Neuroscience and Biobehavioral Reviews 124 (2021) 370–385

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Consistency between self-reported alcohol consumption and biological markers among patients with alcohol use disorder – A systematic review

Dorthe Grüner Nielsen [a,*], Kjeld Andersen [a,b,c], Anette Søgaard Nielsen [a,b], Carsten Juhl [d,e], Angelina Mellentin [a,b,f]

[a] Unit for Clinical Alcohol Research, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark
[b] BRIDGE, Brain Research, Inter-Disciplinary Guided Excellence, Department of Clinical Research, University of Southern Denmark, Odense, Denmark
[c] Open, Odense Patient Explorative Data Network, Odense University Hospital, Odense, Denmark
[d] Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Gentofte, Herlev, Denmark
[e] Department of Sport Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
[f] Tele-Psychiatric Center, Region of Southern Denmark, Odense, Denmark

ARTICLE INFO

Keywords:
Alcohol use disorder
Alcohol consumption
Biomarker
Self-report
Treatment

ABSTRACT

Background: No systematic review has yet examined the consistency between self-reports of alcohol consumption and alcohol biomarkers among patients in treatment for alcohol use disorders (AUD). Therefore, we aimed to provide an overview of the consistency between self-reported alcohol intake and biomarkers among patients in treatment for AUD.

Methods: The electronical databases MEDLINE, PsycINFO, EMBASE, Cochrane Database of Systematic Reviews (CDSR) and CENTRAL were searched for all original studies that examined the validity of self-reported alcohol consumption using a biological marker in samples of patients with AUD. Eligible studies were included in a qualitative synthesis of the outcomes. Quality assessment was conducted with the quality assessment tool for Observational Cohort and Cross-sectional studies, developed by The National Heart, Lung and Blood Institute.

Results: The search identified 7672 hits, and 11 papers comprising 13 eligible studies were included. All the identified studies revealed inconsistencies between self-reporting and biomarkers. Under-reporting was the most common type of inconsistency across short-, intermediate- and long-term biomarkers. For short-term markers, under-reporting was indicated in 7 studies (n = 15–585) in a range from 5.5%–56.0% of the patients, and over-reporting in 2 studies (n = 34–65) in a range from 5.9%–74.1%. Only under-reporting was reported for intermediate-term, direct markers and was indicated in 2 studies (n = 18–54) in a range from 5.0%–50.0% of the patients. Although the results for long-term biomarkers were not reported consistently across the studies, under-reporting was indicated in 3 studies (n = 73–1580) in a range from 0.1%–40.0% of the patients, and over-reporting in 2 studies (n = 15–1580) in a range from 13.0%–70.6%. Correlations between self-reported alcohol consumption and biological markers were strongest for the intermediate-term direct markers, ranging from moderate to strong. For short-term and long-term markers, the correlations were mostly weak. Most of the studies were quality rated as fair.

Conclusion: The findings indicate that inconsistency between self-reported alcohol consumption and biomarkers may occur in a considerable proportion of patients with AUD. However, further studies applying more sensitive, specific, and easily assessable biological markers are warranted to confirm this preliminary synthesis.

PROSPERO registration no.: CRD42018105308

* Corresponding author at: Unit for Clinical Alcohol Research, Institute of Clinical Research, University of Southern Denmark, J.B Winsløws Vej 18, 5000 Odense C, Denmark.
E-mail address: dgnielsen@health.sdu.dk (D. Grüner Nielsen).

https://doi.org/10.1016/j.neubiorev.2021.02.006
Received 22 June 2020; Received in revised form 29 January 2021; Accepted 3 February 2021
Available online 10 February 2021

1. Introduction and background

1.1. Rationale

Alcohol use disorders (AUD) have a major impact on the individual and society, and effective treatment is therefore important. Research is focused on optimizing treatment and a large body of studies examining pharmacological and psychological treatments have identified effective AUD treatments (Drummond et al., 2011; Reus et al., 2018). The main purpose of treatment is to achieve abstinence or to reduce drinking to moderate levels as recommended by the World Health Organization (WHO) (World Health Organization, 2000). The outcome of treatment is measured as the reduction in alcohol intake during treatment and/or at follow up assessment. Consumption is typically assessed using self-report measures, and a number of instruments have been developed and applied in clinical and research settings, such as the timeline follow-back (TLFB) (Litten et al., 1992) and the form 90 interview (Scheurich et al., 2005). These daily estimation techniques involve using a calendar to retrospectively identify daily alcohol consumption during a predetermined time period. They are applied as an outcome measure to determine increase/decrease in alcohol consumption after an intervention.

The psychometric properties of these instruments generally range from fair to excellent. The test-retest reliability has been found to be high across multiple populations of drinkers (Carey et al., 2004; Maisto et al., 2008; Scheurich et al., 2005; Sobell et al., 1996, 1979a; Tonigan et al., 1997), and there is evidence for both criterion and construct validity (Breslin et al., 2001; Carney et al., 1998; Del Boca et al., 1994; Grant et al., 1995; Miller, 1996; O'Farrell and Langenbucher, 1988; Roy et al., 2008; Scheurich et al., 2005; Searles et al., 2000; Sobell et al., 2003; Toll et al., 2006). Concurrent criterion validity has been estimated by correlating self-report data with other subjective sources of information such as collateral informants or official records (e.g. from hospitals and prisons). Findings from these validation studies indicate a high degree of validity for abstinent or light drinking periods, but the strength of the correlation between outcome measures decreases the heavier the drinking (Breslin et al., 2001; O'Farrell and Langenbucher, 1988; Scheurich et al., 2005). Regarding external construct validity, the TLFB has been compared to measures of drinking patterns such as the Quantity-Frequency Questionnaire (Carney et al., 1998; Grant et al., 1995; Searles et al., 2000; Toll et al., 2006), Drinker Profile (Grant et al., 1995), Quick Drinking Screen (Roy et al., 2008; Sobell et al., 2003), the Form-90 to the Lifetime Drinking History Interview (Scheurich et al., 2005), and the Alcohol Use Disorders Identification Test (Del Boca et al., 1994; Miller, 1996), yielding acceptable consistency. Although self-report measures of alcohol consumption have sound psychometric properties and are accepted in the research community, they have mostly been validated against other self-report measures. This may be problematic since several factors may impact self-reported alcohol consumption, e.g. poor episodic memory, other cognitive impairments, social desirability, potential or imagined consequences of reported consumption etc. (Jung and Namkoong, 2014; Le Berre et al., 2017; Welte and Russell, 1993). Further, even when applying collateral reports there is a risk that family members and staff may not be aware of the patient's drinking pattern, due to e.g. lack of continuous monitoring etc. (Connors and Maisto, 2003).

Another way of measuring alcohol consumption and gaining corroborating evidence for the validity of self-report is by using biomarkers, and not relying only on reported intake from either the patient, family or staff. Biomarkers of alcohol consumption are detectable in different biological materials, e.g. blood, breath, and urine (Andresen-Streichert et al., 2018; Ghosh et al., 2019). Biomarkers can be 1) direct markers or metabolites of alcohol, 2) indirect markers almost exclusively specific to alcohol consumption, or 3) indirect markers not specific to alcohol consumption and thus related to many other conditions as well. A direct marker or metabolite of alcohol cannot be produced without the presence of alcohol. Examples of direct markers or metabolites of alcohol, besides ethanol itself, are ethyl glucuronide (EtG), ethyl sulphate (EtS), fatty acid ethyl esters (FAEE), and phosphatidyl ethanol (PEth) (Andresen-Streichert et al., 2018; Luginbühl, 2016; Wozniak et al., 2017). Carbohydrate-deficient transferrin, (CDT), elevated 5-hydroxytryptophol/5-hydroxyindole-3-acetic acid (5-HTOL/5-HIAA) ratio and elevated 5-hydroxytryptophol/creatinine (5-HTOL/CREA) ratio are almost exclusively specific for alcohol consumption but are not direct metabolites of ethanol and may be elevated for other reasons than alcohol consumption. 5-HTOL and CREA may be affected by diet and muscle activity but this is compensated for using 5-HTOL/5-HIAA ratio instead (Carlsson et al., 1993; Ghosh et al., 2019). The only known cause of increased 5-HTOL/5-HIAA ratio besides alcohol itself is treatment with inhibitors of aldehyde dehydrogenase (ALDH) (Helander et al., 1996; Lin et al., 2020). CDT may seem elevated due to type 2 diabetes mellitus and anticonvulsants depending on analytical methods. Also, rare genetic variations, e.g. genetic D-variants of transferrin and congenital disorders of glycosylation (CDG) syndromes may lead to elevated levels even though there has been no excessive consumption of alcohol (Andresen-Streichert et al., 2018; Maenhout et al., 2013; Sillanaukee et al., 2001; Wu et al., 2020). Several conditions, e.g. some liver diseases and kidney and pancreas transplantations may cause false positive levels of CDT, and false elevated levels may also be related to analytical methods (Andresen-Streichert et al., 2018; De Feo et al., 1999; Helander et al., 2016; Sillanaukee et al., 2001). False negative results may also occur due to genetic B-variants of transferrin, high triglycerides and cirrhosis (De Feo et al., 1999; Fleming et al., 2004; Sillanaukee et al., 2001).

Some markers are affected not only by alcohol intake but also by various other conditions. Examples of such indirect non-specific biomarkers are mean corpuscular volume (MCV), gamma-glutamyl transferase (GGT), aspartate aminotransferase (AST), and alanine transaminase (ALT) (Andresen-Streichert et al., 2018; Conigrave et al., 2003; Kunutsor, 2016; Maenhout et al., 2013).

Depending on the biological sample matrix and the nature of the biomarker, biomarkers for alcohol consumption have different detection windows. An example is EtG which can be stored in e.g. both urine and hair causing very different detection times. Some biomarkers are detectable a few hours after consumption, others after months. In this review, biomarkers are described as short- intermediate- and long-term markers. Examples of short-term biomarkers with a window of detection of $\leq$ 24 h are ethanol in breath and urine and 5-HTOL/5-HIAA in urine. Intermediate-term biomarkers, e.g. EtS and EtG in urine, have a window of detection up to 48 h after single ethanol intake, and up to 130 h after excessive consumption. Also, FAEE in blood, which is detectable up to 72 h after alcohol consumption, is categorized as an intermediate-term marker. Examples of long-term biomarkers with a window of detection of weeks are the blood biomarkers GGT, MCV, AST, ALT, CDT and PEth in blood, and EtG and EtS in hair (Andresen-Streichert et al., 2018; Ghosh et al., 2019; Heier et al., 2016). For an overview of the biomarkers used in this review, see Table 1.

Since the majority of evidence-based AUD treatments are based on studies relying on self-report, a still emerging question is whether the validity of these instruments can be corroborated by biomarkers. To date, several studies have addressed this question by applying a variety of direct and indirect markers of alcohol intake among treatment seeking populations with AUD. However, to our best knowledge, no systematic review has yet been conducted to examine the consistency between self-reported alcohol consumption and biomarkers.

1.2. Objective

The objective of this study was to systematically review the empirical literature addressing consistency between self-reported alcohol intake and biomarkers among current and previous treatment-seeking patients with AUD.

Table 1
Biomarkers used in this review.

Biomarker (material)	Characteristics (detection time; type)	Cut off reported in the studies
Ethanol (breath)	(when alcohol is present in the blood; DI)(1)	0‰: (2–4); NR: (5)
Ethanol (urine)	(provides information on BAC when the urine was produced; DI) (1,6)	15 mg/mL: (7); NR: (8)
5-HTOL/5-HIAA (urine)	(up to 24 h; AEM) (9–12)	20 pmol/nmol: (8)
5-HTOL/CREA (urine)	(up to 24 h; AEM) (9–12)	30 pmol/nmol: (8)
EtS (urine)	(up to 48 h after single ethanol intake; DI)(13)	0.1 mg/L: (14)
EtG (urine)	(up to 48 h after single ethanol intake; DI) (13)	0 mg/mL: 0.5 mg/L (14)
AST	(normal levels are expected 2–4 weeks after alcohol cessation; MOC) (6)	NR: (5); Varying: (15)
ALT	(normal levels are expected 2–4 weeks after alcohol cessation; MOC)(6)	NR: (5); Varying: (15)
GGT	(normal levels are expected 2–6 weeks after alcohol cessation; MOC)(6)	NR: (5,16); Varying: (15); F < 30 U/L, M<46 U/L: (17); >28 U/L (19)
CDTect	(normal levels are expected 2–3 weeks after alcohol cessation; AEM)(16,18)	74 mg/mL: (8); M:< 17 U/L, F:< 25 U/L (15); M: 20 U/L (19)
%CDT	(normal levels are expected 2–3 weeks after alcohol cessation; AEM)(16,18)	<2.5%: (19)
MCV	(normal levels are expected 8–16 weeks after alcohol cessation; MOC)(6)	80–96 fl: (17)

Abbreviations: 5-HTOL/5-HIAA = 5-hydroxytryptophol/5-hydroxyindole-3-acetic acid; 5-HTOL/CREA = 5-hydroxytryptophol/creatinine; %CDT =percent carbohydrate deficient transferrin; AEM = almost exclusive marker of alcohol consumption; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BAC = blood alcohol concentration; CDTect = absolute value of carbohydrate deficient transferrin; DI = direct marker of alcohol consumption; EtG = ethyl glucuronide; Ets = ethyl sulphate; F = females; fl = femtoliter; GGT = gamma-glutamyl transferase; M = males; MCV = mean corpuscular volume; mg/L = milligram(s)/liter; mg/mL = milligram(s)/milliliter; MOC = marker affected by multiple conditions, excessive alcohol consumption included; NR = not reported; pmol/nmol = picomole(s)/nanomole; U/L = units/liter.

References: **1**. Maenhout TM, De Buyzere ML, Delanghe JR. Non-oxidative ethanol metabolites as a measure of alcohol intake. Clin Chim Acta. 2013;415:322–9.; **2.** Sobell MB, Sobell LC, VanderSpek R. Relationships among clinical judgment, self-report, and breath-analysis measures of intoxication in alcoholics. J Consult Clin Psychol. 1979;47(1):204–6.; **3.** Midanik L. Over-reports of recent alcohol consumption in a clinical population: a validity study. Drug Alcohol Depend. 1982;9(2):101–10.;
4. Erim Y, Bottcher M, Dahmen U, Beck O, Broelsch CE, Helander A. Urinary ethyl glucuronide testing detects alcohol consumption in alcoholic liver disease patients awaiting liver transplantation. Liver Transpl. 2007;13(5):757–61.; **5.** Whitford JL, Widner SC, Mellick D, Elkins RL. Self-report of drinking compared to objective markers of alcohol consumption. The American Journal of Drug and Alcohol Abuse. 2009;35(2):55–8.; **6.** Andresen-Streichert H, Müller A, Glahn A, Skopp G, Sterneck M. Alcohol Biomarkers in Clinical and Forensic Contexts. Deutsches Arzteblatt international. 2018;115(18):309.; **7.** Peachey JE, Kapur BM. Monitoring drinking behavior with the alcohol dipstick during treatment. Alcohol Clin Exp Res. 1986;10(6):663–6.;
8. Carlsson AV, Hiltunen AJ, Beck O, Stibler H, Borg S. Detection of relapses in alcohol-dependent patients: comparison of carbohydrate-deficient transferrin in serum, 5-hydroxytryptophol in urine, and self-reports. Alcohol Clin Exp Res. 1993;17(3):703–8.; **9.** Ghosh S, Jain R, Jhanjee S, Rao R, Mishra A. Alcohol Biomarkers and their Relevance in Detection of Alcohol Consumption in Clinical Settings. Int Arch Subst Abuse Rehabil. 2019;1(002).; **10.** Helander A, Beck O, Jones AW. Laboratory testing for recent alcohol consumption: comparison of ethanol, methanol, and 5-hydroxytryptophol. Clin Chem. 1996;42(4):618–24.; **11.** Beck O, Helander A, Carlsson S, Borg S. Changes in serotonin metabolism during treatment with the aldehyde dehydrogenase inhibitors disulfiram and cyanamide. Pharmacol Toxicol. 1995;77(5):323–6.; **12.** Johnson RD, Lewis RJ, Canfield DV, Dubowski KM, Blank CL. Utilizing the urinary 5-HTOL/5-HIAA ratio to determine ethanol origin in civil aviation accident victims. J Forensic Sci. 2005;50(3):670–5.; **13.** Heier C, Xie H, Zimmermann R. Nonoxidative ethanol metabolism in humans-from biomarkers to bioactive lipids: Non-oxidative Ethanol Metabolism in Humans. IUBMB life. 2016;68(12):916–23.; **14.** Dahl H, Hammarberg A, Franck J, Helander A. Urinary ethyl glucuronide and ethyl sulphate testing for recent drinking in alcohol-dependent outpatients treated with acamprosate or placebo. Alcohol Alcohol. 2011;46(5):553-7.
15. Babor TF, Steinberg K, Anton R, Del Boca F. Talk is cheap: measuring drinking outcomes in clinical trials. J Stud Alcohol. 2000;61(1):55–63.; **16.** Yoshino A, Kato M. Influence of social desirability response set on self-report for assessing the outcome of treated alcoholics. Alcohol Clin Exp Res. 1995;19 (6):1517–9.; **17.** Keso L, Salaspuro M. Comparative value of self-report and blood tests in assessing outcome among alcoholics. Br J Addict. 1990;85 (2):209–15.; **18.** Niemelä O, Alatalo P. Biomarkers of alcohol consumption and related liver disease. Scand J Clin Lab Invest. 2010;70(5):305–12.; **19.** Mundle G, Ackermann K, Gunthner A, Munkes J, Mann K. Treatment outcome in alcoholism - A comparison of self-report and the biological markers carbohydrate-deficient transferrin and gamma-glutamyl transferase. Eur Addict Res. 1999;5 (2):91–6.

2. Method

2.1. Protocol and registration

The present systematic review was reported according to the guidelines described by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009).

The inclusion criteria and analyses were specified in advance and registered in the international Prospective Register of Systematic Reviews (PROSPERO): registration no. CRD42018105308. The protocol can be accessed at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018105308.

2.2. Eligibility criteria

To be eligible for this systematic review, studies had to (1) be original studies using biomarkers to assess the concurrent criterion validity of self-reported data (case studies were excluded); (2) be published in English-language peer reviewed journals; (3) include adults $\geq$ 18 years; (4) include patients with a diagnosed AUD in current or previous treatment for AUD.

Exclusion criteria were (1) other substance use disorder except for nicotine as the presence of other substance use disorders would affect the patients' cognitive functions significantly; (2) severe psychiatric or neurological comorbidity (e.g. psychotic disorders, intellectual disabilities, dementia or brain damage) as these comorbidities would affect the patients' ability to provide reliable information.

2.3. Information sources

A systematic literature search was performed in the following bibliographic databases: MEDLINE (via PubMed), PsycINFO (via Ovid), EMBASE (via Ovid), CENTRAL, and The Cochrane Database of Systematic Reviews (CDSR), up to March 1st, 2020. There was no period of limitation on publish date.

2.4. Electronical literature search

A systematic search in the electronical databases was performed to identify all relevant studies. The search was based on key words that included subject headings and free text words describing four facets: 1) AUD, 2) self-reported alcohol consumption, 3) biological measure of alcohol consumption, and 4) treatment setting. For an overview of the search terms, see Table 2.

Table 2
Search string.

PubMed/Cochrane							
Facet 1		**Facet 2**	**Facet 3**		**Facet 4**		
AUD/alcohol consumption		**Self-reported alcohol consumption**	**Biological measure of alcohol consumption**		**Treatment setting**		
Acute alcoholic intox*	alcoholics [MeSH Terms]	patient reported outcome measures [MeSH Terms]	Alcohol analy*	fingernail*	alcohol rehabilitat*	outpatient department*	
Alcohol abstinen*	alcohol-induced disorder*	patient reported outcome measure*	alcohol blood level*	hair analy*	alcohol rehabilitation program*	outpatient treatment*	
alcohol abstinence [MeSH Terms]	alcohol-induced disorders [MeSH Terms]	patient reported outcome*	alcohol blood concentration*	hair [MeSH Terms]	ambulatory care*	outpatient*	
alcohol abus*	alcoholism*	patient-reported outcome*	alcohol monitor*	hair*	ambulatory care [MeSH Terms]	outpatients [MeSH Terms]	
alcohol addict*	alcoholism [MeSH Terms]	questionnaire*	alcohol sensor*	metabolite*	hospital patient*	Patient	
alcohol comsum*	alcohol-related disorders [MeSH Terms]	self apprais*	biochemical marker*	nail*	hospitalized patient*	patients [MeSH Terms]	
alcohol dependen*	alcohol-related disorder*	self assess*	biologic marker*	nails [MeSH Terms]	inpatient*	rehabilitati*	
alcohol drinking [MeSH Terms]	binge drink*	self evaluat*	biological marker*	plasma	inpatients [MeSH Terms]	rehabilitation [MeSH Terms]	
alcohol drink*	binge drinking [MeSH Terms]	self report [MeSH Terms]	biomarker*	plasma [MeSH Terms]	interven*	therapy	
alcohol drinking pattern*	chronic alcoholic intox*	self report*	biomarkers [MeSH Terms]	saliva*	outpatient care*	treatment*	
alcohol induced disorder*	ethanol abus*	self-apprais*	blood [MeSH Terms]	saliva [MeSH Terms]			
alcohol intake*	ethanol addict*	self-assessment [MeSH Terms]	blood*	serum			
alcohol intox*	ethanol dependen*	self-assess*	blood alcohol concentration*	serum [MeSH Terms]			
alcohol misus*	ethanol drink*	self-evaluat*	blood alcohol content [MeSH Terms]	sweat*			
alcohol poison*	ethanol intox*	self-report*	blood alcohol content*	sweat [MeSH Terms]			
alcohol related disorder*	ethanol intake*		blood plasma	toenail*			
alcohol us*	ethanol misus*		blood serum	transdermal*			
alcohol-dependen*	ethanol poison*		blood test*	plasmas			
alcoholic intox*	ethanol us*		breath alcohol analy*	serums			
alcoholic intoxication [MeSH Terms]	ethanol-dependen*		breath test*	transdermal alcohol analy*			
alcoholic*	problematic drink*		breath tests [MeSH Terms]	urine*			
			cutane*	urine [MeSH Terms]			
			finger nail*				
Embase							
Facet 1		**Facet 2**	**Facet 3**		**Facet 4**		
AUD/alcohol consumption		**Self-reported alcohol consumption**	**Biological measure of alcohol consumption**		**Treatment setting**		
acute alcoholic intox*	alcoholic*	patient reported outcome measure*	alcohol analyzer (subject heading)	finger nail (subject heading)	alcohol rehabilitation (subject heading)	outpatient care (subject heading)	
alcohol abstinence (subject heading)	alcohol-induced disorder*	patient reported outcome*	alcohol analy*	finger nail*	alcohol rehabilit*	outpatient care*	
alcohol abstinen*	Alcoholism (subject heading)	patient-reported outcome (subject heading)	alcohol blood level (subject heading)	fingernail*	alcohol rehabilitation program (subject heading)	outpatient department (subject heading)	
alcohol abuse (subject heading)	alcoholism*	patient-reported outcome*	alcohol blood level*	hair (subject heading)	alcohol rehabilitation program*	outpatient department*	
alcohol abus*	alcohol-related disorder*	questionnaire*	alcohol concentration*	hair*	ambulatory care (subject heading)	outpatient treatment*	
alcohol addict*	binge drink*	questionnaire (subject heading)	alcohol monitor*	hair analysis (subject heading)	ambulatory care*	patient (subject heading)	
alcohol consumption (subject heading)	binge drinking (subject heading)	self apprais*	alcohol sensor*	hair analy*	hospital patient (subject heading)	patient	
alcohol consum*	chronic alcoholic intox*	self assess*	biochemical marker (subject heading)	metabolite*	hospital patient*	patients	
alcohol dependen*	ethanol addict*	self evaluat*	biochemical marker*	metabolite (subject heading)	hospitalized patient*	rehabilitation (subject heading)	
alcohol drinking pattern*	ethanol consum*	self evaluation (subject heading)	biologic marker*	nail (subject heading)	inpatient*	rehabilitat*	
alcohol drink*	ethanol dependen*	self report (subject heading)	biological marker (subject heading	nail*	interven*	therapy	

(*continued on next page*)

Table 2 (*continued*)

alcohol induced disorder*	ethanol drink*	self report*	biological marker*	plasma	outpatient*	therapy (subject heading)
alcohol intake*	ethanol intake*	self-apprais*	biomarker*	plasma (subject heading)	outpatient (subject heading)	treatment*
alcohol intoxication (subject heading)	ethanol intox*	self-assess*	blood*	plasmas		
alcohol intox*	ethanol misus*	self-evaluat*	blood (subject heading)	saliva (subject heading)		
alcohol misus*	ethanol poison*	self-report*	blood alcohol concentration*	saliva*		
alcohol poison*	ethanol us*		blood alcohol content*	serum (subject heading)		
alcohol related disorder*	ethanol-dependen*		blood plasma	Serum		
alcohol us*	problematic drink*		blood serum	serums		
alcohol-dependen*			blood test*	sweat*		
alcoholic intox*			breath alcohol analyzer (subject heading)	sweat (subject heading)		
alcoholic intoxication (subject heading)			breath alcohol analy*	toenail*		
			breath analysis (subject heading)	transdermal*		
			breath analy*	transdermal alcohol analyzer (subject heading)		
			breath test*	transdermal alcohol analy*		
			cutane*	urine (subject heading)		
				urine*		
Psykinfo						
Facet 1		**Facet 2**	**Facet 3**		**Facet 4**	
AUD/alcohol consumption		**Self-reported alcohol consumption**	**Biological measure of alcohol consumption**		**Treatment setting**	
acute alcoholic intox*	alcoholic*	patient reported outcome measure*	alcohol analy*	Breath test*	alcohol rehabilit*	Outpatient department*
Acute Alcoholic Intoxication (subject heading)	alcohol-induced disorder*	patient reported outcome*	alcohol blood level*	Cutane*	Alcohol Rehabilitation (subject heading)	Outpatient Treatment (subject heading)
alcohol abstinen*	ALCOHOLISM (subject heading)	patient-reported outcome*	alcohol concentration*	finger nail*	alcohol rehabilitation program*	Outpatient treatment*
alcohol abus*	alcoholism*	questionnaires (subject heading)	alcohol monitor*	Fingernail*	Ambulatory care*	outpatient*
Alcohol Abuse (subject heading)	alcohol-related disorder*	questionnaire*	alcohol sensor*	HAIR (subject heading)	hospital patient*	outpatients (subject heading)
alcohol addict*	Binge Drinking (subject heading)	self apprais*	biochemical marker*	Hair*	Hospitalized Patients (subject heading)	patients (subject heading)
alcohol consum*	binge drink*	self assess*	biologic marker*	Hair analy*	hospitalized patient*	Patient
alcohol dependen*	chronic alcoholic intox*	self evaluat*	Biological Markers (subject heading)	METABOLITES (subject heading)	Inpatient*	rehabilitat*
alcohol drink*	Chronic Alcoholic Intoxication (subject heading)	self report*	biological marker*	Metabolite*	interven*	Rehabilitation (subject heading)
Alcohol Drinking Patterns (subject heading)	ethanol abus*	self-apprais*	BLOOD (subject heading)	nail*	intervention (subject heading)	Therapy
alcohol drinking pattern*	ethanol addict*	self-assess*	blood*	Plasma	outpatient care*	Treatment*
alcohol induced disorder*	ethanol consum*	self-evaluat*	Blood Alcohol Concentration (subject heading)	Plasmas		treatment (subject heading)
alcohol intake*	ethanol dependen*	Self-Evaluation (subject heading)	blood alcohol concentration*	SALIVA (subject heading)		
alcohol intox*	ethanol drink*	Self-Report (subject heading)	blood alcohol content*	Saliva*		
Alcohol Intoxication (subject heading)	ethanol intake*	self-report*	blood alcohol level*	Serum		
alcohol misus*	ethanol intox*		Blood Plasma (subject heading)	Serums		
alcohol poison*	ethanol misus*		Blood Serum (subject heading)	SWEAT (subject heading)		
alcohol related disorder*	ethanol poison*		blood test*	Sweat*		
alcohol us*	ethanol us*		blood plasma	toe nail*		
alcohol-dependen*	ethanol-dependen*		Blood serum	Transdermal*		
alcoholic intox*	problematic drink*		breath alcohol analy*			

(*continued on next page*)

Table 2 (*continued*)

	breath analy*	transdermal alcohol analy* urine (subject heading) Urine*

2.5. Study selection

One author (DGN) screened the titles and abstracts of the articles identified by the electronic searches and excluded obviously irrelevant studies. Subsequently, two authors (DGN and AIM) independently read the full text versions of all the remaining articles and excluded those that did not meet the inclusion criteria. The reference lists of the retrieved articles were checked for any further relevant citations. Articles identified as relevant were subjected to full analysis. Also, a search was performed in MEDLINE (via PubMed), PsycINFO (via Ovid) and EMBASE (via Ovid) for articles citing the included articles. Disagreements concerning the eligibility of studies were resolved through discussion and it was not necessary to involve a third researcher to assess eligibility.

2.6. Data collection

Two authors (DGN and AIM) extracted data from the included studies. We attempted to contact all corresponding authors of the included studies to achieve further data for the purpose of conducting a meta-analysis.

2.7. Data items

Data were extracted for study characteristics (author, year published, type of original research), demographic variables (sample size, age, gender), clinical variables (e.g. treatment setting: inpatient or outpatient), type of self-report instrument, type of biological marker, assessment time-points, type of statistical model (e.g. cross tabulation, correlation, regression), and outcome of the statistics modelling, that is, the relationship between self-report and biomarkers.

Correlation coefficients are interpreted as weak (< 0.40), moderate ($0.40 – 0.69$) or strong (> 0.70) (Schober et al., 2018).

2.8. Risk of bias in individual studies

Two authors (DGN, AIM) independently assessed and documented the overall quality of the included studies based on relevant criteria in the Quality of Assessment Tool for Observational Cohort and Cross-sectional studies, developed by the National Heart, Lung and Blood Institute (National Heart Lung and Blood Institute (NHLBI), Accessed 1/4/, 2020). Ten questions were evaluated with either "yes", "no", "not reported" (NR), "cannot determine" (CD) or "not relevant" (NA).

As no explicit guidelines exist, thresholds for the quality assessment rating were established by the quality assessment raters based on a general guidance published by the developers of the quality assessment tool. A score ≥ 9 was considered good, a score ≥ 6 and ≤ 8 was considered fair, and a score ≤ 5 was considered poor.

2.9. Summary measures and qualitative synthesis of results

The outcome is defined as the degree of consistency between self-report and biomarkers of alcohol consumption in the same patient population. Preferably the outcomes are reported using adequate statistical models such as cross tabulation analysis (frequencies/percentages), correlation coefficients or regression coefficients.

Several attempts were made to obtain the original raw data of the included studies to perform a meta-analysis or at least have access to more data. For several reasons (e.g. the raw data no longer exist, or it was not possible to establish contact with corresponding authors) it was not possible to retrieve enough data to perform a meta-analysis.

3. Results

3.1. Study selection

Fig. 1 illustrates the study selection process in the form of a flow chart.

The literature search resulted in 7672 studies. Manual searches of the reference lists in the included studies and of articles which have cited the included studies did not yield any additional studies. After removing duplicates, a total of 5946 studies remained. The titles and abstracts of these studies were subsequently assessed leading to the exclusion of studies. Potentially eligible studies ($n = 277$) were reviewed in more detail. 266 studies were excluded for the following reasons: 45 studies were not original studies, 21 studies did not describe biomarkers or self-report, 7 studies did not comprise an adult population, 96 studies did not report exclusively on patients with AUD, 64 studies did not include alcohol treatment, and 33 studies had the wrong outcomes. A total of 11 papers comprising 13 studies were found to be eligible.

3.2. Study characteristics, measures, and outcomes

An overview of the study characteristics, measures, analyses and outcomes is provided in Table 3.

3.2.1. Study characteristics

Most of the 13 studies had a longitudinal design, but cross-sectional and randomized controlled designs were also included. The sample sizes varied from 15 to 1580, and most patients were males in their thirties or forties. In the majority of studies, AUD was diagnosed according to DSM III, DSM III – R or DSM IV criteria (American Psychiatric Association, 1980, 1987, 1994). Other studies only reported an AUD diagnosis e.g. alcohol dependence.

Five studies (Babor et al., 2000; Carlsson et al., 1993; Dahl et al., 2011; Peachey and Kapur, 1986; Yoshino and Kato, 1995) reported results on outpatients, three studies (Keso and Salaspuro, 1990; Sobell et al., 1979b; Whitford et al., 2009) reported results on inpatients, two studies (Erim et al., 2007; Mundle et al., 1999) reported results from both inpatient and outpatient settings, and three studies (Midanik, 1982; Sobell et al., 1979b) did not report the treatment setting.

Data were obtained at the following time-points: pre-treatment (at admission), during treatment, post-treatment (at the end of treatment), and follow-up (time after treatment). Three studies (Midanik, 1982; Sobell et al., 1979b) collected data at pre-treatment only, one study (Dahl et al., 2011) at pre- and post-treatment (three weeks after admission), and one study at pre-treatment and 12 months follow-up (12 months after discharge) (Babor et al., 2000). In three studies, data were obtained regularly during treatment (Carlsson et al., 1993; Erim et al., 2007; Peachey and Kapur, 1986). One study (Mundle et al., 1999) collected data once during treatment, and five studies included a follow-up assessment (Keso and Salaspuro, 1990; Sobell et al., 1979b; Whitford et al., 2009; Yoshino and Kato, 1995), with a time span ranging from one month to approximately three years.

Two studies (Dahl et al., 2011; Whitford et al., 2009) reported that patients were compensated for participation. One study (Babor et al., 2000) reported that the participants were not compensated, and the remaining studies did not report if compensation was provided.

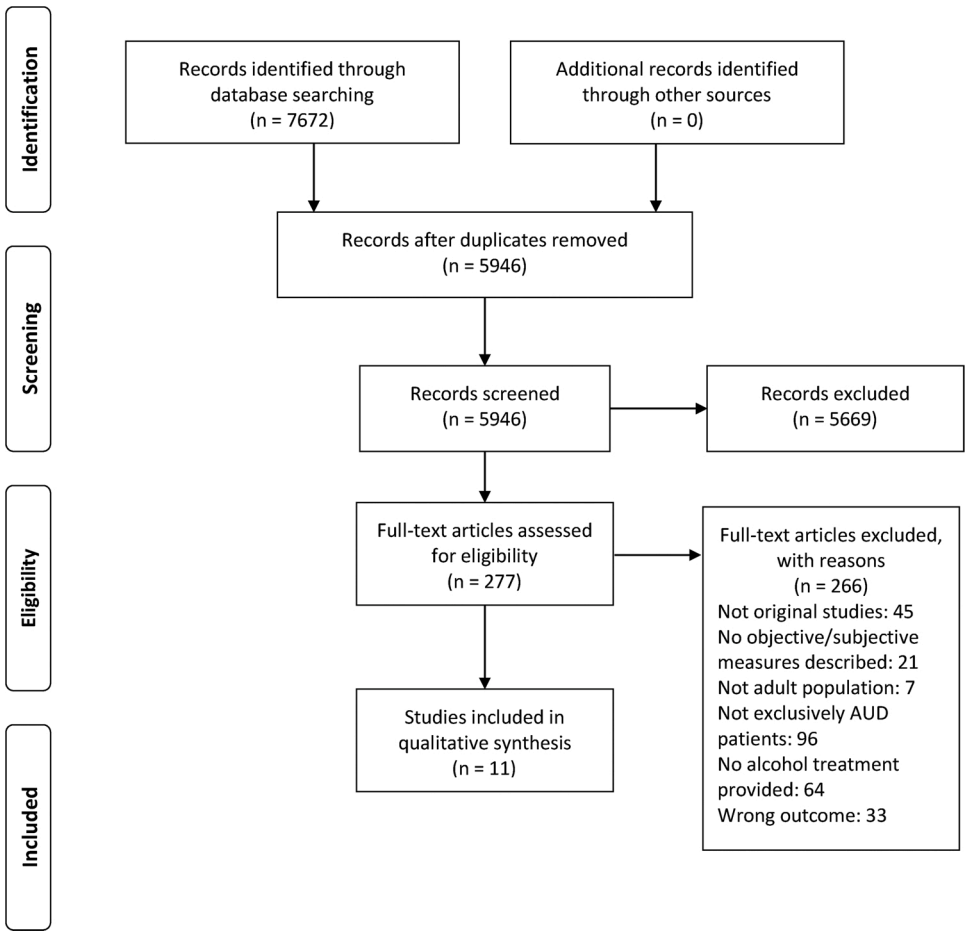

Fig. 1. Study selection.

3.2.2. Study measures

Regarding self-report of alcohol consumption, only four studies applied psychometrically validated measures. These measures were either daily estimation methods such as the TLFB (Dahl et al., 2011) and Form-90 (Babor et al., 2000) or quantity-frequency methods such as the simplified Khavari alcohol test (Keso and Salaspuro, 1990) and the addiction severity index (Whitford et al., 2009). Standardized assessment instruments developed by the research team were applied in another six studies (Midanik, 1982; Mundle et al., 1999; Peachey and Kapur, 1986; Sobell et al., 1979b; Yoshino and Kato, 1995), and the remaining studies simply asked patients about alcohol consumption in a non-standardized manner (Carlsson et al., 1993; Erim et al., 2007; Sobell et al., 1979b).

Regarding biomarkers, short-term markers were used in 8 studies (Carlsson et al., 1993; Erim et al., 2007; Midanik, 1982; Peachey and Kapur, 1986; Sobell et al., 1979b; Whitford et al., 2009), intermediate-term markers in two studies (Dahl et al., 2011; Erim et al., 2007) and long-term markers in 6 studies (Babor et al., 2000; Carlsson et al., 1993; Keso and Salaspuro, 1990; Mundle et al., 1999; Whitford et al., 2009; Yoshino and Kato, 1995).

Short-term markers were ethanol in breath and urine as well as 5-HTOL/5-HIAA and 5-HTOL/CREA in urine. Six studies (Erim et al., 2007; Midanik, 1982; Sobell et al., 1979b; Whitford et al., 2009) used breath ethanol as a marker, two studies (Carlsson et al., 1993; Peachey and Kapur, 1986) used urine ethanol, and one study used urine 5-HTOL/5-HIAA or 5-HTOL/CREA (Carlsson et al., 1993).

Intermediate-term markers were urine EtG and EtS. Urine EtG was evaluated in two studies (Dahl et al., 2011; Erim et al., 2007), and urine EtS in one study (Dahl et al., 2011).

Long-term markers consisted of the blood biomarkers GGT, MCV, AST, ALT, and CDT. GGT was evaluated in five studies (Babor et al., 2000; Keso and Salaspuro, 1990; Mundle et al., 1999; Whitford et al., 2009; Yoshino and Kato, 1995), AST and ALT in two studies (Babor et al., 2000; Whitford et al., 2009), CDT in three studies (CDTect in two studies (Babor et al., 2000; Carlsson et al., 1993), %CDT and CDTect in one study (Mundle et al., 1999)), and MCV in one study (Keso and Salaspuro, 1990),.

3.2.3. Study analysis

Twelve studies analyzed data by means of descriptive statistics, consisting of cross tabulation analysis (seven studies) (Erim et al., 2007; Keso and Salaspuro, 1990; Mundle et al., 1999; Peachey and Kapur, 1986; Sobell et al., 1979b), correlation analysis (one study), (Yoshino and Kato, 1995) or both (four studies) (Babor et al., 2000; Carlsson et al., 1993; Dahl et al., 2011; Midanik, 1982). One study (Whitford et al., 2009) applied inferential statistics, consisting of regression modelling.

3.3. Synthesis of results (outcomes)

3.3.1. Short-term markers

Regarding breath ethanol as a biological marker, inconsistency between self-report and blood alcohol concentration (BAC) was observed in 5.5%–74.1% of the patients (Erim et al., 2007; Midanik, 1982; Sobell et al., 1979b) by means of cross-sectional analysis. Discrepancy tending towards under-reporting was seen in 5.5%–56.0% of the patients (Erim et al., 2007; Midanik, 1982; Sobell et al., 1979b), and towards over-reporting in 23.1%–74.1% (Midanik, 1982).

By means of different elimination rates of alcohol varying between

Table 3
Study characteristics, measures, analysis and outcomes.

Study (year)	Study Type	N[1)]	Age (mean years)	Sex M/F (N)	Setting	Compensation provided	Self-report assessment (who obtained data on self-report)	Timespan for self-report	Biomarker (s) (biologic material)	Time-point for self-report and biomarker	Statistics & general findings
		Short-term markers ≤ 24 h									
Sobell (1979) 1	Long	40	43.7	31/9	I	NR	A (NR)	The preceding 48 h + time of last drink	Ethanol (breath)	At 1,3, & 5 months FUAD	CTA: consistent = 44%–55%
2	Long	585	41.3	493/92	NR	NR	* (research assistant)	Today and yesterday	Ethanol (breath)	Pre-t	CTA: consistent = 44%–55%
3	Long	79	39.4	75/4	NR	NR	* (research assistant)	Today and yesterday	Ethanol (breath)	Pre-t	CTA: consistent = 14.3%–55%[2]
Midanik (1982)	CS	68[3]	36.1	61/7	NR	NR	* (counsellor or researcher)	The preceding 24 h	Ethanol (breath)	Pre-t	CTA: consistent = 43.1%–60% [4a,c] - 11.11%–33.3%[4b,5c]; ur = 6.2%–18.5%[4a,c] - 14.8%–44.4%[4b,c]; or = 23.1%–56.9% [4a,c] - 29.6%–74.1% [4b,c] PC: (r;p) (0.347–0.440; **)[4a)]; (0.208–0.246; NS)[4b)]
Carlsson (1993)	Long	15	40.3	15/0	O	NR	A (treatment team members)	3 times/ week, since last visit	Ethanol 5-HTOL (urine, U)	Daily (dt)	CTA: **Ethanol:** ur (abstinence) = 7%; **5-HTOL:** ur (abstinence) = 47% **SP: 5-HTOL (r;p) = (**0.23;***)
Peachey (1986)	Long	34	40.2	28/6	O	NR	* (by mail)	Daily	Ethanol (urine, U)	Daily (dt)	CTA: ur = 41%; or = 5.9%
Erim (2007)	Long	18	51	9/9	I,O	NR	A (nurses)	Since inter-vention start	Ethanol (breath)	Every 2nd week dt	CTA: ur(abstinence) = 5.5%;
Whitford (2009)	Long	79	NR	NR	I	Y	ASI (NR)	The preceding 30 days	Ethanol (breath)	6 & 12 months FUAD	LR: ethanol was not a predictor of self-reported abstinence.
		Intermediate-term markers ≤ 48 h after single ethanol consumption									
Erim (2007	Long	18	51	9/9	I,O	NR	A (nurses)	During inter-vention time	EtG (urine, S)	Every 2nd week (dt)	CTA: ur (abstinence) = 50%
Dahl (2011)	RCT	54[5]	50	30/26	O	Y	TLFB (study coordinator)	The previous 3 days	EtG, EtS (urine, S/U: NR)	Pre-t & Pt (3 weeks aa)	CTA: EtS & EtG: ur (abstinence) = 5% PC (r;p) AI3 & EtG: (0.66;***); AI3 & EtS: (0.72;***)
		Long-term markers (weeks)									
Keso (1990)	Long	73	42.2	60/13	I	NR	Simplified Khavari alcohol test (NR)	The preceding 2 months	GGT MCV	8 months FUAD	CTA: ur = 40%
Carlsson (1993)	Long	15	40.3	15/0	O	NR	A (treatment team members)	3 times/ week, since last visit	CDTect	Within two weeks aa	CTA: or = 13%
Yoshino (1995)	CS	71	51.9	56/15	O	N	* (researchers)	The preceding 6 months	GGT	2–3 years FUAD	PC = (r;p) FP6M & GGT: (0.47;**); AAP6M & GGT: (0.61; **)
Mundle (1999)	Long	144	40.4	144/0	I,O	NR	* (NR)	The previous 3 weeks	GGT %CDT/ CDTect	7.5 months aa (dt)	CTA: ur (controlled drinking) = **GGT:** 9%; **CDT:** 7%; **Both markers:** 15%
Babor (2000)	RCT	1726[6]	40.2	1307/419	O	N	Form 90 (research assistants)	The preceding 3 months	GGT, AST, ALT, CDTect (CDTect: 12 months only)	Pre-t & 12 months FUAD	CTA: **Baseline:** consistent = GGT: 39.7%; AST: 29.3%; ALT: 30.2%; ur = GGT: 0.1%; AST: 0.1%; ALT: 0.1%; or = GGT: 60.2%; AST: 70.6%; ALT: 69.7%; **15 months follow-up:** consistent = GGT: 51.6%; AST: 46.3%; ALT: 42.6%; CDTect: 55.6%; ur = GGT:3.7%; AST: 2.9%;

(*continued on next page*)

Table 3 (*continued*)

Study (year)	Study Type	N[1)]	Age (mean years)	Sex M/F (N)	Setting	Compensation provided	Self-report assessment (who obtained data on self-report)	Timespan for self-report	Biomarker (s) (biologic material)	Time-point for self-report and biomarker	Statistics & general findings
											ALT: 3.7%; CDTect: 5.3%; or = GGT: 44.4%; AST: 50.8%; ALT: 53.7%; CDTect: 39.1%; PC[7]
Whitford (2009)	Long	79	NR	NR	I	Y	ASI (NR)	The previous 30 days	GGT AST ALT	6 & 12 months FUAD	LR: GGT, AST, ALT were not predictors of self-reported abstinence.

Abbreviations: * = patients were asked about alcohol consumption in a standardized manner; ** = p<0.01; *** = p<0.001; 5-HTOL/5-HIAA = 5-hydroxytryptophol/5-hydroxyindole-3-acetic acid; 5-HTOL/CREA = 5-hydroxytryptophol/creatinine; %CDT = percent carbohydrate deficient transferrin; A = patients were asked about alcohol consumption in a non- standardized manner; aa = after admission; adm = at admission; AAC = Annual alcohol consumption; AA12 = alcohol consumption the previous 12 months; AAP6M = amount of alcohol used the previous 6 months: AI3 = alcohol intake the previous 3 days; ACPM = Alcohol consumption the previous month; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ASI = addiction severity index ; BAC = blood alcohol concentration; BAL = blood alcohol level; CDTect = absolute value of carbohydrate deficient transferrin; CS = cross sectional; CTA = cross tabulation analysis; dt = during treatment; EtG = ethyl glucuronide; EtS = ethyl sulphate; F = fair; FP6M: frequency of alcohol use the previous 6 months; FUAD = follow up after discharge; G = good; GGT = gamma-glutamyl transferase; MCV = mean corpuscular volume; NS = not significant; Long = longitudinal; I = inpatients; LR = logistic regression; N = no; NA = not relevant; NR = not reported; O = outpatients; or = over-reporting; P = poor; PC = Pearson correlation; pt = post treatment (at the end of treatment); Pre-t = pre-treatment (ad admission);RCT = Randomized controlled trial; S = supervised; SP = Spearman correlation; TLFB = time line follow-back; U = unsupervised; ur = under-reporting; Y = yes;

1) Number of patients at the first time point with both subjective and objective measures (except for Midanik, 1982, Babor et al., 2000, and Dahl et al., 2011): see note 3, 5 & 6; 2) The 14.3% is not considered reliable by the authors due to a very low number of participants in a subgroup; 3) Analyses were done on 65 patients;

4) a = Total sample; b = Respondents with positive BAC; c = Depending on how BAC is estimated; 5) analyses were done on test from 40 patients (at both timepoints);

6) Blood analyses were done on 1580 patients at baseline; at follow-up blood analyses were done on 1266 patients in regards of CDT and on 1363 patients in regards of GGT, AST, and ALT.

7) (r;p) Baseline: PDD & GGT: (0.2;***); PDD & AST: (0.15;***); PDD & ALT: (0.1;NS); DDD &: GGT: (015;***); DDD & AST: (0.15;***); DDD & ALT: (0.15;***); At 15 months: PDD & GGT: (.017;***); PDD & AST: (0.20;***); PDD & ALT: (0.14;***); PDD & CDT: (0.32;***) DDD & GGT: (.017;***); DDD & AST: (0.24;***); DDD & ALT: (0.17;***); DDD &CDTect: (0.26;***);

0.01% to 0.03% in blood alcohol reduction per hour, Midanik, 1982, estimated BACs of the patients and compared them with self-reported data. Depending on the elimination rate used, for patients with positive BACs (BAC > 0‰), a tendency towards under-reporting was seen in 14.8%–44.0%, towards consistent reporting in 11.1%–33.3%, and towards over-reporting in 29.6%–74.1%. For the entire group of patients, a tendency towards under-reporting was observed in 6.2%–18.5%, towards consistent reporting in 43.1%–60.0%, and towards over-reporting in 23.1%–56.9% (Midanik, 1982). Only Midanik, 1982, reported correlation coefficients. For the entire sample, correlations between self-report and BAC varied between 0.35 to 0.44 ($p \leq 0.002$), depending on the estimated elimination rate of alcohol. For patients with positive BAC, correlations between self-report and BAC varied between 0.21 to 0.25 and were non-significant (Midanik, 1982). Corroborating these findings, another study (Whitford et al., 2009) found, by means of linear regression, that breath ethanol was not a predictor of self-reported abstinence.

For urine as biological material, when self-report was tested by means of ethanol as a biological marker, ethanol markers indicated drinking in 7.0%–41.0% of the cases where drinking was not reported (Carlsson et al., 1993; Peachey and Kapur, 1986). However, 5.9% of the patients reported alcohol consumption that was not identified by the biological marker (Peachey and Kapur, 1986). By means of 5-HTOL as a biological marker, 5-HTOL markers indicated drinking in 47.0% of the cases where drinking was not reported. The correlation coefficient between self-report and 5-HTOL markers was significant ($p = 0.0001$) but weak ($r = 0.23$) (Carlsson et al., 1993).

Though inconsistency was described, it was also found that when patients reported abstinence, their statement would often be supported by breath ethanol as a biomarker (Sobell et al., 1979b).

Hence, for short-term markers in general, a tendency towards both over- and under-reporting was observed, with under-reporting being the most common type of inconsistency. Results indicating under-reporting were seen in 5.5%–56.0% of the patients and the correlation coefficients were mostly non-significant or weak. Although several of the correlation coefficients were significant, they were mostly weak and not even close to being moderate or strong; the latter would be expected in order to be acceptable when comparing instruments measuring the same construct.

For a graphic illustration of the main results for short-term markers, see Fig. 2.

3.3.2. *Intermediate-term markers*

When self-reported abstinence was validated by means of the long-term, direct markers, EtG and EtS in urine, it was found that the markers indicated drinking in 5.0%–50.0% of the patients (Dahl et al., 2011; Erim et al., 2007). The correlation coefficients between self-report and EtG and between self-report and EtS were 0.66 and 0.72, respectively, and both were significant. Concurrently, this study also found that many cases of self-report were supported by biomarkers, but biomarkers also revealed a large number of occasions when prior drinking had not been reported (Dahl et al., 2011).

For a graphic illustration of the results for intermediate-term markers, see Fig. 3

3.3.3. *Long-term markers*

By means of AST, ALT and GGT, one study (Babor et al., 2000) found that self-report at pre-treatment was consistent in 29.3%–39.7% of the patients, depending on which biomarker was measured. Under-reporting was indicated in 0.1% of the patients, and over-reporting in 60.2%–70.6%. At one-year follow-up (after end of treatment), consistent self-reporting was indicated in 42.6%–55.6% of the patients by means of AST, ALT, GGT and CDT. Under-reporting was indicated in 2.9%–5.3% of the patients, and over-reporting in 39.1%–53.7%. Correlation coefficients were mostly significant but weak, see Table 3.

By combining MCV and GGT, one study (Keso and Salaspuro, 1990)

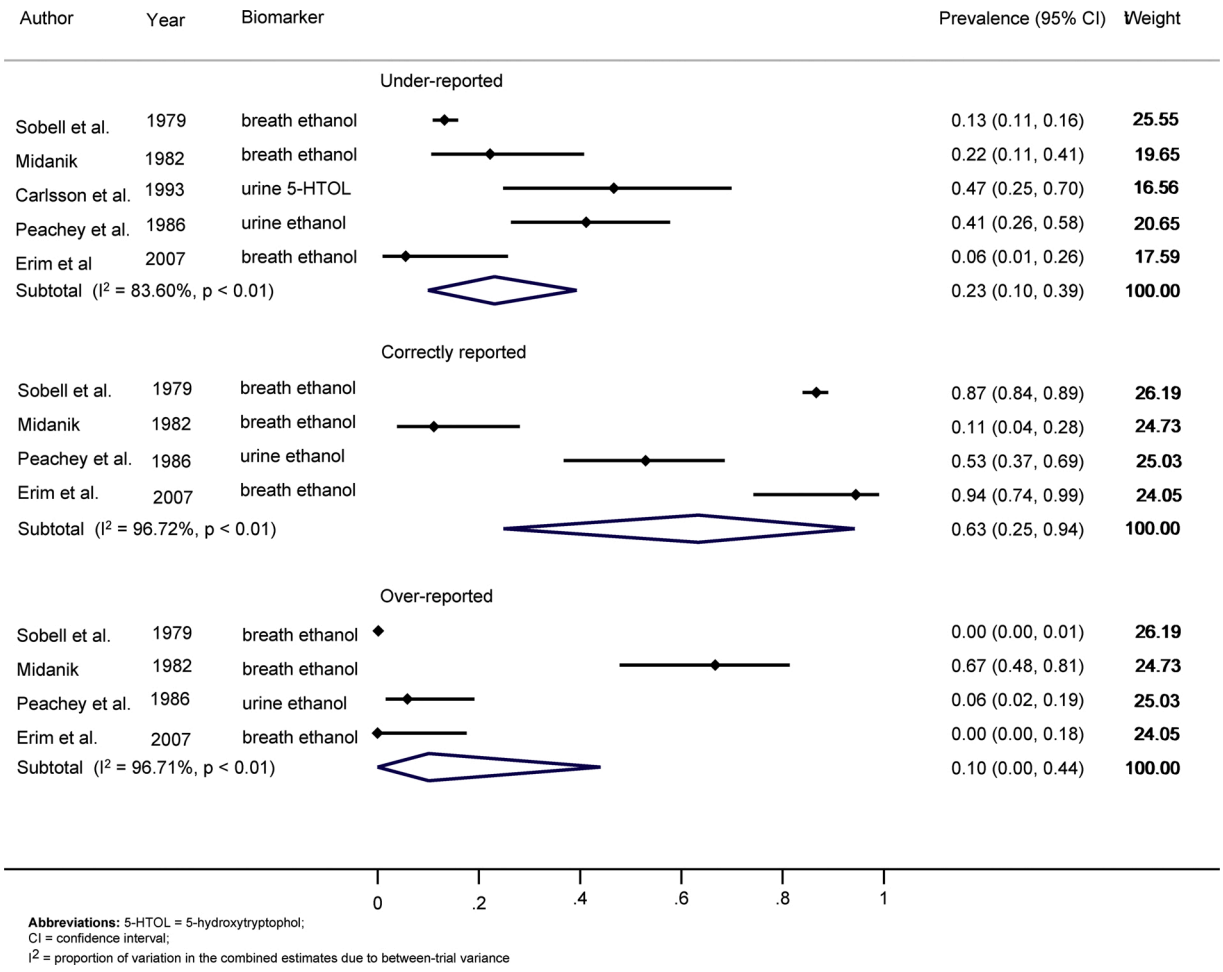

Fig. 2. Graphic illustration of the main results of short-term markers.

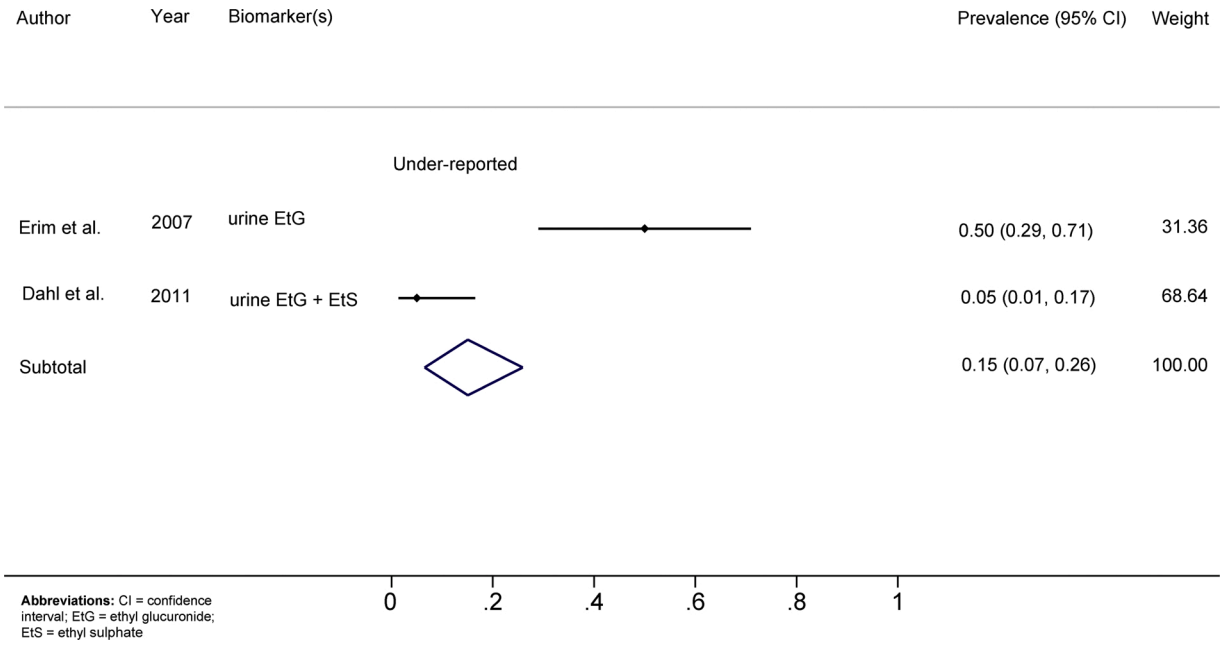

Fig. 3. Graphic illustration of the results of intermediate-term markers.

found that 39.7% of the patients drank controllably, a lower percentage than the one revealed by self-report (72.6%). Altogether, up to 40.0% of the patients may have under-reported their alcohol consumption. By combining GGT and CDT, one study (Mundle et al., 1999) found self-report to be doubted in 15.0% of the cases.

In one study using CDT (Carlsson et al., 1993), it was found that 47.0% of the patients reported never having consumed alcohol. Normal levels of CDT (indicating no drinking) were seen in 60.0% of the patients and over-reporting was thereby indicated in 13.0% of the patients. Twenty per cent of the patients reported frequent drinking (> 10.0% of positive reports). CDT levels corresponding to frequent drinking were seen in 20.0% of the patients as well. Thirty-three per cent of the patients reported sporadic drinking (> 0.0%–10.0% positive reports) and corresponding CDT levels were found in 20.0% of the patients.

When used as an external reference, GGT was significantly higher among drinkers than non-drinkers and correlation coefficients between self-report and GGT were significant, varying between 0.47 and 0.61 (Yoshino and Kato, 1995). One study (Whitford et al., 2009) found that AST, ALT and GGT were not significant indicators of self-reported abstinence.

Though inconsistency was described, it was also reported, that when abstinence was claimed, heavy drinking was ruled out to some extent (Keso and Salaspuro, 1990), self-reported alcohol consumption in general was valid (Babor et al., 2000; Mundle et al., 1999), and that indirect, long-term biomarkers do not add any further information to self-reported information on alcohol consumption (Babor et al., 2000).

Summing up using long-term markers, there were indications of consistent, under-, and over-reporting, and the correlation coefficients were mostly significant but weak.

For a graphic illustration of the main results for long-term markers, see Fig. 4.

3.4. Risk of bias in individual studies

Regarding the quality rating, specifically two risk of bias items varied between the studies: whether a validated questionnaire was used to obtain data on self-reported alcohol consumption and whether patients were asked about their alcohol consumption prior to assessment of the biological marker. For the last-mentioned, however, this item was not relevant for all the studies, e.g. in the studies where the biomarker was collected by the patients themselves on a daily basis and information on alcohol consumption was reported by mail daily or by interview three times a week (Carlsson et al., 1993; Peachey and Kapur, 1986). As can be seen in Table 4, the studies were quality rated as poor (Erim et al., 2007; Sobell et al., 1979b; Whitford et al., 2009), fair (Carlsson et al., 1993; Dahl et al., 2011; Keso and Salaspuro, 1990; Midanik, 1982; Mundle et al., 1999; Peachey and Kapur, 1986; Sobell et al., 1979b; Yoshino and Kato, 1995) or good (Babor et al., 2000).

4. Discussion

4.1. Summary of evidence

This study aimed to provide an overview of the consistency between self-reported alcohol intake and biomarkers among patients in current or previous AUD treatment. A total of thirteen eligible studies were identified, reporting outcomes on both short-, intermediate- and long-term biomarkers. Eight studies were based on short-term markers, two studies were based on intermediate markers and six studies used long-term markers. The results of the synthesis indicate that under-reporting of alcohol consumption was the most common type of discrepancy across the different types of biomarkers. Under-reporting was for short-term markers indicated in 7 studies (n = 15–585) in a range from 5.5%–56.0% of the patients, and over-reporting was indicated in 2 studies (n = 34–65) in a range from 5.9%–74.1%. For

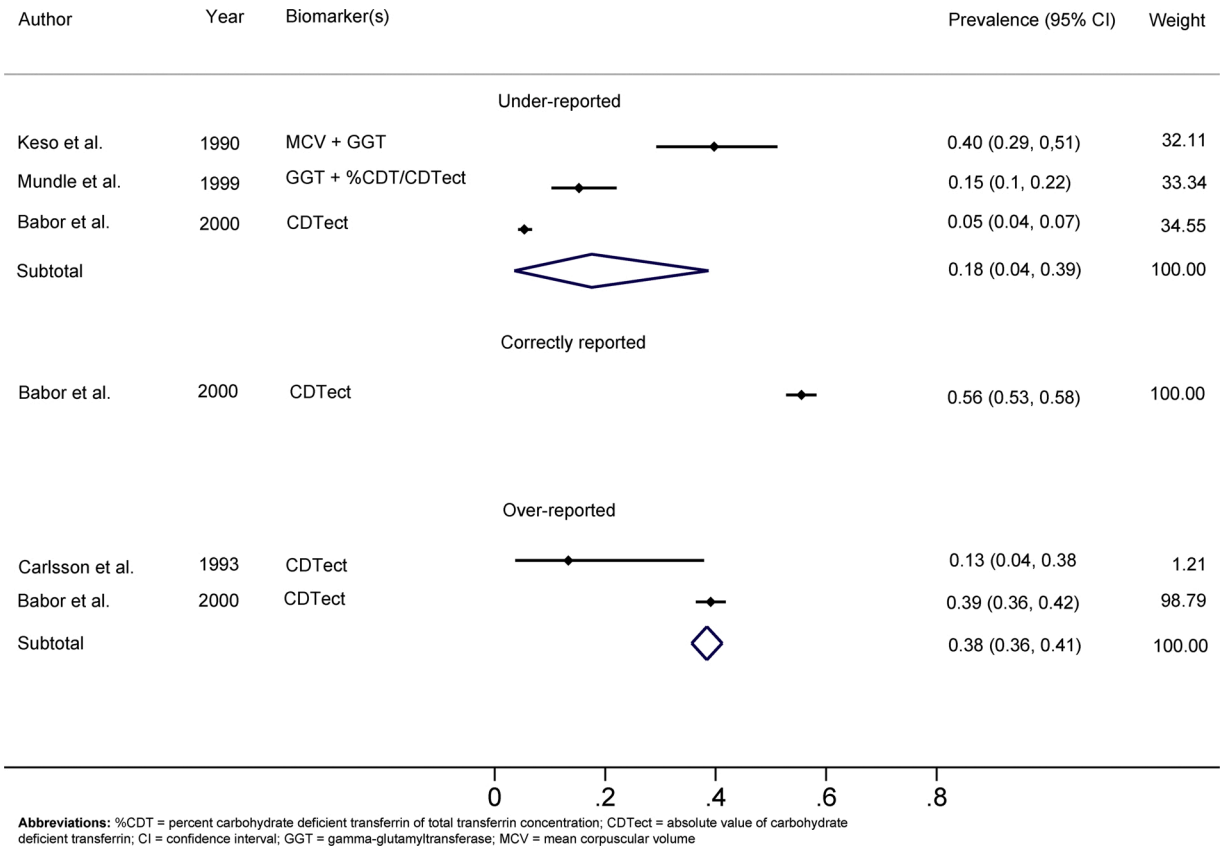

Fig. 4. Graphic illustration of the main results of long-term markers.

Table 4
Quality rating.

Quality item	Sobell et al. 1 1979	Sobell et al. 2 1979	Sobell et al. 3 1979	Midanik, 1982	Peachey et al., 1986	Keso et al., 1990	Carlsson et al., 1993	Yoshino et al., 1995	Babor et al., 2000	Erim et al., 2007	Whitford et al., 2009	Mundle et al., 2009	Dahl et al., 2011
1. Was the research question or objective in this paper clearly stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. Was the study population clearly specified and defined?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
3. Was the participation rate of eligible persons at least 50%?	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
4. Were all the subjects selected or recruited from the same or similar populations? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	CD	CD	CD	Y	Y	CD	CD	Y	Y	Y	CD	CD	Y
5. Was a sample size justification, power description, or variance and effect estimates provided?	N	N	N	N	N	N	N	N	Y	N	N	N	N
6. Were the exposure (s) of interest measured prior to the outcome(s) being measured?	Y	Y	Y	Y	NA	NR	NA	NR	Y	NR	NR	NA	NR
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome?	N	Y	Y	Y	Y	Y	Y	NR	Y	N	Y	Y	Y
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	N*	N*	N*	N*	N*	Y	N	Y	Y	N	Y	N*	Y
10. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Rating	Poor	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Good	Poor	Poor	fair	Fair

Abbreviations: Y = yes; N = no; NR = not reported; CD = cannot determine; NA = not relevant.
Question 9: * indicates that self-report measures were assessed by a standardized, but not validated instrument.

intermediate-term, direct markers, only under-reporting was reported and indicated in 2 studies (n = 18–54) in a range from 5.0%–50.0% of the patients. Although the results for CDT and indirect, long-term biomarkers were not reported consistently across the studies, under-reporting was indicated in 3 studies (n = 73–1580) in a span from 0.1%–40.0% of the patients, and over-reporting in 2 studies (n = 15–1580) in a range from 13.0%–70.6%. Correlations between self-reports of alcohol consumption and biological markers were strongest for the intermediate-term, direct markers, ranging from moderate to strong, whereas for short-term markers, CDT and long-term, indirect markers, the correlations were mostly weak. Most of the studies were quality rated as fair.

4.1.1. Self-report

Though most studies did not apply standardized self-report measures, three studies (Midanik, 1982; Sobell et al., 1979b) assessed drinking during very short time spans (up to 48 h) and another two (Carlsson et al., 1993; Peachey and Kapur, 1986) very frequently (at least three times/week). Thus, the measures of alcohol consumption used in these five studies may provide estimates that are as good as or even better than those provided by retrospective methods assessing intake over a long period. Therefore, although it is preferable to apply well-known methods, we believe that the majority of the self-reports are as trustworthy as the standardized measures.

4.1.2. Self-report and short-term biomarkers

Correlation coefficients between self-reports and breath ethanol were weak (Midanik, 1982). Since the patients were asked about alcohol consumption during the previous 24 h, it may be difficult to cover the entire time span using breath ethanol. However, the correlation coefficient between 5-HTOL/5-HIAA (covering a 24 h time span) and self-report was also weak (Carlsson et al., 1993). Despite this, several more cases suggesting under-reporting were revealed by means of 5-HTOL/5-HIAA and 5-HTOL/CREA than by ethanol (Carlsson et al., 1993). Hence, breath ethanol seems appropriate when it comes to ensuring that patients are sober during e.g. therapy sessions, but because it can only be detected when alcohol is present in the body, it covers a very short time span.

Ethanol in urine reveals information on BAC at the time the urine was produced (Andresen-Streichert et al., 2018; Maenhout et al., 2013) making it cover a longer time span than breath ethanol, and using 5-HTOL increases the detection time even more (Beck et al., 1995; Ghosh et al., 2019; Helander et al., 1996; Johnson et al., 2005). Though not mentioned as a common concern among the authors of the included studies, one article describes the possibility of patients providing "bogus" samples (Peachey and Kapur, 1986). This conduct, though considered to occur rarely, may blur the results. However, it is worth noticing the 5.9% of the patients whose alcohol consumption was not detected by ethanol in urine (Peachey and Kapur, 1986).

4.1.3. Self-report and intermediate-term biomarkers

Several cases indicating under-reporting were also assessed by means of the direct, intermediate-term markers, EtG and EtS (Dahl et al., 2011; Erim et al., 2007). These markers have a window of detection of up to 48 h after single alcohol intake (Heier et al., 2016). The correlation coefficients between self-report and EtG and between self-report and EtS were moderate and strong, respectively (Dahl et al., 2011). However, one should take into consideration that not all the patients provided urine samples, and, though not considered a major issue, there is still a risk of bogus samples. Erim et al., 2007, found indications of under-reporting to a larger extent than Dahl et al., 2011, and had reduced the risk of bogus samples in their collection of samples, but it is also important to note that different groups of patients with AUDs were assessed: patients in treatment for AUD who were on a waiting list for a liver transplant (Erim et al., 2007) and outpatients in AUD treatment under more usual circumstances (Dahl et al., 2011).

Regarding the monitoring of abstinence, EtG in urine is considered appropriate, and sensitivity and specificity have been shown to be 89.0% and 99.0%, respectively (Andresen-Streichert et al., 2018; Heier et al., 2016).

4.1.4. Self-report and long-term biomarkers

Correlation coefficients between self-reports and long-term biomarkers were mostly significant but weak (Babor et al., 2000; Yoshino and Kato, 1995). This could be due to that these markers may be affected by other conditions than alcohol consumption. Further, it takes a certain amount of alcohol consumption over the course of weeks before an elevation in the markers can be seen. Finally, these markers do not immediately return to normal levels after alcohol consumption is discontinued. Reaching normal levels usually take weeks after discontinuation of excessive alcohol consumption. Conditions that can affect indirect biomarkers may be certain types of medication, obesity, non-alcohol induced hepatic disorders, genetic variance, and endocrine disorders (Andresen-Streichert et al., 2018; Conigrave et al., 2003; Kunutsor, 2016; Maenhout et al., 2013). These markers are thus insecure markers for the detection of any alcohol consumption but may reflect heavy drinking. They are not suitable for monitoring abstinence (Andresen-Streichert et al., 2018), though normal levels of these markers seem to rule out heavy drinking to some extent (Andresen--Streichert et al., 2018; Conigrave et al., 2003; Kunutsor, 2016; Maenhout et al., 2013). On the other hand, GGT is not necessarily elevated despite excessive alcohol consumption (Andresen-Streichert et al., 2018).

In general, long-term markers are difficult to use as objective biomarkers of alcohol intake, a point also raised by the authors of some of the included studies (Babor et al., 2000; Whitford et al., 2009; Yoshino and Kato, 1995). Discrepancies between self-reports of alcohol consumption and these markers, representing tendencies towards both under- and over-reporting, have been found (Babor et al., 2000; Keso and Salaspuro, 1990; Mundle et al., 1999; Yoshino and Kato, 1995), and one study (Yoshino and Kato, 1995) reported a moderate correlation coefficient between GGT and self-report. An additional explanation for the moderate correlation coefficient may be that the patients were asked about alcohol consumption during the previous six months (Yoshino and Kato, 1995) but GGT only covers approximately the previous 6 weeks (Andresen-Streichert et al., 2018).

To sum up long-term markers are considered suitable for the detection of excessive chronic alcohol consumption, but they are not appropriate for detecting abstinence. Though considered suitable as biomarkers of excessive consumption, the sensitivity and specificity of these markers may be rather low; however, sensitivity and specificity may be improved when combined with CDT (Andresen-Streichert et al., 2018).

4.1.5. Self-report and biomarkers in general

In general, one should be careful when relying only on self-report or biomarkers as an evaluation of treatment outcome. Biomarkers for the confirmation or rejection of self-report need to be chosen carefully and the results should be interpreted with caution as well. An interesting example of an apparently almost perfect correspondence between self-report and breath ethanol is seen in the study by Erim et al., 2007b, in which discrepancy tending towards under-reporting was observed in only 5.5% of the patients (Erim et al., 2007). However, when the direct, intermediate-term marker, urine EtG, was examined, it indicated that 50.0% of the patients had been drinking at least once without reporting it. Also, it is worth noting that 33.0% of the patients denied providing a

sample at least once. A similar pattern of varying discrepancy depending on the detection window of the markers was reported by Carlsson et al., 1993, when comparing 5-HTOL and CDT with self-report. 5-HTOL has a detection window of up to 24 h (Beck et al., 1995; Ghosh et al., 2019; Helander et al., 1996; Johnson et al., 2005) and CDT does not detect smaller amounts of drinking (Andresen-Streichert et al., 2018; Niemelä and Alatalo, 2010). In this study, abstinence was reported by 47.0% of the patients, though 5-HTOL was found in at least one urine sample of any patient and CDT levels were normal in 60.0% of the patients, indicating that 13.0% of the patients were over-reporting. In this example, it must be taken into consideration not only the window of detection but also the missing ability of CDT to detect any drinking. For the 5.9% of the patients whose consumption was not identified by the biomarker in the study by Peachey et al., 1986, this may be explained by a relatively short detection time of ethanol in urine (Andresen-Streichert et al., 2018).

In addition to the factors specifically relating to the measures, many other factors may have influenced the inconsistency in the included studies. Obviously, the study setting is important when it comes to both over- and under-reporting. Indications of over-reporting were often seen when heavy drinking was a criterion for getting treatment (Midanik, 1982). Although under-reporting was the most common type of inconsistency and often did not have any consequences for treatment, it was identified in 50.0% of the patients in a study applying abstinence as a criterion for getting access to a liver transplant (Erim et al., 2007). In this context, the order in which the measures are conducted as well as the staff in charge of collecting the measures should also be mentioned. The order of collecting self-reported data and the biomarkers is particularly relevant when it comes to breath ethanol where the optimal scenario is that data on self-report is collected prior to the biomarker (Midanik, 1982; Sobell et al., 1979b), preventing patients from getting information on their blood alcohol level and subsequently answering in accordance with the test result. Also, due to social desirability (Welte and Russell, 1993), it may be advantageous that information on self-reported consumption is collected by independent research assistants who are not involved in patient treatment, though treatment staff might also be able to do this depending on the interaction and confidence between treatment staff and patients.

Inconsistency may also be due to elevated BAC causing changed perception of reality (Welch, 2011) and excessive alcohol consumption also causes gastro-intestinal effects leading to vomiting, which may result in the misdetection of alcohol consumption (Vonghia et al., 2008). Lastly, AUD causes poor episodic memory and general cognitive impairment (Jung and Namkoong, 2014; Le Berre et al., 2017), and cognitive impairment could be worsened by a possible hepatic encephalitis often found in patients with severe AUD (Butterworth, 2019; Vilstrup et al., 2014).

In sum, differences between self-report measures and biomarkers with respect to sensitivity and specificity and the limited window of detection of the biomarkers prevents us from deriving an unambiguous conclusion, which is also affected by differences in treatment setting and patient population. However, based on the most specific biomarkers, we suggest that the patients' current needs and conditions are reflected in their self-reports as either under-reporting (e.g. if a new liver is needed) or over-reporting (if it is a prerequisite for getting treatment for AUD). Therefore, the use of validated biomarkers is of utmost importance in clinical and research settings as add-on information to self-report, though there may be circumstances in research settings where the inconsistency may be balanced, e.g. in blinded pharmacological intervention trials.

In this review, no studies used transdermal alcohol sensors, PEth in blood, EtG or EtS in hair, but these may be important future markers. EtG in hair (if possible, in combination with PEth in blood) would make it possible to increase the window of detection by months. Furthermore, regular alcohol consumption would be detected and reduction in drinking for a shorter time period would not cover excessive drinking behavior as it would when using biomarkers in e.g. urine with a detection time of 24–48 hours. However, occasional drinking would not necessarily be detected (Andresen-Streichert et al., 2018; Maenhout et al., 2013). To detect occasional drinking, transdermal alcohol sensors could be considered (Fairbairn and Kang, 2019). Applications to other settings where psychology services are offered may also be possible. For example, transdermal monitoring may be a useful tool for verifying abstinence among individuals undergoing custody or competency evaluations. Monitoring may also be useful for verifying abstinence among individuals taking medications or being evaluated for medical procedures or treatments for which concurrent alcohol consumption may be contraindicated.

4.2. Limitations

This review has some limitations. First, few studies used validated questionnaires to assess self-reported alcohol consumption. Second, several of the studies were small. Third, the studies used different biological measures with various possibilities. Fourth, the time frame of the studies varied. Fifth, except for MCV, indirect, long-term markers may return to normal within a few weeks after alcohol cessation. Used with a self-reported measure of months, levels may be normal at the time the sample is provided. Lastly, it was not possible to retrieve enough raw data to perform a meta-analysis.

5. Conclusion

Self-reported alcohol consumption and biomarkers may be inconsistent in a considerable proportion of patients with AUD, which poses a possible threat to the current evidence-base for AUD treatments. Nonetheless, more studies applying more sensitive, specific, and easily accessible biological measures are warranted to confirm this preliminary synthesis. Future studies should consider using a biomarker as add-on to self-reports, preferably validated questionnaires, during and after alcohol treatment. First and foremost, biomarkers must be direct (e.g. ethanol, EtG, PEth or FAEE) or almost exclusive markers (e.g. 5-HTOL in urine) of alcohol consumption. The time frame for the self-reported measure should correspond with the window of detection of the biomarker and biological material. For detecting occasional drinking, transdermal alcohol sensors could be a possibility.

Contributors

Authors DGN, KA, ASN, CJ and AIM designed the study. Author DGN and AIM conducted literature searches and identified the included studies, extracted the data, and conducted the quality assessment. Authors DGN and AIM wrote the first draft of the manuscript and all authors contributed to and have approved the final manuscript.

Appreciations

The authors wish to thank Flemming Nielsen, PhD & Senior Researcher, University of Southern Denmark for his helpful comments and Lotte Skøt for proofreading the manuscript.

Declaration of Competing Interest

None.

Role of funding sources

The study was unconditionally funded by the Lundbeck Foundation, Denmark (grant ID: R155-2012-11282 and Region of Southern Denmark, Denmark (grant ID: D.NielsenA2716 1.r.18).

References

American Psychiatric Association, 1980. Diagnostic and Statistical Manual of Mental Disorders: DSM-III. American Psychiatric Association, Washington.

American Psychiatric Association, 1987. Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R, 3, revis ed. American Psychiatric Association, Washington.

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV. American Psychiatric Association, Washington.

Andresen-Streichert, H., Muller, A., Glahn, A., Skopp, G., Sterneck, M., 2018. Alcohol biomarkers in clinical and forensic contexts. Arztebl. Int. 115, 309–315.

Babor, T.F., Steinberg, K., Anton, R., Del Boca, F., 2000. Talk is cheap: measuring drinking outcomes in clinical trials. J. Stud. Alcohol 61, 55–63.

Beck, O., Helander, A., Carlsson, S., Borg, S., 1995. Changes in serotonin metabolism during treatment with the aldehyde dehydrogenase inhibitors disulfiram and cyanamide. Pharmacol. Toxicol. 77, 323–326.

Breslin, F.C., Borsoi, D., Cunningham, J.A., Koski-Jännes, A., 2001. Help-seeking timeline followback for problem drinkers: preliminary comparison with agency records of treatment contacts. J. Stud. Alcohol 62, 262–267.

Butterworth, R.F., 2019. Hepatic encephalopathy in cirrhosis: pathology and pathophysiology. Drugs 79, 17–21.

Carey, K.B., Carey, M.P., Maisto, S.A., Henson, J.M., 2004. Temporal stability of the timeline followback interview for alcohol and drug use with psychiatric outpatients. J. Stud. Alcohol 65, 774–781.

Carlsson, A.V., Hiltunen, A.J., Beck, O., Stibler, H., Borg, S., 1993. Detection of relapses in alcohol-dependent patients: comparison of carbohydrate-deficient transferrin in serum, 5-hydroxytryptophol in urine, and self-reports. Alcohol. Clin. Exp. Res. 17, 703–708.

Carney, M.A., Tennen, H., Affleck, G., Del Boca, F.K., Kranzler, H.R., 1998. Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time "electronic interviews. J. Stud. Alcohol 59, 447–454.

Conigrave, K.M., Davies, P., Haber, P., Whitfield, J.B., 2003. Traditional markers of excessive alcohol use. Addiction 98, 31–43.

Connors, G., Maisto, S., 2003. Drinking reports from collateral individuals. Addict., Suppl. 98, 21–29.

Dahl, H., Hammarberg, A., Franck, J., Helander, A., 2011. Urinary ethyl glucuronide and ethyl sulphate testing for recent drinking in alcohol-dependent outpatients treated with acamprosate or placebo. Alcohol Alcohol. 46, 553–557.

De Feo, T.M., Fargion, S., Duca, L., Mattioli, M., Cappellini, M.D., Sampietro, M., Cesana, B.M., Fiorelli, G., 1999. Carbohydrate-deficient transferrin, a sensitive marker of chronic alcohol abuse, is highly influenced by body iron. Hepatology 29, 658–663.

Del Boca, F.K., Babor, T.F., McRee, B., 1994. Reliability enhancement and estimation in multisite clinical trials. J. Stud. Alcohol Suppl. 12, 130–136.

Drummond, C., Pilling, S., Brown, A., Copello, A., Day, E., Dyer, M., Flanagan, E., Fry, J., Georgeson, B., Gilvarry, E., Gosnall, J., Harris, L., Lewis, J., Lingford-Hughes, A., McCarthy, T., Morgan, M., Noble, S., Phillips, T., Roberts, P., Dervan, J., 2011. Alcohol use disorders: the NICE guideline on diagnosis. Assessment and Management of Harmful Drinking and Alcohol Dependence.

Erim, Y., Bottcher, M., Dahmen, U., Beck, O., Broelsch, C.E., Helander, A., 2007. Urinary ethyl glucuronide testing detects alcohol consumption in alcoholic liver disease patients awaiting liver transplantation. Liver Transpl. 13, 757–761.

Fairbairn, C.E., Kang, D., 2019. Temporal dynamics of transdermal alcohol concentration measured via new-generation wrist-worn biosensor. Alcohol. Clin. Exp. Res. 43, 2060–2069.

Fleming, M.F., Anton, R.F., Spies, C.D., 2004. A review of genetic, biological, pharmacological, and clinical factors that affect carbohydrate-deficient transferrin levels. Alcohol. Clin. Exp. Res. 28, 1347–1355.

Ghosh, S., Jain, R., Jhanjee, S., Rao, R., Mishra, A., 2019. Alcohol biomarkers and their relevance in detection of alcohol consumption in clinical settings. Int. Arch. Subst. Abuse Rehabil. 1.

Grant, K.A., Tonigan, J.S., Miller, W.R., 1995. Comparison of three alcohol consumption measures: a concurrent validity study. J. Stud. Alcohol 56, 168–172.

Heier, C., Xie, H., Zimmermann, R., 2016. Nonoxidative ethanol metabolism in humans-from biomarkers to bioactive lipids: nonoxidative Ethanol Metabolism in Humans. IUBMB Life 68, 916–923.

Helander, A., Beck, O., Jones, A.W., 1996. Laboratory testing for recent alcohol consumption: comparison of ethanol, methanol, and 5-hydroxytryptophol. Clin. Chem. 42, 618–624.

Helander, A., Wielders, J., Anton, R., Arndt, T., Bianchi, V., Deenmamode, J., Jeppsson, J.-O., Whitfield, J.B., Weykamp, C., Schellenberg, F., on behalf of the International Federation of Clinical, C., Laboratory Medicine Working Group on Standardisation of Carbohydrate-Deficient, T., International Federation of Clinical, C., Laboratory Medicine Working Group on Standardisation of Carbohydrate-Deficient, T., 2016. Standardisation and use of the alcohol biomarker carbohydrate-deficient transferrin (CDT). Clin. Chim. Acta 459, 19–24.

Johnson, R.D., Lewis, R.J., Canfield, D.V., Dubowski, K.M., Blank, C.L., 2005. Utilizing the urinary 5-HTOL/5-HIAA ratio to determine ethanol origin in civil aviation accident victims. J. Forensic Sci. 50, 670–675.

Jung, Y.-C., Namkoong, K., 2014. Alcohol: intoxication and poisoning - diagnosis and treatment. Handb. Clin. Neurol. 125, 115.

Keso, L., Salaspuro, M., 1990. Comparative value of self-report and blood tests in assessing outcome among alcoholics. Br. J. Addict. 85, 209–215.

Kunutsor, S.K., 2016. Gamma-glutamyltransferase—friend or foe within? Liver Int. 36, 1723–1734.

Le Berre, A.P., Fama, R., Sullivan, E.V., 2017. Executive functions, memory, and social cognitive deficits and recovery in chronic alcoholism: a critical review to inform future research. Alcohol. Clin. Exp. Res. 41, 1432–1443.

Lin, Z., Wang, H., Jones, A.W., Wang, F., Zhang, Y., Rao, Y., 2020. Evaluation and review of ways to differentiate sources of ethanol in postmortem blood. Int. J. Legal Med. 134, 2081–2093.

Litten, R.Z., Allen, J.P., SpringerLink, 1992. Measuring Alcohol Consumption: Psychosocial and Biochemical Methods. Humana Press, Totowa, NJ.

Luginbühl, 2016. Determination of Fatty Acid Ethyl Esters in Dried Blood Spots by LC–MS/MS As Markers for Ethanol Intake: Application in a Drinking Study.

Maenhout, T.M., De Buyzere, M.L., Delanghe, J.R., 2013. Non-oxidative ethanol metabolites as a measure of alcohol intake. Clin. Chim. Acta 415, 322–329.

Maisto, S.A., Conigliaro, J.C., Gordon, A.J., McGinnis, K.A., Justice, A.C., 2008. An experimental study of the agreement of self-administration and telephone administration of the timeline followback interview. J. Stud. Alcohol Drugs 69, 468–471.

Midanik, L., 1982. Over-reports of recent alcohol consumption in a clinical population: a validity study. Drug Alcohol Depend. 9, 101–110.

Miller, W.R., 1996. Form 90: a Structured Assessment Interview for Drinking and Related Behaviors: Test Manual.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, P., The, P.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6, e1000097.

Mundle, G., Ackermann, K., Gunthner, A., Munkes, J., Mann, K., 1999. Treatment outcome in alcoholism - A comparison of self-report and the biological markers carbohydrate-deficient transferrin and gamma-glutamyl transferase. Eur. Addict. Res. 5, 91–96.

National Heart Lung and Blood Institute (NHLBI), 2020. Accessed 1/4/2020. Study Quality Assessment Tools.

Niemelä, O., Alatalo, P., 2010. Biomarkers of alcohol consumption and related liver disease. Scand. J. Clin. Lab. Invest. 70, 305–312.

O'Farrell, T., Langenbucher, J., 1988. Time-line drinking behavior interview. Dictionary of Behavioral Assessment Techniques, pp. 477–479.

Peachey, J.E., Kapur, B.M., 1986. Monitoring drinking behavior with the alcohol dipstick during treatment. Alcohol. Clin. Exp. Res. 10, 663–666.

Reus, V.I., Fochtmann, L.J., Bukstein, O., Eyler, A.E., Hilty, D.M., Horvitz-Lennon, M., Mahoney, J., Pasic, J., Weaver, M., Wills, C.D., McIntyre, J., Kidd, J., Yager, J., Hong, S.-H., 2018. The american psychiatric association practice guideline for the pharmacological treatment of patients with alcohol use disorder. Am. J. Psychiatry 175, 86–90.

Roy, M., Dum, M., Sobell, L.C., Sobell, M.B., Simco, E.R., Manor, H., Palmerio, R., 2008. Comparison of the quick drinking screen and the alcohol timeline followback with outpatient alcohol abusers. Subst. Use Misuse 43, 2116–2123.

Scheurich, A., Müller, M., Anghelescu, I., Lörch, B., Dreher, M., Hautzinger, M., Szegedi, A., 2005. Reliability and validity of the form 90 interview. Eur. Addict. Res. 11, 50–56.

Schober, P., Boer, C., Schwarte, L.A., 2018. Correlation coefficients: appropriate use and interpretation. Anesth. Analg. 126, 1763–1768.

Searles, J.S., Helzer, J.E., Walter, D.E., 2000. Comparison of drinking patterns measured by daily reports and timeline follow back. Psychol. Addict. Behav. 14, 277–286.

Sillanaukee, P., Strid, N., Allen, J.P., Litten, R.Z., 2001. Possible reasons why heavy drinking increases carbohydrate-deficient transferrin. Alcohol. Clin. Exp. Res. 25, 34–40.

Sobell, L.C., Maisto, S.A., Sobell, M.B., Cooper, A.M., 1979a. Reliability of alcohol abusers' self-reports of drinking behavior. Behav. Res. Ther. 17, 157–160.

Sobell, M.B., Sobell, L.C., VanderSpek, R., 1979b. Relationships among clinical judgment, self-report, and breath-analysis measures of intoxication in alcoholics. J. Consult. Clin. Psychol. 47, 204–206.

Sobell, L.C., Brown, J., Leo, G.I., Sobell, M.B., 1996. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. Drug Alcohol Depend. 42, 49–54.

Sobell, L.C., Agrawal, S., Sobell, M.B., Leo, G.I., Young, L.J., Cunningham, J.A., Simco, E. R., 2003. Comparison of a quick drinking screen with the timeline followback for individuals with alcohol problems. J. Stud. Alcohol 64, 858–861.

Toll, B.A., Cooney, N.L., McKee, S.A., O'Malley, S.S., 2006. Correspondence between Interactive Voice Response (IVR) and Timeline Followback (TLFB) reports of drinking behavior. Addict. Behav. 31, 726–731.

Tonigan, J.S., Miller, W.R., Brown, J.M., 1997. The reliability of Form 90: an instrument for assessing alcohol treatment outcome. J. Stud. Alcohol 58, 358–364.

Vilstrup, H., Amodio, P., Bajaj, J., Cordoba, J., Ferenci, P., Mullen, K.D., Weissenborn, K., Wong, P., 2014. Hepatic encephalopathy in chronic liver disease: 2014 practice guideline by the american association for the study of liver diseases and the european association for the study of the liver. Hepatology 60, 715–735.

Vonghia, L., Leggio, L., Ferrulli, A., Bertini, M., Gasbarrini, G., Addolorato, G., Group, A. T.S., 2008. Acute alcohol intoxication. Eur. J. Intern. Med. 19, 561–567.

Welch, K.A., 2011. Neurological complications of alcohol and misuse of drugs. Pract. Neurol. 11, 206–219.

Welte, J.W., Russell, M., 1993. Influence of socially desirable responding in a study of stress and substance abuse. Alcohol. Clin. Exp. Res. 17, 758–761.

Whitford, J.L., Widner, S.C., Mellick, D., Elkins, R.L., 2009. Self-report of drinking compared to objective markers of alcohol consumption. Am. J. Drug Alcohol Abuse 35, 55–58.

World Health Organization, 2000. International Guide for Monitoring Alcohol Consumption and Related Harm. World Health Organization, Geneva.

Wozniak, M.K., Wiergowski, M., Namiesnik, J., Biziuk, M., 2017. Biomarkers of alcohol consumption in body fluids - possibilities and limitations of application in toxicological analysis. Curr. Med. Chem.

Wu, J., Haitjema, C.H., Heger, C.D., Boge, A., 2020. A proof-of-concept analysis of carbohydrate-deficient transferrin by imaged capillary isoelectric focusing and in-capillary immunodetection. Biotechniques 68, 85–90.

Yoshino, A., Kato, M., 1995. Influence of social desirability response set on self-report for assessing the outcome of treated alcoholics. Alcohol. Clin. Exp. Res. 19, 1517–1519.