Alcohol Biomarkers in Applied Settings: Recent Advances and Future Research Opportunities

Raye Z. Litten, Ann M. Bradley, and Howard B. Moss

During the past decade, advances have been made in the identification, development, and application of alcohol biomarkers. This is important because of the unique functions that alcohol biomarkers can serve in various applied settings. To carry out these functions, biomarkers must display several features including validity, reliability, adequacy of temporal window of assessment, reasonable cost, and transportability. During the past two decades, several traditional alcohol biomarkers have been studied in multiple human studies. Meanwhile, several new, promising biomarkers, including various alcohol metabolites and alcohol biosensors, are being explored in human studies. In addition, researchers have explored using biomarkers in combination and using biomarkers in combination with self-reports, resulting in increased sensitivity with little sacrifice in specificity. Despite these advances, more research is needed to validate biomarkers, especially the new ones, in humans. Moreover, recent advances in high-throughput technologies for genomics, proteomics, and metabolomics offer unique opportunities to discover novel biomarkers, while additional research is needed to perfect newly developed alcohol sensors. Development of more accurate biomarkers will help practicing clinicians to more effectively screen and monitor individuals who suffer from alcohol use disorders.

Key Words: Alcohol Biomarkers, Alcohol Sensors, Alcohol Use Disorders.

The alcohol use disorders (AUDs), alcohol abuse, and alcohol dependence are complex, sometimes devastating, disorders responsible for a host of medical, psychological, and social problems (Li, 2007). During the past decade, new behavioral and pharmacological interventions have been introduced that can help to arrest AUDs and related harms before they become severe (Fuller and Hiller-Sturmhofel, 1999; Johnson, 2008; Litten et al., 2005; Miller et al., 2001). In particular, the use of medications in alcohol-dependent patients can help to reduce the likelihood and severity of early relapse (Johnson, 2008; Litten et al., 2005). With these advances in treatment, biomarkers can be instrumental to identify persons with alcohol problems who may be in need of treatment and to monitor a patient’s progress during treatment. Indeed, the potential uses of alcohol biomarkers extend beyond clinical settings to encompass multiple public safety, criminal justice, and research applications.

While considerable progress has been made in the discovery and development of alcohol biomarkers (Hannuksela et al., 2007; Litten and Fertig, 2003; Niemela, 2007), efforts to improve existing biomarkers and identify new biomarkers continue to be an important National Institute on Alcohol Abuse and Alcoholism (NIAAA) research priority, as evidenced by the 2005 Request for Applications Identification of Alcohol Biomarker Signatures (RFA-AA-06-002) and Genomic, Proteomic, and Metabolomic Fingerprints as Alcohol Biomarkers (SBIR/STTR) (RFA-AA-06-001), and the June 2008 conference “Workshop on Alcohol Biomarkers.” This article reviews current and potential uses of alcohol biomarkers in applied and research settings with a focus on chemical compounds in the body that are part of either normal biological processes or a pathogenic process. In addition, we describe desirable features of biomarkers, summarize research knowledge about traditional, new, and promising biomarkers, and update knowledge of their clinical utility. The article is intended to highlight future research opportunities in this rapidly evolving component of the alcohol research field.

POTENTIAL USES OF ALCOHOL BIOMARKERS IN APPLIED SETTINGS

Alcohol biomarkers perform several vital functions in medicine, public safety, and research (Litten and Fertig, 2003). In clinical settings, they serve as objective means of identifying problem drinking and estimating the extent of alcohol-related tissue damage, thereby enabling clinicians to target treatments to AUD severity. In addition, alcohol biomarkers can be used to diagnose and monitor alcohol-related medical conditions.
Increasingly, alcohol biomarkers also serve as objective measures of AUD treatment outcome. Whereas the accuracy of patient self-reports is well supported by research evidence (Del Boca and Darkes, 2003) and, at present, the most commonly used means to ascertain patient relapse, biomarkers could be used either to verify or, if sufficiently accurate, replace patient self-reports altogether. Clinicians also use biomarkers to provide patients with feedback about drinking effects and to foster motivation to cut back or refrain from drinking. In settings that also conduct clinical trials, biomarkers may be used to gauge the effectiveness of experimental interventions. In such applications, the use of valid, objective markers of treatment outcome can contribute to both clinical trial efficiency and confidence in outcome reports (Anton et al., 2002).

In individual and public safety applications, biomarkers are used to monitor abstinence in high-risk individuals (e.g., actively drinking pregnant women, persons previously convicted of alcohol-related offenses) and situations (e.g., medical, transportation, other occupations that affect public well-being) (Litten and Fertig, 2003).

In alcohol research, biomarkers have the potential to serve as trait markers of AUD phenotypes (i.e., observable physical and biochemical characteristics of an individual determined by genetic makeup and environmental factors). The identification and categorization of alcohol dependence phenotypes is expected to provide models for the multiple subtypes of alcohol dependence (Ait-Daoud et al., 2009; Mayfield and Harris, 2009; Oslin et al., 2003), enabling clinicians to identify patients who are likely to respond positively or negatively to specific treatments, especially medications. Thus, biomarkers will have a central role in personalized medicine for AUDs.

**DESIRABLE PSYCHOMETRIC CHARACTERISTICS OF ALCOHOL BIOMARKERS**

To perform these functions successfully, alcohol biomarkers must possess certain attributes including a high degree of validity, i.e., the ability to measure accurately the condition of interest (Litten and Fertig, 2003). Biomarker validity is commonly determined by its criterion validity, that is, the correlation between the biomarker and a criterion variable (or variables) taken as representative of the construct of interest. These criterion variables may include concurrent measures of alcohol consumption, indicators of alcohol dependence, or indicators of organ damage. In addition, certain alcohol biomarkers may be examined for predictive validity by testing the manner in which the current biomarker level reflects some future biological process, such as advanced organ damage or severity of AUD. Last, useful alcohol biomarkers should demonstrate ecological validity, wherein they demonstrate utility in real-world situations. Biomarkers must be easily implemented in and transported to a variety of settings, including primary care, medical specialty, and workplace settings as needed. They should be easily administered, with samples easily collected and analyzed, and affordable, with reasonable costs for instruments, other materials, and labor. Above all, they must be acceptable to the clinicians who use them and the patients to whom tests are administered. Finally, a desirable test should display a practicable window of assessment. It is important to determine how long a biomarker will remain in the body as a function of the amount of alcohol consumed after drinking is stopped. Ideally, a biomarker should persist for at least several days after cessation of drinking and return to normal values only after a period of abstinence that affords sufficient time to screen for the condition.

**Test accuracy** is another important characteristic of a biomarker that reflects the quality and usefulness of the biomarker test. Tests for alcohol biomarkers must accurately measure the marker of interest. For example, if a milliliter of blood contains 30 units of a marker, a valid test would identify proportionally the same or a similar quantity.

The accuracy of biomarker is reflected in its sensitivity (i.e., the probability that the biomarker will produce a true positive result when used on an affected population) and specificity (i.e., the probability that a test will produce a true negative result when used in a nonaffected population) (see Table 1). Reliable cutoff values (i.e., threshold values that define the presence of a condition) are essential in determining sensitivity and specificity and depend on the condition being measured and the setting in which it is measured. For example, a high cutoff value will make it more difficult to detect the condition (decreasing the sensitivity) but will reduce the number of false positives.

**Table 1. Diagnostic Test Characteristics and Definitions**

<table>
<thead>
<tr>
<th>Condition or disease</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>TP</td>
<td>FP</td>
</tr>
<tr>
<td>-</td>
<td>FN</td>
<td>TN</td>
</tr>
<tr>
<td>All with condition</td>
<td>TP+FN</td>
<td>FP+TN</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>TP/(TP+FN)</td>
<td>FN/(FP+TN)</td>
</tr>
<tr>
<td>Specificity</td>
<td>TN/(FP+TN)</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>(TP + TN)/(TP+FN+FP+TN)</td>
<td>Prevalence = (TP + FN)/(TP+FN+FP+TN)</td>
</tr>
</tbody>
</table>

| All positive tests   | TP+FP| All negative tests | TN+FN |
| Positive predictive  | TP/(TP+FP) | Negative predictive | TN/(TN+FN) |
positives (increasing the specificity). In contrast, a low cutoff value results in a higher sensitivity and a lower specificity. In settings such as primary care screening for problematic drinking where high sensitivity is desirable and false positives are acceptable, a lower cutoff value is used. By contrast, in criminal justice settings, a higher cutoff value may be prudent because false positives can have devastating effects. The determination of cutoff values should be validated by rigorous research studies.

Positive and negative predictive value of a given biomarker test is another measure of test accuracy. When using alcohol biomarkers, professionals must take into account the positive predictive value (PPV) (i.e., the percentage of positive tests in which the condition occurs) and the negative predictive value (NPV) (i.e., the percentage of negative tests in which the condition does not occur) for the biomarker used. Predictive values depend not only on the sensitivity and specificity of a test, but also on the prevalence of the disorder in the population tested. A population with low prevalence of the disorder results in an increase in false positive tests, whereas a higher prevalence rate yields more false-negative tests. Table 2 presents examples of this concept: assuming that a biomarker has a sensitivity and specificity of 90%, if the prevalence of heavy drinkers is 10% (as in a primary care screening situation), the PPV is only 50%, while the NPV value is 99%. On the other hand, if the prevalence of heavy drinkers is 90% (as in an addiction clinic), then the PPV is 99%, while the NPV is only 50%. Thus, the interpretation of tests will vary from one setting to another depending on the prevalence of the problem being screened. This concept is not only a concern for screening alcohol problems, but for all biomarkers used to screen any medical disorder.

Useful alcohol biomarkers must also demonstrate a high degree of reliability. Reliability of a biomarker is only of interest when the biomarker has first demonstrated its validity. Biomarker reliability refers to the degree to which the biomarker is internally consistent and stable in measuring that which it is intended to measure over time. There are several general types of test reliability in the psychometrics literature, but only a few are applicable to biomarker analysis. Perhaps, the most important of these for biomarkers is test–retest reliability. Here, the biomarker is evaluated in the same sample of individuals on two different occasions. This approach assumes that there is no substantial change in the underlying condition of interest (e.g., level of alcohol consumption, magnitude of organ damage) between the two occasions. Because these conditions may vary over time, the amount of time allowed between measures is critical. Thus, the shorter the time gap, the higher the expected test–retest reliability; the longer the time gap, the lower the expected reliability.

## UPDATE ON ALCOHOL BIOMARKERS

During the past decade, many alcohol biomarkers have been studied in humans. Among clinical applications, alcohol biomarkers have been used primarily to detect alcohol consumption and less frequently to detect tissue damage. This section reviews traditional, new, and promising alcohol biomarkers that have been used in various human studies (see Table 3).

### Traditional Alcohol Biomarkers

The most direct way to ascertain alcohol consumption is to measure alcohol’s presence in blood, urine, or breath. Although such measures may be useful under certain conditions (e.g., to test for recent consumption in the workplace), the relatively short half-life of alcohol prohibits their utility in most settings (Swift, 2003). To address this problem,
<table>
<thead>
<tr>
<th>Alcohol biomarker/device</th>
<th>Drinking behavior targeted</th>
<th>Sample source</th>
<th>Window of assessment&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Primary indication</th>
<th>Sensitivity</th>
<th>False positives</th>
<th>Population</th>
<th>Cost/type of testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGT</td>
<td>Chronic heavy drinking</td>
<td>Blood</td>
<td>2 to 3 weeks</td>
<td>Screening</td>
<td>Moderate (low to high)</td>
<td>Many sources of false positives</td>
<td>Adults (ages 30 to 60)</td>
<td>Low/routine testing</td>
</tr>
<tr>
<td>AST and ALT</td>
<td>Chronic heavy drinking</td>
<td>Blood</td>
<td>2 to 3 weeks</td>
<td>Screening</td>
<td>Lower than GGT</td>
<td>Many sources of false positives</td>
<td>Adults (ages 30 to 60)</td>
<td>Low/routine testing</td>
</tr>
<tr>
<td>MCV</td>
<td>Chronic heavy drinking</td>
<td>Blood</td>
<td>Up to several months</td>
<td>Screening</td>
<td>Lower than GGT</td>
<td>Several conditions of false positives</td>
<td>Adults (ages 30 to 60)</td>
<td>Low/routine testing</td>
</tr>
<tr>
<td>% CDT</td>
<td>Moderate to high (heavy drinking for 7 to 10 days)</td>
<td>Blood</td>
<td>2 to 3 weeks</td>
<td>Screening/relapse</td>
<td>Similar to GGT</td>
<td>Few conditions of false positives</td>
<td>Adults (influenced by gender)</td>
<td>Moderate to high/specialized testing</td>
</tr>
<tr>
<td>5-HTOL/5-HIAA</td>
<td>~ 4 drinks</td>
<td>Urine</td>
<td>1 day</td>
<td>Screening/relapse</td>
<td>High</td>
<td>Few conditions of false positives</td>
<td>More research on patient variability</td>
<td>High/specialized testing</td>
</tr>
<tr>
<td>EtG</td>
<td>1 to 2 drinks</td>
<td>Urine</td>
<td>Several days</td>
<td>Abstinence/relapse</td>
<td>Few sources of false negatives Unknown</td>
<td>Few sources of false positives</td>
<td>More research on individual variability</td>
<td>Moderate/specialized testing</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>Hair</td>
<td>Up to several months</td>
<td>Screening/Abstinence</td>
<td>Unknown</td>
<td>Unknown</td>
<td>More research on individual variability</td>
<td>High/specialized testing</td>
</tr>
<tr>
<td>EIS</td>
<td>1 to 2 drinks</td>
<td>Urine</td>
<td>1 to 2 days</td>
<td>Abstinence/Relapse</td>
<td>High</td>
<td>Few sources of false positives</td>
<td>More research on individual variability</td>
<td>Moderate/specialized testing</td>
</tr>
<tr>
<td>FAEE</td>
<td>Unknown</td>
<td>Hair</td>
<td>Up to several months</td>
<td>Screening/Abstinence</td>
<td>Unknown</td>
<td>Unknown</td>
<td>More research on individual variability</td>
<td>High/specialized testing</td>
</tr>
<tr>
<td></td>
<td>At least several drinks</td>
<td>Blood</td>
<td>2 days</td>
<td>Abstinence/relapse</td>
<td>Unknown</td>
<td>Unknown</td>
<td>More research on individual variability</td>
<td>High/specialized testing</td>
</tr>
<tr>
<td>PEth</td>
<td>Heavy drinking for ~5 days</td>
<td>Blood</td>
<td>1 to 2 weeks</td>
<td>Screening/relapse</td>
<td>High for screening/unknown for relapse</td>
<td>Unknown</td>
<td>More research on individual variability</td>
<td>Moderate/specialized testing</td>
</tr>
<tr>
<td>SCRAM</td>
<td>1 to 5 drinks</td>
<td>Transdermal alcohol vapor</td>
<td>Continuously records</td>
<td>Abstinence/relapse</td>
<td>Moderate</td>
<td>False positives approach zero</td>
<td>More research on individual variability</td>
<td>High/high-tech device worn around ankle</td>
</tr>
</tbody>
</table>

GGT, gamma-glutamyl transpeptidase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; MCV, mean corpuscular volume; EtS, ethyl sulfate; CDT, carbohydrate-deficient transferrin; EIG, ethyl glucuronide; FAEE, fatty acid ethyl esters; PEth, phosphatidyl ethanol; HIAA, hydroxyindoleacetic; HTOL, hydroxytryptophol; SCRAM, secure continuous remote alcohol monitor.

<sup>a</sup>Window of assessment: time biomarker remains measurable in body after drinking is stopped.

Primary indications: Abstinence – detecting any alcohol intake; Relapse – monitoring drinking, especially heavy drinking during treatment; Screening – detecting drinking, especially high-risk heavy drinking.

Sources: Substance Abuse and Mental Health Services Administration, 2006; Kissack et al., 2008; Conigrave et al., 1995, 2003; Litten et al., 1995; Niemela, 2007; Javors et al., 1997; Fleming et al., 2004; Jeppsson et al., 2007.
Researchers have identified biomarkers of alcohol consumption with longer windows of assessment than direct alcohol measures of breath and body fluids. These so-called traditional biomarkers measure alcohol consumption indirectly by detecting tissue damage or other physiological reactions to heavy drinking over time.

The most common traditional marker is gamma glutamyl transferase (GGT). GGT is found in the cell membranes of multiple tissues, including the liver, kidney, pancreas, spleen, heart, and prostate (Javors et al., 1997). Chronic heavy drinking elevates GGT levels in the blood because of increased GGT synthesis and/or GGT leakage from liver cells that have been damaged or destroyed by alcohol. GGT has a long window of assessment, with values that remain elevated for 2 to 3 weeks after heavy-drinking cessation (Javors et al., 1997). GGT usually is not elevated in heavy-drinking adolescents and young adults. Although the reason for this is still unknown, it may be related to the need for individuals to reach a certain physiological age before noticeable GGT changes occur or to drink heavily for many years to elevate GGT (Conigrave et al., 2003). Overall, GGT sensitivity for screening heavy drinking is moderate, ranging from low to high values depending on the population and setting where it is used (Conigrave et al., 1995, 2003; Litten et al., 1995). Although GGT is measured routinely at reasonable cost in most clinical laboratories, its clinical utility is limited by a high rate of false positives because of non-alcohol-related liver diseases, obesity, diabetes, smoking, and such medications as anticonvulsants, anticoagulants, and barbiturates (Conigrave et al., 2003).

Aspartate aminotransferase (AST) and alanine aminotransferase (ALT), liver enzymes often measured in routine screening for liver damage, also are used as biomarkers of heavy alcohol consumption. AST and ALT have characteristics similar to those of GGT but with lower sensitivity (Conigrave et al., 2003). Consequently, AST and ALT are used less frequently than GGT to screen for heavy alcohol consumption.

Another traditional alcohol biomarker is mean corpuscular volume (MCV), a measurement of the size of red blood cells. Chronic heavy drinking increases the size of red blood cells. As with the liver enzymes, MCV is routinely measured in most laboratories and works best to detect heavy drinking in adults aged 30 to 60. Overall, for reasons that are still unclear, MCV appears to have lower sensitivity in males and higher sensitivity in females than GGT and is more specific than GGT (Conigrave et al., 2003). Because red blood cells exhibit a long half-life (e.g., 13 to 27 weeks), MCV values can remain elevated up to several months after cessation of drinking (Niemela, 2007). Numerous sources of false positives include folate and B₁₂ deficiencies, nonalcoholic liver diseases, hemolysis, bleeding disorders, hypothyroidism, medications that produce marrow toxicity, and bone marrow disorders (Conigrave et al., 2003; Niemela, 2007).

Most recent among traditional biomarkers is carbohydrate-deficient transferrin (CDT) (Anton, 2001; Niemela, 2007). Transferrin, a glycoprotein synthesized and secreted by the liver, transports iron throughout the body. Heavy drinking (50 to 80 g alcohol/day) for 7 to 10 days decreases the carbohydrate content of transferrin, including sialic acid, galactose, and N-acetylglucosamine, by a mechanism that remains unclear but most likely involves decreased activity of glycosyltransferases and increased activity of sialidases, enzymes that add and remove carbohydrate groups, respectively (Litten and Allen, 1998; Stüber, 1991). After cessation of drinking, 2 to 3 weeks are required for serum CDT values to return to normal. CDT has been tested as a screening marker of alcohol intake and recently has been demonstrated to detect relapse (Anton et al., 2002). Whereas CDT sensitivity is similar to that of GGT (Litten et al., 1995), its real strength is its high specificity. The few identified sources of false positives include rare genetic transferrin variants, primary biliary cirrhosis, chronic end-stage liver disease, and hepatocarcinoma (Fleming et al., 2004; Javors and Johnson, 2003). In addition, CDT is influenced by smoking, body weight, and female gender (Fleming et al., 2004). During the past several years, methods to measure CDT have improved, resulting in a further increase in sensitivity and specificity (Bergstrom and Helander, 2008). These improvements have also minimized the effects of smoking, body weight, and gender on CDT values. As a result of these improvements, the International Federation of Clinical Chemistry and Laboratory Medicine Working Group made the following recommendations (Jeppsson et al., 2007): (i) to normalize the variability of the transferrin content, especially in women, CDT should be expressed as relative amount (CDT/total transferrin); (ii) CDT is a complex compound consisting of several glycoforms, among which the disialotransferrin glycoform displays the best correlation with alcohol intake (Bergstrom and Helander, 2008; Jeppsson et al., 2007). Therefore, the disialotransferrin glycoform is recommended as the primary target for CDT measurement. This glycoform can be measured with high performance liquid chromatography (HPLC), which has been suggested to be the reference method for % CDT measurement (Jeppsson et al., 2007). Finally, % CDT has been approved by the FDA in the United States as a marker of heavy alcohol consumption.

New Biomarkers

A serotonin metabolite that is elevated during alcohol consumption, 5-hydroxytryptophol (5-HTOL) is elevated at approximately 50 + g of alcohol during a drinking occasion (Beck and Helander, 2003). 5-HTOL displays high sensitivity and specificity and appears uninfluenced by age, gender, liver diseases, or medications other than disulfram (Beck and Helander, 2003). 5-HTOL is expressed as a ratio to 5-hydroxyindole-3-acetic acid (5-HIAA), another metabolite of serotonin, an attribute that reduces false positives, especially from foods rich in serotonin such as bananas, pineapples, and tomatoes (Borg et al., 1992). This ratio
measurement is important because increases in serotonin from these foods elevate the serotonin metabolites 5-HTOL and 5-HIAA equally. Limitations of 5-HTOL include 1-day window of assessment (Beck and Helander, 2003; Hoiseth et al., 2008). In addition, cumbersome methodologies, such as gas chromatography–mass spectroscopy (GS–MS), liquid chromatograph-mass spectroscopy (LCMS), or HPLC, must be employed to measure the metabolites (Beck and Helander, 2003). Recently, the glucuronidated form of 5-HTOL, glucuronidated 5-hydroxytryptophol (GTOL), has been used as a biomarker (Beck et al., 2007). GTOL is the predominant metabolite of 5-HTOL in the urine and may be measured using simpler methods such as the enzyme-linked immunosorbent assay (ELISA) (Dierkes et al., 2007).

Recent research has focused on the more direct measurement of alcohol metabolites, an important advance in alcohol biomarkers. The most advanced such biomarker is ethyl glucuronide (EtG), formed from a conjugation reaction of alcohol with glucuronic acid in the presence of the enzyme uridine diphosphate glucuronyl transferase (UGT) (Wurst et al., 2003). This metabolite is measurable in blood, hair, and urine, most commonly urine. There are 16 functional human genes of UGT, raising the possibility that the formation rate of EtG may vary by genotype (Miners et al., 2002; Wurst et al., 2003).

At present, EtG usually is measured by LC–MS, although less costly methods are in development (Bottcher et al., 2008). Detection time for EtG in the urine is longer than in the blood (14 to 24 hours) (Borucki et al., 2007; Hoiseth et al., 2007), so that urine measures have potential to both screen for problematic drinking and monitor relapse in treatment and recovery programs (Kissack et al., 2008; Wurst et al., 2004). The detection time of EtG in the urine varies ranging from 20 hours up to 80 to 102 hours (Kissack et al., 2008; Wojcik and Hawthorne, 2007; Wurst et al., 2004), depending on the amount of alcohol consumed and individual variability. Methodology also is in development to measure EtG in the hair. So far, EtG can be detected in the hair for several months after drinking is stopped. This has implications for forensic purposes as well as for monitoring abstinence among individuals convicted of driving while intoxicated and women who are at risk for drinking during pregnancy (Bendroth et al., 2008; Pragst and Yegles, 2008; Wurst et al., 2008a,b).

EtG appears to be highly sensitive and specific to alcohol intake. In fact, it is so sensitive that it can measure incidental alcohol exposure including alcohol in foods (e.g., cooking wines, flavoring extracts), over-the-counter cold medications, mouthwash, hand sanitizer gel, and other hygiene products (Costantino et al., 2006; Rosano and Lin, 2008). In addition, some urine samples may contain yeast that can convert urine glucose to alcohol and subsequently EtG if stored at room temperature for more than 12 hours (Kissack et al., 2008; Saady et al., 1993); this may lead to problems when EtG is used with to test persons who are diabetic and have high levels of glucose in the urine. False negatives also can arise from E. coli hydrolysis of EtG in urinary tract infections (Helander and Dahl, 2005) or from ingestion of chloral hydrate medications (Arndt et al., 2009). Another source of false negatives is urine dilution. To counter this, it is recommended that either urinary creatinine be measured with a cutoff of 25 mg/dl to indicate dilution (Goll et al., 2002) or EtG be expressed in ratio to creatinine (Dahl et al., 2002).

Finally, EtG varies among individuals (Sarkola et al., 2003). Factors that may underlie this variability include gender, age, ethnic group, medical or psychiatric comorbidity, genetic polymorphism of UGT, and possible others. Characterizing this variation will be important in establishing reliable cutoff values to determine sensitivity and specificity in different populations and settings.

**Promising Biomarkers**

In addition to EtG, researchers are exploring several other promising alcohol metabolites, including phosphatidylethanol (PEth), ethyl sulfate (EtS), fatty acid ethyl esters (FAEEs), and acetaldehyde adducts. PEth, which is formed by the action of phospholipase D, is a group of phospholipids with a common nonpolar phosphoethanol head group (Helander and Zheng, 2009). PEth is detected in the blood after consuming approximately 1000 g of alcohol, usually over about a 2-week period (Varga et al., 1998). PEth has a 1- to 2-week window of assessment, depending on the level of drinking (Stewart et al., 2009; Varga et al., 2000). In addition, its sensitivity and specificity appear better than for traditional biomarkers (Aradottir et al., 2006; Hartmann et al., 2006). Because PEth appears not to be influenced by liver diseases (Stewart et al., 2009), it may be promising in detecting and monitoring heavy drinkers with hepatic pathology because most traditional biomarkers are elevated by non-alcohol-induced liver disease. PEth can be measured using high-performance LC and evaporative light scattering and electrospray MS technologies (Gunnarsson et al., 1998). Nonetheless, new methods are being developed to measure PEth. PEth has recently been measured by combining LC–ESI-MS/MS which is more sensitive than previous method of combining HPLC with evaporative light scattering detection (ELSD), allowing a >2 weeks of window assessment (Gnann et al., 2009). Helander and Zheng (2009) recently identified PEth species in the blood by using an electrospray ionization (ESI) LC–MS approach. More research is required to clinically characterize PEth, in particular to determine its variability among individuals (Varga et al., 2000). EtS is another minor alcohol metabolite formed from the sulfate conjugation of alcohol from 3'-phosphoadenosine 5'-phosphosulfate alcohol, a reaction catalyzed by sulfotransferase (Helander and Beck, 2005). EtS is detectable in urine around 30 hours after the last drink, a short window of assessment. However, because EtS and EtG appear to be highly correlated and LC–MS can simultaneously measure each (Helander and Beck, 2005; Wurst et al., 2006), measuring both may serve to corroborate results. Junghanns and colleagues (2009) recently
reported that all patients who were positive for urinary EtG were also positive for urinary EtS.

FAEES, such as ethyl palmitate, ethyl oleate, and ethyl stearate, are nonoxidative metabolites of alcohol measurable by GC–MS (Pragst et al., 2001). In very heavy drinkers, FAEES can be detected up to 99 hours in the blood (Borucki et al., 2007). FAEES also are present in meconium fluid, enabling detection of alcohol use during pregnancy (Bearer et al., 2003; Moore et al., 2003; Swift, 2003). A limitation of this practice is the inability to accurately determine when and how much alcohol consumption occurred. Recent studies (Pragst and Yegles, 2008; Pragst et al., 2001; Wurst et al., 2008a) also have measured FAEES in the hair. The advantage of this procedure is that it can detect alcohol in the hair up to 2 months after abstinence (Pragst et al., 2001), useful in monitoring drinking in pregnant women (Pragst and Yegles, 2008; Wurst et al., 2008a) and offenders convicted of driving under the influence (DUI) (Wurst et al., 2008b). EtG in hair often is measured simultaneously with FAEES because GC–MS is employed for the measurement of both (Pragst and Yegles, 2008). Further research is needed to determine the relationship between drinking and alcohol metabolite levels in the hair, rate of change in levels in the hair following abstinence, distribution of alcohol metabolites along hair shaft, and sources of false positives and false negatives, such as use of hair lotions, and racial or ethnic variations in FAEE deposition.

Acetaldehyde adducts are formed by the reaction of acetaldehyde, the major metabolite of alcohol, with proteins. Acetaldehyde forms both stable and unstable adducts with proteins. The stable ones remain in the blood for weeks, rendering acetaldehyde adducts a promising marker of alcohol consumption (Litten and Allen, 1998). Sensitivity and specificity appear to be as least as accurate as with traditional biomarkers (Niemela, 2007; Swift, 2003). Other than drinking, additional sources of acetaldehyde, especially from smoking (Salaspuro, 2007), must be identified and characterized. Although, to this time, researchers have been unable to develop a routine method to accurately measure these adducts, new approaches are still being tested, including the use of the ELISA assay to measure specific immunoglobulin A (IgAs) against acetaldehyde–protein adducts (Hietala et al., 2006a).

Finally, other biomarkers have been explored, including sialic acids, sialic acid index of plasma apolipoprotein J, beta-hexosaminidase, platelet MAO B activity, high density lipoprotein (HDL), and dolichol, but these appear no better than the traditional biomarkers (Hannuksela et al., 2007). The measurement of platelet MAO B protein levels rather than activity is also being investigated and early results appear promising (Tabakoff et al., 2009).

**Biomarker Combinations**

Because all biomarkers exhibit some limitations, one approach to improve accuracy is to use them in combination. The most common combination studied, so far, is CDT measured in conjunction with GGT (CDT + GGT) (Litten et al., 1995). Employing the rule that subjects are labeled positive if either test exceeds its cutoff values, CDT + GGT consistently and often impressively increases sensitivity, with little loss of specificity, compared with using either biomarker alone (Litten et al., 1995). This finding was consistent in screening for problematic drinking in multiple populations, including alcoholics with and without liver disease, heavy drinkers, different ethnic groups, males and females, and college students (Litten et al., 1995). Some investigators have used a mathematically formulated equation of CDT + GGT and reported higher sensitivities for the combined markers (Anttila et al., 2003; Hietala et al., 2006b; Sillanaukee and Olsson, 2001). Lastly, Anton and colleagues (2002) found that CDT + GGT performed better than either biomarker alone in detecting relapse to heavy drinking.

Recent studies have used the traditional markers CDT and GGT and the newer markers EtG, EtS, PEth, and FAEES in combination with the self-report Alcohol Use Disorders Identification Test (AUDIT) questionnaire. Using alcohol metabolite biomarkers with CDT or GGT, in combination with the AUDIT, improved detection of drinking in emergency rooms and workplace settings and with pregnant women (Hermansson et al., 2000; Kip et al., 2008; Neumann et al., 2008; Wurst et al., 2008a). Finally, several investigators have developed a program that incorporates the combination of many clinical tests (Harasymiw and Bean, 2001; Korzec et al., 2005). For example, the Bayesian Alcoholism Test (BAT) consists of 8 to 15 clinical and biochemical indicators including GGT, AST, and CDT. Korzec and colleagues (2009) recently reported that BAT was more sensitive than CDT, GGT, and AST alone in distinguishing harmful alcohol users from controls. Another example is the Early Detection of Alcohol Consumption (EDAC), a program that uses a linear discriminant function analysis on 10 to 36 routine laboratory tests (Harasymiw and Bean, 2001, 2007). Harasymiw and Bean (2007) found that EDAC test performed better than GGT in detecting heavy drinking in subjects from treatment and community settings.

**Alcohol Sensor Devices**

From biochemical markers of alcohol consumption (except for direct measurement of blood alcohol levels), it is impossible to determine precisely when drinking occurred, the exact amount of alcohol consumed, how many drinking episodes were required to produce the value rendered, or whether that value has increased or decreased from the last drink. Nor do biochemical compounds correlate accurately with amount of alcohol intake. For these reasons, alcohol sensor devices provide important alternative. With alcohol sensor devices, one could determine the time that drinking occurred; such devices also have the potential to measure accurately the amount of alcohol intake during each drinking episode and provide an estimate of blood alcohol concentration.
Several new noninvasive alcohol devices are now in development, the most advanced of which is the Secure Continuous Remote Alcohol Monitor (SCRAM). The SCRAM is a bracelet worn around the ankle that measures alcohol vapor electrochemically in a semi-quantitative manner (Sakai et al., 2006). A unique property of the SCRAM is that it measures alcohol intake over a 24-hour period, enabling one to determine the approximate time that drinking occurred. So far, its specificity appears to be excellent and its sensitivity moderate for the detection of drinking (Wojcik, 2008). However, a 2007 National Highway Traffic Safety Administration report identified several problems. For example, water accumulates in the device, resulting in reduced sensitivity and lower signals in females than males. In addition, the SCRAM should be more precise in measuring alcohol intake, less expensive, and more comfortable to wear. Further research is needed to determine interindividual variability and to test SCRAM utility in various settings.

Another device in development is the Giner WrisTAS. In contrast to the SCRAM, this device is worn around the wrist. Like the SCRAM, the Giner WrisTAS measures alcohol vapor electrochemically and measures drinking continuous with time (Swift, 2003). It is still at the research stage and is not ready for commercial use (National Highway Traffic Safety Administration, 2007).

**Biomarkers in Specific Settings**

Although there is, at present, no ideal biomarker, biomarkers, nevertheless, are in use in multiple applied settings. This section reviews some of these applications and reports current knowledge of and strategies for using alcohol biomarkers (Table 4). In all such applications, it is critical that biomarkers be validated by adequately controlled research studies.

**Recovery Programs**

Many U.S. recovery programs are expressly geared to treat health care professionals, while others monitor airline and maritime pilots, air traffic controllers, and railway engineers. For example, physician health programs currently monitor abstinence in over 9,000 physicians who have been identified as having alcohol problems (Skipper et al., 2004). A recent survey of physician health programs in 46 states found that many use traditional and new alcohol biomarkers to monitor abstinence (Jansen et al., 2004). Determining whether program participants have relapsed to drinking is a special concern for such programs, given the public service performed by these professionals. In such settings, biochemical tests that monitor abstinence must be sensitive to alcohol consumption while keeping the number of false positives at a minimum. This is important because positive tests can lead to severe consequences. Whereas, at present, many of these programs use EtG, its sensitivity to incidental alcohol exposure remains problematic. Accordingly, a 2006 Substance Abuse and Mental Health Services Administration (2006) included a warning about using EtG test in recovery programs. More research is needed to accurately determine reliable cutoff values that account for individual variability and incidental alcohol exposure.

Possible future research for this setting includes exploring the combination of biomarkers with alcohol sensor devices, such as the SCRAM. EtG and EtS can detect one to two drinks, whereas the SCRAM appears sensitive to several drinks with few false positives. In addition, the SCRAM can record the approximate time that drinking occurs.

**Criminal Justice Settings**

Forty percent of violent crimes committed in the United States involve the consumption of alcohol (U.S. Department of Justice, 1998). Moreover, 4 in 10 fatal motor vehicle accidents involve alcohol, while approximately 1.5 million DUIs occur each year (U.S. Department of Justice, 1998; Wojcik, 2008). Total alcohol abstinence is a common pretrial, probation, or parole order for many of these offenders. Because alcohol is rapidly eliminated from the body, monitoring required abstinence is a challenge to multiple criminal justice settings.

| Table 4. Use of Biomarkers in Applied Settings |
|-------------------------------|-----------------|----------------|----------------|
| Setting                        | Primary indication | Type of alcohol biomarker needed | Possible alcohol biomarkers |
| Recovery program               | Abstinence       | Sensitive to low amounts of alcohol; no false positives | EtG, EtS, SCRAM |
| Criminal justice               | Abstinence       | Sensitive to low amounts of alcohol; no false positives | EtG, EtS, SCRAM |
| Primary care                   | Screening/relapse | Sensitive to heavy drinking | CDT+GGT |
| Alcohol/drug specialty treatment | Relapse         | Sensitive to drinking–especially heavy drinking | CDT+GGT EtG, EtS, SCRAM |
| Other specialty treatment      | Screening       | Sensitive to any drinking–especially to heavy drinking | CDT+GGT EtG, EtS |
| Workplace: employee assistance programs | Abstinence | Sensitive to low levels of alcohol | EtG, EtS SCRAM |
| Workplace: health and safety screening | Screening   | Sensitive to drinking–especially heavy drinking | EtG, EtS CDT, GGT |

GGT, gamma-glutamyl transpeptidase; EtG, ethyl glucuronide; EtS, ethyl sulfate; CDT, carbohydrate-deficient transferring; SCRAM, secure continuous remote alcohol monitor.
During recent years, more than 1600 courts in 45 states have monitored more than 65,000 offenders using the SCRAM (Wojcik, 2008). Although the SCRAM is not sufficiently sensitive to measure total abstinence, early results suggest that it is sensitive in detecting five to six drinks (Wojcik, 2008). In addition, Wurst and colleagues (2008b) suggest that measuring EtG and FAEE in hair may be a reasonable approach to monitor drinking in this population. Although determining alcohol metabolites in hair samples will not reveal amount or duration of drinking, it does reveal whether any drinking occurred over the past few months. Similar to the recovery programs, alcohol biomarkers in criminal justice applications should be sensitive to alcohol intake with the number of false positives approaching zero.

**Primary Care Settings**

Although only 10 to 13% of patients with AUDs seek alcohol/drug specialty treatment (McLellan, 2007; Willenbring, 2009), it is likely that many more are seen by primary care physicians for health problems related to drinking. Most primary care clinicians, however, do not screen for alcohol problems (Dawson et al., 2005; Willenbring, 2009). However, primary care physicians often conduct a battery of clinical tests, including tests of the liver enzymes GGT, AST, and ALT, to evaluate liver function. Elevation of these enzymes could alert physicians to possible drinking problems. As more effective medications receive FDA approval, physicians are more likely to screen and more confidently offer treatment to this population. To that purpose, traditional biomarkers, such as GGT and CDT alone and in combination, may be useful tools to screen for patients who drink heavily over a prolonged period, although these biomarkers without self-reports may not be very sensitive in screening for alcoholism (Conigrave et al., 1995). Encouragingly, in recent studies, Dilie and colleagues (2005) and Kapoor and colleagues (2009) found that adding CDT screening to patient self-report in primary care settings results in significant savings in medical and legal costs. Alcohol biomarkers used in such settings should be sensitive to heavy at-risk drinking, defined for men as 5 or more and for women as 4 or more drinks per occasion (National Institute on Alcohol Abuse and Alcoholism, 2007). For example, Anton and Youngblood (2006) found that CDT could detect heavy drinking that occurred 4 to 5 days a week. The addition of a biomarker not only would confirm the self-report but also would provide results from an objective biochemical test to help physicians to motivate patients to either stop drinking or cut back to low-risk levels (Miller et al., 2004).

**Alcohol/Drug Specialty Treatment Settings**

Although most alcohol/drug specialty facilities do not, at present, use biomarkers for screening, some existing biomarkers could be used for that purpose. For example, researchers (Anton et al., 2002) reported that the combination of CDT and GGT was successful in monitoring relapse to heavy drinking in inpatient alcoholics. Biomarkers, such as EtG and EtS, also have the potential to monitor relapse to any drinking while patients are in treatment. Wurst and colleagues (2008c) reported that EtG measurement in hair and urine of methadone maintenance patients detected alcohol use in those who scored negatively on the AUDIT. As in other settings, the combined use of ethanol metabolites and traditional biomarkers with patient self-reports might be most efficient in monitoring drinking in this population.

**Other Specialty Treatment Settings**

Because alcohol affects every system in the body, many problematic drinkers present themselves in other specialty treatment settings with physical symptoms including liver, cardiovascular, gastroenterological, and sleep disorders (Li, 2007). Reports (Kip et al., 2008; Neumann et al., 2008) suggest that the alcohol metabolites EtG and PEth are useful in emergency medicine to detect drinking in injured and clinically nonintoxicated patients, as well as in those with thoracic or gastrointestinal complaints. In contrast, Neumann and colleagues (2009) found that adding the traditional biomarkers % CDT, MCV, and GGT to the AUDIT was not useful for screening trauma patients with alcohol problems. Fleming and colleagues (2009) reported that measuring blood alcohol levels in patients admitted for trauma was helpful in identifying subjects who denied excessive drinking as well as predicting the development of alcohol withdrawal and other adverse health events. The usefulness of CDT in these patients was, however, unclear and required larger samples. In another medical setting, Wurst and colleagues (2008a) demonstrated that, in a second-trimester ultrasound screening at a hospital setting, pregnant women were most likely to be detected for drinking if a combination of alcohol metabolites (EtG, FAEE, EtS) and the AUDIT were employed. Recently, Pichini and colleagues (2009) found, for the first time, EtG and EtS in the meconium. This raises the possibility of measuring the combination of EtG and EtS along with FAEES, which has already been measured in the meconium to monitor drinking in pregnant women who are at a high risk for drinking. As in primary care settings, it is important that biomarkers used in other specialty settings identify patients engaged in high-risk drinking. A prime example is patients with severe liver disease who may be awaiting transplantation or other treatments where it is crucial to differentiate recent alcohol use from underlying liver pathology.

**The Workplace**

Monitoring abstinence in workers who are problematic drinkers is essential in high-risk work settings, especially in settings that affect public safety such as the transportation sector. For this and less critical workplace purposes,
combining biomarkers with self-reports can be a practical monitoring method. For example, Hermansson and colleagues (2000, 2002) found that the CDT and the AUDIT combined are complementary in identifying risky drinking for employees in the transportation sector. Ideally, in most workplace settings, biomarkers should be sensitive to low levels of alcohol consumption and demonstrate high specificity.

**RESEARCH DIRECTIONS**

Although progress has been made in alcohol biomarker discovery and development, a need remains for relatively inexpensive tests that more accurately measure alcohol intake. All biomarkers developed for clinical use must be fully characterized, with strengths and limitations clearly delineated, so that they can be optimally used in applied settings. This includes identifying factors that can affect a biomarker response to alcohol, including severity of problematic drinking, pattern and amount of drinking, genetic differences, race, gender, age, body weight, and physical diseases (Fleming et al., 2004). In addition, reliable cut off values must be established that reflect optimal sensitivity and specificity as it relates to different populations and settings. Because, at present, no biomarker is ideal, it is essential to explore and characterize combinations of biochemical markers, alcohol devices, and self-reports in research studies.

Aided by recent advances in high-throughput technologies for genomics, proteomics, and metabolomics, opportunities abound to discover novel biomarkers (Hodgkinson et al., 2008; Mayfield and Harris, 2009). To identify individuals who are sensitive to alcohol intake, new bioinformatic approaches and computational tools should be developed that will integrate thousands of RNAs, proteins, and metabolites. This is necessary to be able to identify those that are sensitive to alcohol intake and alcohol-induced tissue damage. For the development of biomarkers to detect alcohol-induced tissue damage, it is essential to define the degree of tissue damage being assessed and also its specificity to alcohol. Finally, it is also important to develop appropriate in vitro and animal models to detect alcohol biomarker signatures of acute and chronic alcohol consumption and alcohol-induced tissue damage.

Research also should continue to explore new technologies, including noninvasive imaging techniques and alcohol biosensors, the latter with a focus on developing more comfortable devices that also are less expensive and more quantitative in measuring alcohol intake.

**FINAL COMMENTS**

Although more research is needed for validation, alcohol biomarkers already serve several vital functions in the prevention, screening, and treatment of AUDs, primarily in recovery and criminal justice settings where monitoring abstinence is the primary goal. In various treatment settings, including primary care settings and trauma services, biomarkers are either not being used or being used only sparingly. This can be expected to change as more effective treatments, especially medications, emerge. However, before any biomarkers are used in applied settings, it is essential that they are characterized rigorously in research studies to better gauge their strengths and weaknesses.

Overall, the future of biomarkers is bright. New high-throughput technologies increase the possibility of discovering biomarker panels or signatures with the potential to be more sensitive and specific. Biomarker signatures, composed of multiple parameters, are sought to monitor either alcohol consumption or alcohol-induced organ damage. Also, advances in technology have been made in developing new alcohol sensors that have the potential not only to measure quantitatively the amount of drinking but also to determine when drinking occurred. Development of more accurate biomarkers will allow clinicians to better identify and monitor individuals who suffer from problematic drinking.

**ACKNOWLEDGMENTS**

The authors appreciate the review of the manuscript by Drs. Kathy Jung and Daniel Falk and the preparation of the tables by Megan Ryan.

**REFERENCES**


improve verification of abstinence in alcohol-dependent patients during in-patient treatment and at follow-ups. Addiction 104:921–926.


