Alcohol and Alcoholism Vol. 48, No. 6, pp. 712–719, 2013 doi: 10.1093/alcalc/agt134
Advance Access Publication 17 August 2013

REVIEW

Use of Novel Technology-Based Techniques to Improve Alcohol-Related Outcomes in Clinical Trials

Eugenia M. Gurvich[1], George A. Kenna[2] and Lorenzo Leggio[3,4,5,*]

[1]Biomedical Engineering Undergraduate Program, Brown University School of Engineering, Providence, RI, USA, [2]Center for Alcohol and Addiction Studies, Brown University Medical School, Providence, RI, USA, [3]Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology, Laboratory of Clinical and Translational Studies, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, USA, [4]Intramural Research Program, National Institute on Drug Abuse, National Institutes of Health, Baltimore, MD, USA and [5]Department of Behavioral and Social Sciences, Brown University, Providence, RI, USA

*Corresponding author: Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology, NIAAA & NIDA, NIH, 10 Center Drive (10CRC/15330) MSC 1108, Room 1-5429, Bethesda, MD 20892-1108, USA. Tel.: +1-301-435-9398; Fax: +1-301-402-0445; E-mail: lorenzo.leggio@nih.gov

(*Received* 4 *March* 2013; *first review notified* 4 *April* 2013; *in revised form* 4 *July* 2013; *accepted* 15 *July* 2013)

Abstract — With a better understanding of the biologic basis of alcohol dependence and the considerable financial burden of alcohol abuse and dependence, the number of alcohol-related clinical pharmacotherapy trials has been on the rise. Subsequently, the potential to find efficacious treatments is more promising. Unfortunately, alcohol-related trials face a number of challenges, as a result of the difficulties that arise from traditional and outdated methods to collect data and ensure medication adherence. Novel technology-based assessments, such as ecological momentary assessment, interactive voice response, transdermal sensor and medication-event monitoring system provide a prospective solution—albeit not without possible concerns—to the difficulties faced in alcohol-related clinical trials. Clinical trials are meant to define the efficacy of the treatment and to determine an effective and safe dosage. However, due to lack of adherence a drug could inappropriately or mistakenly be judged as ineffective for treating a specific disorder. The described technologies may be important tools to prevent false negatives in validating drug efficacy, to provide consistency in clinical trials and to improve available data regarding the study of pharmacotherapies for alcohol dependence.

INTRODUCTION

In recent years, pharmacotherapy for alcohol dependence has garnered a great amount of attention, as the biologic basis of the disease has been demonstrated (Dackis and O'Brien, 2005). Alcohol dependence is now recognized as a complex multifactorial disorder whose etiology includes biologic and environmental factors. By examining the physiologic basis of the disease, it has become apparent that a pharmacologic approach can be effective in treating alcohol dependence. For example, it has been shown that alcohol misuse is driven by increased levels of dopamine in the nucleus accumbens, which lead to feelings of pleasure and euphoria (Koob and Moal, 2001). There is also clear evidence that a variety of other neurotransmitters—e.g. gamma-amino butyric acid (GABA), glutamate, endorphins, serotonin—play a key role in the neurobiology of alcohol dependence; these systems provide a number of potential targets for pharmacotherapies (Lingford-Hughes and Nutt, 2003). The overall concept is that pharmacologic agents may be capable of modifying the functions of neurotransmitter systems and hence modify dependent drinking behavior (Edwards and Koob, 2010). The increased interest in using pharmacotherapies to treat alcohol dependence is further driven by the recent development of pharmaceutical agents that reduce alcohol consumption in animal models of alcohol dependence, which may be effectively applied to clinical trials (Leggio *et al.*, 2010). After several decades when only one medication (i.e. disulfiram) was approved, clinical trials have become more numerous for new pharmacotherapies to treat alcohol dependence (Kranzler, 2000). Indeed, since 1994 three additional medications have been approved by the Food and Drug Administration (naltrexone tablets, naltrexone intramuscular and acamprosate) and several other pharmacotherapies are currently under investigation (Edwards *et al.*, 2011).

There are several important reasons why the development of novel effective medications for alcoholism is important. For example, most alcoholic patients never seek treatment using more time-intensive psychosocial interventions (e.g. alcoholics anonymous, cognitive behavioral therapy), and efficacious medications that can be prescribed in primary care settings could significantly improve the use of alcoholism interventions. Additionally, efficacious medications can be combined with psychosocial therapy and this combined approach can play a crucial role in improving addiction treatment, as it can address both the biologic and psychosocial aspects of addiction (Edwards *et al.,* 2011). Unfortunately, pharmacotherapy trials face difficulties, with low medication adherence, difficulty monitoring alcohol use and inconsistent collection of outcome measures (Weiss, 2004; Swift *et al.*, 2011). Lack of adherence to medications among all areas of medicine in the USA alone has been projected to cost $100 billion annually (Osterberg and Blaschke, 2005). Low patient adherence can distort the validity of clinical trials and make it difficult to analyze drug effectiveness (Haynes, 1987).

Current methods to assess medication adherence include pill counting, assaying medications or their metabolites in body fluids (e.g. blood, urine) or tagging medications with inert tracers. Measuring drug concentrations can be inaccurate due to either the cumulative effects of long-lasting chemical compounds in the drug or the idiosyncrasies in drug metabolism. Measurement does not allow adherence in the placebo arm to be assessed and does not permit real-time adherence problems to be incorporated into behavioral interventions designed to improve medication adherence. While tagging both drug and placebo with a biologic inert tracer (typically riboflavin) can be a powerful indicator of adherence, the amount of riboflavin added must be high enough to be distinguishable from dietary riboflavin, and patients may have variations in their riboflavin excretion (Babiker *et al.*, 1989). The incorporation of riboflavin onto the drug requires additional costs (Young *et al.*, 1984). Riboflavin can determine that the drug was consumed, but not that the participant complied with

accurate consumption (Besch, 1995). Pill counting, while inexpensive, does not validate whether the medication was actually taken yet is still the main means for determining adherence in clinical trials (Spilker, 1991). Pill dumping, whether intentional or unintentional, can occur in an attempt to avoid investigator disapproval or to maintain a high enough level of adherence to continue the study.

A systematic review of alcoholism clinical trials analyzed 22 randomized, double-blind, placebo-controlled clinical trials with naltrexone that reported adherence rates (Swift *et al.*, 2011). Interestingly, of these 22 trials, only 3 (14%) met criteria for high levels of adherence assurance, 5 (23%) met medium adherence assurance criteria and 14 (64%) met low adherence criteria. The Spearman's correlation between risk ratios for return to heavy drinking (for naltrexone vs. placebo) and the level of adherence assurance (low vs. medium vs. high) was significant ($r = -0.62$, $P = 0.025$) (Swift *et al.*, 2011). An additional issue that researchers face in clinical trials is that random or systematic error in self-reported data collection can result in biased data (Del Boca and Noll, 2000); as such, patient reported outcomes have come under scrutiny (Wilke *et al.*, 2004). Typical primary outcomes in alcoholism trials focus on abstinence and/or reduction. An extensively used method is the timeline follow-back (TLFB) method, based upon the respondent retrospectively estimating daily alcohol consumption (Sobell and Sobell, 1992). TLFB is not usually used to differentiate variations in alcohol consumption behavior within the day (e.g. mornings vs. afternoons), and is usually used only to report information on daily drinking. Errors made in self-report data collection may be systematic in that the patients' mood can affect the information retrieved (Tennen and Affleck, 1996), and some errors may be the consequence of deliberate and denial-based distortions (Trivedi *et al.*, 2011). Unlike for other drugs, urine/breath tests are not feasible substitutes for alcohol TLFB (Preston *et al.*, 1997). Additionally, data assessing mood, anxiety and craving can be collected, but can carry the bias of being self-reported retrospectively.

In summary, there is a need to find additional approaches, such as novel technology-based techniques in order to address the issues described above. New technologies that may improve medication adherence and data collection will be discussed in this review. Studies were searched using PubMed and Google Scholar until January 2013, with the following search terms: alcoholism pharmacotherapy, compliance, adherence, clinical trials, addiction, medication-event monitoring system (MEMS), transdermal sensor (TDS), ecologic momentary assessment (EMA) and interactive voice response (IVR) (Table 1). Additional publications were found based on the references of the papers originally reviewed.

Table 1. Number of studies reviewed

Technology	Number of studies
IVR	17
TDS	5
MEMS Cap	8
EMA	11
Total studies	41

IVR, interactive voice response; TDS, transdermal sensor; MEMS, medication event monitoring system; EMA, ecological momentary assessment.

INTERACTIVE VOICE RESPONSE

IVR is a self-report technology used to monitor both medication adherence and alcohol use. Participants receive previously recorded messages and report information about their drinking and/or medication adherence through their phone's voice response technology or touch tone keypad.

The IVR system administers survey questions that are predetermined and pre-recorded (Alemi *et al.*, 1994; Kranzler *et al.*, 2004). IVR exists as either a 'call-in' system, where participants call the system to initiate the interview, or a 'call-out' system, where participants receive a computerized call (Midanik and Greenfield, 2010). Furthermore, 'call-in' and 'call-out' functions may be combined. IVR questionnaires record information regarding the number, type, size and time of drinks, and can also provide other information, e.g. mood, craving, confidence to resist drinking, adherence and drug side effects (Kranzler *et al.*, 2004).

Recent advances allow respondents to verbally input answers, with speech recognition accuracy reaching as high as 90–96% (Abu-Hasaballah *et al.*, 2007). Examples of alcohol data collected via IVR include a trial of naltrexone, where IVR provided information on whether the medication was taken, patient's daily drinking and mood (Kranzler *et al.*, 2004). Further studies have used IVR to measure drinking and medication adherence (Simpson *et al.*, 2009; Kranzler *et al.*, 2012). While there are no studies directly comparing pill counting and IVR, it is reasonable to conclude that IVR is superior to pill counting because it reminds patients to take the pills on a daily basis, therefore improving compliance compared with a simple retrospective compliance measure such as pill counting.

Cranford *et al.* (2010) used IVR to monitor daily drinking and mood in alcoholic individuals and demonstrated that daily adherence with IVR exceeded 90%. Perrine *et al.* (1995) also utilized IVR with male social drinkers; participants' partners called in and reported on the participant's daily alcohol consumption. While this format invites a new bias (participant's partner), it also solves some issues, i.e. evidence that respondents are alcohol free when completing the IVR (Sobell and Sobell, 1992).

Studies comparing IVR to TLFB show that IVR allows participants to sustain a higher level of awareness of their drinking over extended lengths of time (Searles *et al.*, 2002). Thus, drinking was reported in higher levels by IVR than in the TLFB data. In other words, participants consistently underestimated their drinking with the TLFB. Additional studies have confirmed these differences between IVR and TLFB (Searles *et al.*, 2000; Tucker *et al.*, 2002). On the other hand, Toll *et al.* (2006) found moderate-to-high correlations between drinking data assessed by IVR and TLFB, and lower adherence to the IVR than the TLFB in non-alcoholic smokers. A new version of IVR is the interactive web response platform (IWR). With a user-friendly interface, the IWR allows subjects to enter their responses using a computer clicking their answers on an online survey (Jordan *et al.*, 2011).

Strengths and weaknesses of IVR (or IWR) are discussed on the basis of several criteria, e.g. flexibility, ease of use, impact of use, possibility of error and immediacy. Further criteria evaluated are patient honesty, patient-to-researcher relationship, retention of participants and possible sources of bias. IVR allows patients to provide data regardless of location or time constraints. Such flexibility may help in retention of

participants. Additionally, the data acquired by employing IVR are instantaneously accessible to the researchers. Touch-tone replies are immediately entered into the computer system, hence eliminating potential sources of data entry error (Bardone *et al.*, 2000).

IVR expands the population that can be reached; for example, Alemi *et al.* (1994) reported that historically difficult-to-reach populations (e.g. drug users and alcoholic individuals) are more receptive to the IVR than to mailed written questionnaires. IVR data are also more accurate as they reduce recall time and bias (Bardone *et al.*, 2000). Another strength in utilizing IVR to collect data is that participants feel more confident and comfortable in answering questions honestly because their perception of privacy is increased (Tourangeau *et al.*, 2002). As a consequence, participants may provide more valid data than when responding to a live interviewer, presumably because they are less concerned about negative appraisals by a computer. However, it should be mentioned that others have argued that study participants tend to respond in a socially desirable way when face to face with an interviewer (de Leeuw and van der Zouwen, 1998; de Leeuw, 2005). Hence respondents may be more likely to respond in a similar manner over the phone, as opposed to a paper questionnaire. For example, Dillman *et al.* (2009) compared the data-collection methods of telephone interview, mail, IVR and internet and concluded that visual and aural responses differ as a means of data collection. Furthermore, the study tested whether response rates could be improved by offering a second survey mode after multiple attempts to collect data by the first mode; it was observed that significant enhancements to data acquisition were accomplished by switching from IVR to telephone interviews (Dillman *et al.*, 2009).

While IVR has several strengths, as noted above, some weaknesses should also be taken into account. For example, the technology and software necessary to effectively utilize IVR can be expensive. However, it is now possible to replace proprietary IVR platforms with industry-standard computer systems, thus driving down costs (Abu-Hasaballah *et al.*, 2007).

Other weaknesses include participants who may be without phone coverage or might not own a cell phone at all. However, lack of phone service was not a top reason for lack of response in adolescents with alcohol-use disorders (Kaminer *et al.*, 2006). Similarly, IVR can run into issues if participants do not have computers or Internet access.

IVR also has the potential to institute bias because the participants' answers can be influenced by the tone of the interviewer (Corkrey and Parkinson, 2002); this can be avoided by a computer-generated recording. Another weakness is that participants' answers may differ if they are under the influence of alcohol. In addition, IVR might increase attrition of study patients because financial incentives are delayed, and therefore gratification is not instantaneous. Since alcoholic individuals may seek such instantaneous payment, they may be less likely to remain in the study (Bardone *et al.*, 2000). However, attrition in IVR-based trials has been reported only in one smoking study (Toll *et al.*, 2008); therefore, based on the currently available literature, attrition does not seem to be an issue for using IVR-based technology in clinical trials.

Possible improvements of the IVR methodology might be obtained providing participants with phone/internet coverage during the duration of the study.

TRANSDERMAL SENSOR

Ingested alcohol can be measured transdermally with accuracy because transdermal alcohol concentrations are highly correlated with blood alcohol concentrations (BACs). Approximately 1% of alcohol is eliminated transdermally, which is significant enough to allow for accurate measurement (Swift, 2003).

The TDS is an electrochemical detector and wearable device that measures alcohol consumption over long periods of time by measuring ethanol vapor at the skin surface. A person wears the sensor and the device can hence collect objective information continuously. The device is relatively complex, consisting of an electrochemical cell that produces an uninterrupted current signal and a thermistor that monitors contact of the skin and sensor (Swift *et al.*, 1992).

The TDS is portable, and samples data every 30 s to 10 min that is stored for later retrieval. Examples of TDSs include the Wrist Transdermal Alcohol Sensor (WrisTAS™) and the Secure Continuous Remote Alcohol Monitor (SCRAM®).

The WrisTAS™ device (Giner, Inc.) is affixed to the wrist with a Velcro strap and uses proton exchange membrane technology, in which an electrode oxidizes ethanol into acetic acid and the level of oxidation current is measured. This watch-like device provides nearly continuous data. A serial port interface is used to download the data to a computer, but the device itself can store data for 21 days. WrisTAS also measures body temperature and skin resistance (Marques and McKnight, 2009). The SCRAM® device (Alcohol Monitoring Systems, Inc.) uses a fuel cell sensor with a battery lasting between 30 and 45 days (Marques and McKnight, 2009). The bracelet itself is locked around the ankle, but the device also consists of a modem to allow for data uploading and the SCRAMNetwork, which is a remote server that aggregates data. The device is programmed to sample more frequently if alcohol is detected. The data, consisting of alcohol content, skin surface reflectivity and proximity, and body temperature, become available to researchers only minutes after being uploaded.

Strengths and weaknesses of TDS are discussed based on several criteria, i.e. sources of bias, the type of information collected, patient comfort and cost. The TDS may provide objective measures of alcohol consumption because indirectly it tracks BAC (Swift *et al.*, 1992). The correlation between measurements recorded by the sensor and alcohol quantities consumed has been found to be high, but the delay can range from 30 to 120 min. This correlation in amplitude and shape is strong enough to make TDSs a viable option for accurate acquisition of data. TDS data are comparable to measurements determined by breathalyzers and blood samples (Hawthorne and Wojcik, 2011). TDSs are non-invasive, continuous and objective, as the patient has no control over the data collection. Sakai *et al.* (2006) indicated that TDSs show discriminative legitimacy as a semi-quantitative measure of alcohol ingestion and are neither comfortable nor uncomfortable, as rated by the patients. This can be an improvement over the discomfort felt by patients who need to have their blood drawn to test BAC. However, while no bruising was observed in a study by Barnett *et al.* (2011), over half of the participants reported some sort of mark on their skin from the TDS; still, interference was reported as minimal. As for the weaknesses of TDS, for example various brands of TDSs can have peculiarities that reduce performance (Marques and McKnight, 2009). The faults, which can be the result of the highly integrated

technologies involved, caused missed readings and occasional false positives (Marques and McKnight, 2007). The complexity of TDSs and the need for their improvement increase costs. The steep price might deter researchers from utilizing the device, especially for large-scale and long-term treatment trials. Rather, it is more likely that TDS devices may be much more useful for smaller etiologic studies investigating predictors and products of alcohol use in real time, as well as for small proof-of-concept clinical studies testing novel medications for alcoholism.

While studies have demonstrated that the comfort level of the device is neutral, there is a stigma associated with wearing the TDS, as some patients might find it embarrassing (Marques and McKnight, 2007). Barnett *et al*. (2011) reported that out of 20 subjects, the use of the bracelet caused 2 patients to drop out and a third patient to cut off the sensor altogether. A possible improvement of this weakness could be simply obtained improving the cosmetic appearance of the TDS devices, thus increasing participants' acceptance. Finally, TDSs provide no information on medication adherence, as they only measure alcohol intake. The implication is that clinical trials for alcoholism pharmacotherapies would need a second technology to measure medication adherence.

MEDICATION EVENT MONITORING SYSTEM

The aim of the MEMS is to retrofit a standard pill bottle with a microprocessor in the cap that is capable of documenting the frequency, duration and time of bottle openings (Olivieri *et al.*, 1991). MEMS units are special caps that fasten to typical pill vials and so the MEMS bottle does not look any different from conventional bottles. MEMS caps may also be provided with digital displays of the number of hours since the last opening and the total daily openings. The data collected are objective and downloaded to a computer using special software (Malta *et al.*, 2008). The data can be displayed in various forms, such as calendar plots, lists of bottle openings and tables of dose intervals. Dosing patterns can readily be accessed from the data recorded by the microprocessor (Cramer and Rosenheck, 1999). Thus, it is possible to determine and retrieve data for the presumptive dates and times that all doses were or were not taken. In theory, this correlates with the days on which the medication was or was not consumed.

Strengths and weaknesses of MEMS are discussed on the basis of several criteria, such as specificity of data collected, sources of error, patient adherence, misinformation, type of data collected and potential for malfunction will all be discussed. The data collected by the MEMS cap are specific; not only can researchers determine the date and time of bottle opening, but they can also measure the duration of the opening and the hours between dosages (Marques and McKnight, 2007). This information is entirely computerized, and errors from transcription, recall bias and biases involved in self-report are eliminated.

MEMS is useful in improving and monitoring patient adherence. Cramer *et al*. (2003) found that over half of the patients reported that the digital displays of the number of hours since the last opening and the total daily openings were useful in properly taking dosages of naltrexone. This study demonstrated that visual reminders on the caps contribute to improving medication adherence among participants. The increased medical adherence via MEMS has been demonstrated across many clinical trials (Rivers *et al.*, 1998; de Bruin *et al.*, 2005; Vriensendorp *et al.*, 2007; Knafl *et al.*, 2010). Furthermore, MEMS caps may include a signal to the patient to remind him/her to take the medication at a pre-determined time, thus further improving patient's adherence.

On the other hand, MEMS is a device that exclusively provides information on medication adherence, and cannot provide insight into monitoring alcohol use nor collect other types of outcome measures. Furthermore, while MEMS measures when bottles are opened, it cannot assess whether the medication was actually consumed. In theory, a patient could open the bottle and discard the dose. However, it has been demonstrated that if a participant opens a bottle it is likely that she/he will take the medication. If they wish to be non-compliant, participants will simply stop opening the bottle and stop consuming the medication (Cramer and Rosenheck, 1999).

Because MEMS is unable to label missed doses that are rectified by the patient taking a double dose at a later time, patient diaries are still essential in clinical trials (Olivieri *et al.*, 1991). However, it can be noted that pill counts are also unable to distinguish between missed and double doses. Another limitation is that if a patient opens the MEMS cap after midnight, the data will be recorded as an event that happened the following day; however, this limitation can easily be rectified in data cleaning; therefore it does not represent a sizeable limitation for the use of MEMS.

An additional concern is cap malfunctions; for example, in a study of 17 patients treated for HIV infection and using MEMS caps, 4 patients required a replacement of the cap. This is not only costly, but also causes data points to be missed (Martin *et al.*, 2009). A possible improvement of this weakness could be obtained by developing a wireless system that permits data transfer in real time. A wireless system might allow investigators, in case they do not receive data transfer in a certain day, to contact the patient and verify if pills were taken or not if there is a malfunction; such a system, however, might not be practical in the real world. Additionally, MEMS can only measure pill consumption and cannot measure drinking behaviors. Ideally, a technology-based assessment that is utilized in trials involving alcoholic patients would be able to provide information on both variables. On the other hand, MEMS is the only technology discussed that specifically focuses on pill consumption and may be necessary to employ if other technologies cannot provide such information. Finally, another drawback to be noted is the high cost of MEMS (Martin *et al.*, 2009).

ECOLOGIC MOMENTARY ASSESSMENT

EMA assesses clinical trial participants through the use of hand-held data-collection devices (Epstein and Preston, 2003). The device can range from telephones to electronic diaries (palm-top computers) to, in the most simple of cases, written diaries (Shiffman, 2009). EMA allows researchers to gain insight into medication adherence and events, such as alcohol use and/or other outcome measures. Participants must respond to a survey on their cell phone, telephone, laptop or palm pilot. These data are then downloaded or wirelessly transmitted to a computer. There are several designs for EMA data

collection. One format is event-based recording that captures events in real time. The second design relies on time-based assessments, i.e. subjects are asked questions at a specific time and must retrospectively answer about the events that have occurred since the previous assessment (Shiffman, 2009). The third type involves the recording of data throughout random points determined by the researcher. This means that the participant fills out the EMA, in real time, when prompted. The data may or may not coincide with an event. For example, it evaluates the mood of the participant at a point in time, but may also ask the participant to retrospectively describe alcohol consumption occasions. EMA can be considered a form of PROs, and must therefore be carefully designed to provide meaningful data and to limit the possibility of skewed qualitative results (Basch *et al.*, 2011).

Unlike the random and standardized sampling, the event-sampling procedure allows investigators to collect data in real time, therefore the potential issue of recall bias is eliminated (however, it should be noted that random and standardized sampling procedures can query patients about their present state, therefore the weakness of recall bias is not universally applicable to these sampling procedures). As such, in theory the event-sampling procedure might be the best, but it does rely on participants' compliance to collect the requested data when the event happens. In contrast, the random or standardized sampling procedures allow investigators to sample moments based on a regular or random time schedule, and the use of an EMA computer sample events at random ensures that the resulting assessments are representative (Shiffman, 2009). Furthermore, standardized time samples for EMA data collection provide some advantages in terms of statistical modeling. In summary, while standardized time-based assessments have been often written off as not particularly valuable, in reality different sampling procedures hold both strengths and weaknesses; as such, the best EMA sampling method should be decided based specifically on the question(s) that the investigators are asking. Additionally, event-related and standard-time intervals can be combined in useful ways (e.g. 15 min, 30 min and 1 h into a drinking episode) and this approach may overall represent a significant improvement of the EMA methodology.

EMA involves repeated sampling, and is hence a longitudinal technology that examines and documents temporal sequences and cascades of events. Forms of EMA, including wristwatches and palm-top computers, sound an alarm to prompt data recording (McGeary *et al.*, 2006). The alarm clock function in EMA is also used to eliminate prompting when the subject is sleeping. Prompts can be scheduled randomly throughout the day so that the data collected paint a multi-faceted image of the subjects' daily experience (Stone, 2000). Henker *et al.* (2002) utilized Palm III hand-held computers to collect their data, e.g. moods, activities, social settings, dietary intake, smoking and alcohol use. While not specifically focusing on alcohol abuse, the study was a confirmation that EMA could be successful in collecting specific information. Collins *et al.* (1998) studied participants with self-reported excessive drinking using an electronic diary on a small hand-held computer to enter data at random moments throughout the day, in addition to entering data whenever they consumed alcohol. Collins *et al.* (2003) also demonstrated the feasibility of collecting EMA data on alcohol consumption using cellular phones with an IVR system, which demonstrates the interdependence of both IVR and EMA. Recent technologic advancements have allowed for the implementation of EMA into cell phones, mobile devices and smartphones. For example, Mays *et al.* (2010) examined the feasibility of using wireless mobile devices to collect daily alcohol information among college students. Their study indicated that the use of mobile devices produced results comparable to paper-based assessments (Mays *et al.*, 2010). EMA in mobile devices has also been utilized successfully in patients with other addictions, e.g. crack-cocaine addicts (Freedman *et al.*, 2006).

Strengths and weaknesses of EMA are discussed in the context of its use, the type of information collected, ease of collection and potential for bias. Electronic diaries are popular in clinical trials as a result of the ease with which a participant can enter a data point, as well as the fact that the computer adds a time and date stamp (Raymond and Ross, 2000). A study examining naltrexone effects on drinking utilized EMA and found that the ability to conduct multiple assessments continuously over 35 days strengthens the power to test the effects of naltrexone and provides the power to test the effects of potential moderators (Tidey *et al.*, 2008).

Information regarding craving, mood, drinking triggers and drinking behavior can be collected without the need for retrospective analysis. In addition, by implementing EMA data collection, gaps in data collection are minimized and the potential for invalid responses is eliminated. Palm-computer EMA does not permit a participant to skip questions, and offers responses in a fixed multiple choice format so that the response would be valid. There is hence quality control of data (Stone, 2000; Shiffman *et al.*, 2008). EMA also prevents participants from faking an entry because the entries are individually recorded with a date and time stamp. Adherence is also tracked because failures to respond are recorded (Shiffman *et al.*, 2008). The Palm computer also allows for data entry even when not prompted. A participant can enter craving at any point in the day, hence expanding data available for analyses (Gwaltney *et al.*, 2008). There are, however, some weaknesses with EMA that need to be highlighted as well. Participants must be trained extensively to familiarize them with the technology. The device and computer software are also costly. Thus, EMA is both time and cost intensive because of the technology involved (Raymond and Ross, 2000; Freedman *et al.*, 2006; Shiffman *et al.*, 2008). The data must also be extensively managed, which can drive up the cost. The participant must come into the research facility on a weekly basis to download the data and to also change the batteries (Freedman *et al.*, 2006). This is inconvenient and might impact patient participation. As is the case with using any sort of technology, there is always the risk of technical problems (Raymond and Ross, 2000; Gwaltney *et al.*, 2008). It should be noted, however, that some of these weaknesses have been recently overcome by the use of EMA via cell phones or other more recently developed platforms that reduce both costs and time consumed for training and in-person visits.

Another potential irreparable weakness is the fact that quality and reliability of responses decreases with drinking. But, before drinking, the data are believed to be reliable and can provide insight into triggers of drinking (Litt *et al.*, 1998). There has been skepticism that alcoholic participants will not comply with protocols, will steal or break the PDA's or will provide falsified information. However, studies of homeless crack-cocaine addicts and regular ecstasy users have disproved

such concerns, for the number of devices lost was minimal and illicit behavior was regularly lacking (Epstein *et al.*, 2009).

CONCLUSION

Although not without important limitations and weaknesses, technology-based techniques, can be utilized to significantly improve the quality and accuracy of alcohol-related results in clinical trials. For example, in the systematic review mentioned before, only 3 trials (14%) met criteria for high levels of adherence assurance, 5 (23%) met medium adherence assurance criteria and 14 (64%) met low adherence criteria (Swift *et al.*, 2011). On the other hand, an alcoholism trial with naltrexone using MEMS reported high medication compliance in all participants (Anton *et al.*, 2005).

An additional important aspect to consider is the possibility of combining technology-based assessment methods with standardized approaches. For example, contingency management in addiction programs has shown a large statistical effect (Schumacher *et al.,* 2007), and the inclusion of contingency management procedures to programs that use technology-based methods may improve alcohol-related outcomes (e.g. Barnett *et al.*, 2011).

Current methods for data collection and medication adherence measurement pose issues such as recall bias and missing data points. Technology-based assessments aim to address these problems, and hence hold the possibility for eliminating error and bias from clinical pharmacotherapy trials of alcoholism (Table 2).

Another important aspect to consider is that technology-based methods may actually become an intervention. For example, IVR may enhance brief interventions for alcohol (Helzer *et al.*, 2008) and reduce relapse following discharge from residential treatment (Mundt *et al.*, 2006). This might represent an additional advantage of using these techniques, but, on the other hand, it might add a confounding factor that increases the difficulty to evaluate the real effect of a pharmacotherapy. This aspect, however, is controversial and deserves additional studies in order to be fully addressed.

An important aspect is to consider which techniques are useful for the type of data interested in. For example, if one is most interested in medication adherence rates, then the MEMS technique is recommended because of the fact that it is objective and not subject to bias. If one is most interested in obtaining data on alcohol use, then the TDS technique is recommended because of the fact that it is the most objective in terms of information provided. However, the TDS cannot provide exact time of alcohol consumption and so can at times be interchanged with EMA. The selection of technique is subject to the nature of the participant, most important question(s) the investigators want to address and budget restriction. EMA should be utilized to evaluate triggers of alcohol consumption, as it prompts the participant to fill out detailed questionnaires in real time. As such, future research may also focus on the strengths, weaknesses and cost-effectiveness of the combination of several of these techniques in the same samples enrolled in a clinical trial.

FUTURE DIRECTIONS

Daily improvements are made to internet, computer devices, cellular devices and software. With such computer science growth, it would be feasible to improve the discussed devices and make them more convenient for average users. As a result, the use of these technology-based assessments has only just reached its infancy. The hope is that improvements can serve to drive down the cost of these technologies.

In general, it is possible to argue that measuring and maintaining a very high adherence during a clinical trial will not translate into ‘real life’ adherence. In the translation from clinical trials to clinical practice, the technology-based assessments described will not be available to the average consumer purchasing a prescribed pharmacotherapy for alcoholism. Thus, medication adherence will never be as high in daily life. However, clinical trials are meant to define the efficacy of the treatment and to determine a safe dosage. The described technologies are necessary to prevent false negatives in validating drug efficacy. It is possible that drugs are discarded as ineffective in treating a disorder, when in actuality, the adherence in

Table 2. Strengths and weaknesses of the technology-based assessments discussed

Technology	Strengths	Weaknesses
IVR	Automatic	Answers socially biased
	Convenient and allows for flexibility	Extensive software necessary
	Provides information on patient compliance and drinking behavior	Financial incentives delayed
	Reaches a larger population	Tone can influence response
	Collects in real time and prevents patient bias	
TDS	Objective and quantitative	Does not provide information on medication compliance
	Continuous and instantaneous	Expensive
	Eliminates bias	May cause some discomfort and embarrassment
		Can malfunction
MEMS	Specific information on medication compliance	Cannot verify medication consumption
	Digital displays improve medication compliance	Does not provide information on drinking behavior
	Eliminates recall bias and transcription errors	High cost
		Potential for cap malfunctions
EMA	Specific (time, date stamp)	Software can malfunction
	Can assess drinking behavior and medication compliance	Requires extensive training for use
	Real time to eliminate recall bias	Expensive
	Prevents faking compliance and incomplete answers	Answer reliability decreases with drinking behaviors
	Provides information on mental state	

the trial was low and inaccurately measured. In other words, technologic methods such as EMA, IVR and TDS may increase the accuracy of data collections as these techniques reduce false negatives by dramatically increasing the number of data points in a study, thus improving statistical power. Furthermore, MEMS and IVR may reduce false negatives through improving medication compliance.

In summary, it is important to keep in mind that the use of rigorous adherence monitoring methods, such as those described in the present review, may reduce the possibility of Type II error and increase the internal validity of a clinical trial. On the other hand, however, this rigorous adherence monitoring may reduce the external validity of a clinical trial, so that its results may not generalize to 'real-world' clinical practice settings (Swift *et al.*, 2011). As such, the use of technology-based assessments to address the lack of medication adherence in clinical trials is the first step in addressing the overarching problem.

Funding — Dr Leggio's research is supported by the Division of Intramural Clinical and Biological Research of the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Intramural Research Program of the National Institute on Drug Abuse (NIDA).

Conflict of interest statement. None declared.

REFERENCES

Abu-Hasaballah K, James A, Aseltine R. (2007) Lessons and pitfalls of interactive voice response in medical research. *Contemp Clin Trials* **28**:593–602.

Alemi F, Stephens R, Parran T *et al*. (1994) Automated monitoring of outcomes: application to treatment of drug abuse. *Med Decis Making* **14**:180–7.

Anton RF, Moak DH, Latham P *et al*. (2005) Naltrexone combined with either cognitive behavioral or motivational enhancement therapy for alcohol dependence. *J Clin Psychopharmacol* **25**:349–57.

Babiker IE, Cooke PR, Gillett MG. (1989) How useful is riboflavin as a tracer of medication compliance*? J Behav Med* **12**:25–38.

Bardone AM, Krahn DD, Goodman BM *et al*. (2000) Using interactive voice response technology and timeline follow-back methodology in studying binge eating and drinking behavior: different answers to different forms of the same question? *Addict Behav* **25**:1–11.

Barnett NP, Tidey J, Murphy JG *et al*. (2011) Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. *Drug Alcohol Depend* **118**:391–9.

Basch E, Abernethy AP, Reeve BB. (2011) Assuring the patient centeredness of patient-reported outcomes: content validity in medical product development and comparative effectiveness research. *Value Health* **14**:965–6.

Besch CH. (1995) Compliance in clinical trials. *AIDS* **9**:1–10.

Collins RL, Morsheimer ET, Shiffman S *et al*. (1998) Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* **6**:306–15.

Collins LR, Kashdan TB, Gollnisch G. (2003) The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Exp Clin Psychopharmacol* **11**:73–8.

Corkrey SR, Parkinson L. (2002) Interactive voice response: review of studies 1989–2000. *Behav Res Methods Instrum Comput* **34**:342–53.

Cramer JA, Rosenheck R. (1999) Enhancing medication compliance for people with serious mental illness. *J Nerv Ment Dis* **187**:53–5.

Cramer J, Rosenheck R, Kirk G *et al*. (2003) Medication compliance feedback and monitoring in a clinical trial: predictors and outcomes. *Value Health* **6**:566–73.

Cranford JA, Tennen H, Zucker RA. (2010) Feasibility of using interactive voice response to monitor daily drinking, moods, and relationship processes on a daily basis in alcoholic couples. *Alcohol Clin Exp Res* **34**:499–508.

Dackis C, O'Brien C. (2005) Neurobiology of addiction: treatment and public policy ramifications. *Nat Neurosci* **8**:1431–6.

de Bruin M, Hospers HJ, van den Borne HW *et al*. (2005) Theory- and evidence-based intervention to improve adherence to antiretroviral therapy among HIV-infected patients in the Netherlands: a pilot study. *AIDS Patient Care STDS* **19**:384–94.

de Leeuw ED. (2005) To mix or not to mix data collection modes in surveys. *J Official Statist* **21**:233–55.

de Leeuw ED, van der Zouwen J. (1988) Data quality in telephone and face to face surveys: a comparative meta-analysis. In Groves RM, Biemer PP, Lyberg LE (eds). *Telephone Survey Methodology*. New York: John Wiley & Sons, Inc., 283–99.

Del Boca FK, Noll JA. (2000) Truth or consequences: the validity of self-report data in health services research on addictions. *Addiction* **95**:347–60.

Dillman DA, Phelps G, Tortora R *et al*. (2009) Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Soc Sci Res* **38**:3–20.

Edwards S, Koob GF. (2010) Neurobiology of dysregulated motivational systems in drug addiction. *Future Neurol* **5**:393–410.

Edwards S, Kenna GA, Swift RM *et al*. (2011) Current and promising pharmacotherapies, and novel research target areas in the treatment of alcohol dependence: a review. *Curr Pharm Des* **17**:1323–32.

Epstein DH, Preston KL. (2003) The reinstatement model and relapse prevention: a clinical perspective. *Psychopharmacology* **168**:31–41.

Epstein DH, Willner-Reid J, Vahabzadeh M *et al*. (2009) Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. *Arch Gen Psychiatry* **66**:88–94.

Freedman MJ, Lester KM, McNamara C *et al*. (2006) Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *J Subst Abuse Treat* **30**:105–11.

Gwaltney CJ, Bartolomei R, Colby SM *et al*. (2008) Ecological momentary assessment of adolescent smoking cessation: a feasibility study. *Nicotine Tob Res* **10**:1185–90.

Hawthorne JS, Wojcik MH. (2011) Transdermal alcohol measurement: a review of the literature. *Can Soc Forensic Sci J* **39**:65–71.

Haynes R. (1987) Patient compliance and the conduct and interpretation of therapeutic trials. *Control Clin Trials* **8**:12–9.

Helzer JE, Rose GL, Badger GJ *et al*. (2008) Using interactive voice response to enhance brief alcohol intervention in primary care settings. *J Stud Alcohol Drugs* **69**:251–8.

Henker B, Whalen CK, Jamner LD *et al*. (2002) Anxiety, affect, and activity in teenagers: monitoring daily life with electronic diaries. *J Am Acad Child Adolesc Psychiatry* **41**:660–70.

Jordan J, Teschendorf B, Langel K. (2011) Which ePRO solution is right for your study? *CRF Health* 1–8.

Kaminer Y, Litt MD, Burke RH *et al*. (2006) An interactive voice response (IVR) system for adolescents with alcohol use disorders: a pilot study. *Am J Addict* **15**(Suppl. 1):122–5.

Knafl GJ, Bova CA, Fennie KP *et al*. (2010) An analysis of electronically monitored adherence to antiretroviral medications. *AIDS Behav* **14**:755–68.

Koob GF, Le Moal M. (2001) Drug addiction, dysregulation of reward, and allostastis. *Neuropsychopharmacology* **24**:97–129.

Kranzler HR. (2000) Pharmacotherapy of alcoholism: gaps in knowledge and opportunities for research. *Alcohol* **35**:537–47.

Kranzler HR, Abu-Hasaballah K, Tennen H *et al*. (2004) Using daily interactive voice response technology to measure drinking and related behaviors in a pharmacotherapy study. *Alcohol Clin Exp Res* **28**:1060–4.

Kranzler HR, Armeli S, Tennen H. (2012) Post-treatment outcomes in a double-blind, randomized trial of sertraline for alcohol dependence. *Alcohol Clin Exp Res* **36**:739–44.

Leggio L, Cardone S, Ferrulli A *et al*. (2010) Turning the clock ahead: potential preclinical and clinical neuropharmacological targets for alcohol dependence. *Curr Pharm Des* **16**:2159–81.

Lingford-Hughes A, Nutt D. (2003) Neurobiology of addiction and implications for treatment. *Br J Psychiatry* **182**:97–100.

Litt MD, Cooney NL, Morse P. (1998) Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* **17**:48–52.

Malta M, Strathdee SA, Magnanini MM *et al*. (2008) Adherence to antiretroviral therapy for human immunodeficiency virus/acquired immune deficiency syndrome among drug users: a systematic review. *Addiction* **103**:1242–57.

Marques PR, McKnight SA. (2007) United States National Highway Traffic Safety Administration & Pacific Institute for Research and Evaluation. Evaluating transdermal alcohol measuring devices: final report. www.nhtsa.gov.

Marques P, McKnight AS. (2009) Field and laboratory alcohol detection with two types of transdermal devices. *Alcohol Clin Exp Res* **33**:703–11.

Martin S, Elliott-DeSorbo DK, Calabrese S *et al*. (2009) A comparison of adherence assessment methods utilized in the United States: perspectives of researchers, HIV-infected children, and their caregivers. *AIDS Patient Care STDS* **23**:593–601.

Mays D, Cremeens J, Usdan S *et al*. (2010) The feasibility of assessing alcohol use among college students using wireless mobile devices: implications for health education and behavior research. *Health Educ J* **69**:311–20.

McGeary JE, Monti PM, Rohsenow DJ *et al*. (2006) Genetic moderators of naltrexone's effects on alcohol cue reactivity. *Alcohol Clin Exp Res* **30**:1288–96.

Midanik LT, Greenfield TK. (2010) Reports of alcohol-related problems and alcohol dependence for demographic subgroups using interactive voice response (IVR) versus telephone (CATI) surveys: the 2005 U.S. National Alcohol Survey. *Drug Alcohol Rev* **29**:392–8.

Mundt JC, Moore HK, Bean P. (2006) An interactive voice response program to reduce drinking relapse: a feasibility study. *J Subst Abuse Treat* **30**:21–9.

Olivieri NF, Matsui D, Hermann C *et al*. (1991) Compliance assessed by the medication event monitoring system. *Arch Dis Child* **66**:1399–402.

Osterberg L, Blaschke T. (2005) Adherence to medication. *N Engl J Med* **353**:487–97.

Perrine MW, Mundt JC, Searles JS *et al*. (1995) Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology. *J Stud Alcohol* **56**:487–90.

Preston KL, Silverman K, Schuster CR *et al*. (1997) Assessment of cocaine use with quantitative urinalysis and estimation of new uses. *Addiction* **92**:717–27.

Raymond SA, Ross RN. (2000) Electronic subject diaries in clinical trials. *Appl Clin Trials* **9**:48–58.

Rivers PH, Ardagh-Walter N, Wright EC. (1998) Measurement of anticonvulsant adherence behaviour in the community using a Medication Events Monitoring System (MEMS). *Health Care Anal* **6**:308–16.

Sakai JT, Mikulich-Gilbertson SK, Long RJ *et al*. (2006) Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. *Alcohol Clin Exp Res* **30**:26–33.

Schumacher JE, Milby JB, Wallace D *et al*. (2007) Meta-analysis of day treatment and contingency-management dismantling research: Birmingham Homeless Cocaine Studies (1990–2006). *J Consult Clin Psychol* **75**:823–8.

Searles JS, Helzer JE, Walter DE. (2000) Comparison of drinking patterns measured by daily reports and timeline followback. *Psychol Addict Behav* **14**:277–86.

Searles JS, Helzer JE, Rose GL *et al*. (2002) Concurrent and retrospective reports of alcohol consumption across 30, 90 and 366 days: interactive voice response compared with the timeline follow back. *J Stud Alcohol* **63**:352–62.

Shiffman S. (2009) Ecological momentary assessment (EMA) in studies of substance use. *Psychol Assess* **21**:486–97.

Shiffman S, Stone AA, Hufford MR. (2008) Ecological momentary assessment. *Annu Rev Clin Psychol* **4**:1–32.

Simpson TL, Saxon AJ, Meredith CW *et al*.. (2009) A pilot trial of the alpha-1 adrenergic antagonist, prazosin, for alcohol dependence. *Alcohol Clin Exp Res* **33**:255–63.

Sobell LC, Sobell MB. (1992) Timeline followback: a technique for assessing self-reported alcohol consumption. In Litten RZ, Allen JP (eds). *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*. Totowa: Humana, 41–72.

Spilker B. (1991) Teaching courses in clinical trial research methods. *J Clin Pharmacol* **31**:496–508.

Stone AA. (2000) Real-time self report of momentary states in the natural environment: computerized ecological momentary assessment: Chapter 16. The Science of Self-report: Implications for Research and Practice. Mahwah, NJ: Lawrence Erlbaum.

Swift RM. (2003) Direct measurement of alcohol and its metabolites. *Addiction* **98**(Suppl. 2):73–80.

Swift RM, Martin CS, Swette L *et al*. (1992) Studies on a wearable, electronic, transdermal alcohol sensor. *Alcohol Clin Exp Res* **16**:721–5.

Swift R, Oslin DW, Alexander M *et al*. (2011) Adherence monitoring in naltrexone pharmacotherapy trials: a systematic review. *J Stud Alcohol Drugs* **72**:1012–8.

Tennen H, Affleck G. (1996) Daily processes in coping with chronic pain: methods and analytic strategies. In Zeidner M, Endler NS (eds). *Handbook of Coping*. New York: Wiley, 151–80.

Tidey JW, Monti PM, Rohsenow DJ *et al*. (2008) Moderators of naltrexone's effects on drinking, urge, and alcohol effects in non-treatment-seeking heavy drinkers in the natural environment. *Alcohol Clin Exp Res* **32**:58–66.

Toll BA, Cooney NL, McKee SA *et al*. (2006) Correspondence between interactive voice response (IVR) and timeline followback (TLFB) reports of drinking behavior. *Addict Behav* **31**:726–31.

Toll BA, Cooney JL, McKee SA *et al*. (2008) Correspondence of interactive voice response (IVR) reports of nicotine withdrawal, craving, and negative mood with questionnaire ratings. *Nicotine Tob Res* **10**:1057–64.

Tourangeau R, Steiger DM, Wilson D. (2002) Self-administered questions by telephone: evaluating interactive voice responses. *Public Opin Q* **66**:265–78.

Trivedi MH, Greer TL, Potter JS *et al*. (2011) Determining the primary endpoint for a stimulant abuse trial: lessons learned from STRIDE (CTN 0037). *Am J Drug Alcohol Abuse* **37**:339–49.

Tucker JA, Vuchinich RE, Rippens PD. (2002) Predicting natural resolution of alcohol-related problems: a prospective behavioral economic analysis. *Exp Clin Psychopharmacol* **10**:248–57.

Vriensendorp R, Cohen A, Kristanto P *et al*. (2007) Adherence to HAART therapy measured by electronic monitoring in newly diagnosed HIV patients in Botswana. *Eur J Clin Pharmacol* **63**:1115–21.

Weiss RD. (2004) Adherence to pharmacotherapy in patients with alcohol and opioid dependence. *Addiction* **99**:1382–92.

Wilke RJ, Burke LB, Erickson P. (2004) Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved product labels. *Control Clin Trials* **25**:535–52.

Young LM, Haakenson CM, Lee KK *et al*. (1984) Riboflavin use as a drug marker in veterans administration cooperative studies. *Control Clin Trials* **5**(4 Suppl.):497–504.