29-0.81 * GGT for heavy drinking: se 0.77, sp 0.81 in one study but limited utility for MCV, AST, ALT   Abuse/dependence:   * CAGE most effective for abuse/dependence: se 0.43-0.94, sp 0.70-0.97 * CAGE >=2 for abuse/dependence: se 0.21-0.94, sp 0.77-0.97 * CAGE >=1 for abuse/dependence: se 0.60-0.71, sp 0.84-0.88 * AUDIT for abuse/dependence: se 0.33-0.93, sp 0.89-0.97 * SMAST >=2 for abuse/dependence: se 0.48-1, sp 0.85-0.97 * Cyr/Wartman: se 0.48-0.91, sp 0.76- 0.93 (vs MAST as referent) * Single question: se 0.40-0.70, sp 0.93-0.99 * TWEAK: se 0.75, sp 0.90 * quantity-frequency: se 0.20- 0.50, sp 0.87-0.97 based on cutoff * Alcohol Clinical Index: se 0.28, sp 0.86 * Health Screening Survey: se 0.78, sp 0.71 | **Conclusions**   * AUDIT was most effective for at-risk, hazardous, harmful * CAGE was most effective for abuse and dependence * Formal screening instruments performed better than QF questions   **Limitations**   * Authors state few studies performed comparisons among multiple screening instruments |

Abbreviations: ALT = alanine transaminase; ARPS = Alcohol-Related Problems Survey; AST = aspartate transaminase; AUDADIS = Alcohol Use Disorder and Associated Disabilities Interview Schedule; AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test - Consumption; auROC = area under receiving operator characteristic; BMAST = brief Michigan Alcohol Screening Test; CAGE = Cut down, Annoyed, Guilty, Eye opener questionnaire; CDP = carbohydrate deficient transferrin; DSM = *Diagnostic and Statistical Manual of Mental Disorders*; ED = emergency department; GGT = gamma glutamyl transferase; ICD = International Classification of Diseases; LR = likelihood ratio; MAST = Michigan Alcoholism Screening Test; MAST-G = Michigan Alcoholism Screening Test – geriatric version; NET = Normal drinker, Eye opener, Tolerance questionnaire; OR = odds ratio; NR = not reported; PC = primary care; QF = quantity/frequency; RCT = randomized controlled trial; se = sensitivity; shARPS = shortened Alcohol-Related Problems Survey; SMAST = short Michigan Alcoholism Screening Test; SMAST-G = short Michigan Alcoholism Screening Test – geriatric version; sp = specificity; T-ACE = Tolerance, Annoyed, Cut-down, Eye-opener questionnaire; TWEAK = Tolerance; Worried; Eye opener; Amnesia; Kut down