Received: 17 August 2020 | Accepted: 1 March 2021

DOI: 10.1111/acer.14597

ALCOHOLISM CLINICAL & EXPERIMENTAL RESEARCH

ORIGINAL ARTICLE

Accuracy of Consumer-marketed smartphone-paired alcohol breath testing devices: A laboratory validation study

Mucio Kit Delgado[1,2,3,4] | Frances Shofer[1] | Reagan Wetherill[5] | Brenda Curtis[6] | Jessica Hemmons[1] | Evan Spencer[1] | Charles Branas[3,7] | Douglas J. Wiebe[2,3,4] | Henry R. Kranzler[5,8]

[1]Behavioral Science & Analytics For Injury Reduction (BeSAFIR) Lab, Department of Emergency Medicine & the Center for Health Incentives and Behavioral Economics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[2]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[3]Penn Injury Science Center, University of Pennsylvania, Philadelphia, PA, USA

[4]Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA

[5]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[6]Technology and Translational Research Unit, National Institute on Drug Abuse, National Institutes of Health, Baltimore, MD, USA

[7]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA

[8]VISN 4 Mental Illness Research, Education, and Clinical Center (MIRECC), Crescenz VA Medical Center, Philadelphia, PA, USA

Correspondence
M. Kit Delgado, University of Pennsylvania, 933 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, USA.
Email: mucio.delgado@pennmedicine.upenn.edu

Funding information
Research reported in this publication was supported by the Centers for Disease Control and Prevention under award R49 CE002474 and the National Institutes of Health under award numbers K23 HD090272001 and UL1 TR001878. This research was also supported in part by the VISN 4 Mental Illness Research, Education and Clinical Center of the Department of Veterans Affairs and Abramson Family Foundation Fund for Acute Care and Injury Prevention Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Centers for Disease Control and Prevention, the National Institutes of Health, the Department of Veterans Affairs, Abramson Family Foundation, or the US government.

Abstract

Background: Although alcohol breath testing devices that pair with smartphones are promoted for the prevention of alcohol-impaired driving, their accuracy has not been established.

Methods: In a within-subjects laboratory study, we administered weight-based doses of ethanol to two groups of 10 healthy, moderate drinkers aiming to achieve a target peak blood alcohol concentration (BAC) of 0.10%. We obtained a peak phlebotomy BAC and measured breath alcohol concentration (BrAC) with a police-grade device (Intoxilyzer 240) and two randomly ordered series of 3 consumer smartphone-paired devices (6 total devices) with measurements every 20 min until the BrAC reached <0.02% on the police device. Ten participants tested the first 3 devices, and the other 10 participants tested the other 3 devices. We measured mean paired differences in BrAC with 95% confidence intervals between the police-grade device and consumer devices.

Results: The enrolled sample (N = 20) included 11 females; 15 white, 3 Asian, and 2 Black participants; with a mean age of 27 and mean BMI of 24.6. Peak BACs ranged from 0.06–0.14%. All 7 devices underestimated BAC by >0.01%, though the BACtrack Mobile Pro and police-grade device were consistently more accurate than the Drinkmate and Evoc. Compared with the police-grade device measurements, the BACtrack Mobile Pro readings were consistently higher, the BACtrack Vio and Alcohoot measurements similar, and the Floome, Drinkmake, and Evoc consistently lower. The BACtrack Mobile Pro and Alcohoot were most sensitive in detecting BAC driving limit thresholds, while the Drinkmate and Evoc devices failed to detect BAC limit thresholds more than 50% of the time relative to the police-grade device.

Conclusions: The accuracy of smartphone-paired devices varied widely in this laboratory study of healthy participants. Although some devices are suitable for clinical and research purposes, others underestimated BAC, creating the potential to mislead intoxicated users into thinking that they are fit to drive.

KEYWORDS
alcohol breath test, alcohol consumption, blood alcohol concentration, smartphone

INTRODUCTION

Alcohol-impaired driving kills approximately 29 people per day and costs more than $121 billion per year in the United States (National Highway Traffic Safety Administration, 2017; Zaloshnja et al., 2013). After years of progress in reducing alcohol-impaired driving fatalities, efforts began to stall in 2009 and fatalities started increasing again in 2015. As a result, in January 2018, the U.S. National Academies of Sciences, Engineering, and Medicine released a report calling for a reduction in the legal blood alcohol concentration (BAC) for driving from 0.08% to 0.05% to be consistent with most other industrialized countries (National Academies of Sciences, Engineering, and Medicine, 2018). Furthermore, based on several studies demonstrating that drinkers cannot accurately estimate their BAC (Beirness, 1987; Beirness et al., 1993; Martin et al., 2016; Thombs et al., 2003), the report concluded: "Consumer marketed personal breath testing devices are an emerging technology with the potential to reduce alcohol-impaired driving by promoting more accurate BAC self-estimation. However, these technologies require further investigation of their accuracy and effects on behavior before promoting widespread use." (National Academies of Sciences, Engineering, and Medicine, 2018, p. 155).

Handheld alcohol breath testing devices, also known as "breathalyzers," have been used by law enforcement since the 1970 for roadside screening of alcohol intoxication (Chambers et al., 1976) and have been validated against BACs obtained by venipuncture (Schechtman & Shinar, 2011; Van Tassel et al., 2004). Although BrAC as measured by police-grade breath testing devices is highly correlated with BAC, comparison studies show that the BrAC estimates are consistently lower than actual venous BAC by up to 15% (Jones & Andersson, 2003; Kriikku et al., 2014). Beginning in the 2000s, personal breath testing devices have been marketed directly to consumers for self-monitoring of estimated BAC (Ashdown et al., 2014). Typically, users are instructed to blow into the device at least 20 min after their last alcohol consumption and they are provided an estimated BAC within a few seconds. In 2012, France passed a law requiring all drivers to keep a personal alcohol breath testing device in the car at all times (Légifrance, 2012). A study of consumer-marketed personal breath testing devices found wide variability in the accuracy of these devices relative to law enforcement devices, cautioning that they may provide false reassurance to drivers who are intoxicated (Ashdown et al., 2014; Gornall, 2014).

The latest generation of personal alcohol breath testing devices marketed to consumers now pair with smartphone apps. Users blow into the device, and the estimated BAC reading is displayed in the smartphone app. These apps can track readings over time and provide messaging and prompts according to the estimated BAC level.

The consumer appeal and increasing popularity of these devices have been covered broadly in lay media, with specific questions as to the accuracy of the devices (Jolly, 2015). In January 2017, the U.S. Federal Trade Commission successfully filed suit to pull from the market one such device, Breathometer, because of deceptive claims of its accuracy (FTC, 2017). There are now dozens of alcohol breath testing devices sold on online market places such as Amazon.com with limited to no information supporting their accuracy (Alexandra Berzon, 2019). Still, if accurate, the smartphone connectivity of newer personal breath testing devices opens up broad new opportunities for remote monitoring and mobile interventions to reduce hazardous drinking (Luczak & Ramchandani, 2019; Quanbeck et al., 2014; Tofighi et al., 2019). One recent study showed that smartphone-prompted breath testing measurements using the BACTrack Mobile Pro were more accurate than traditional self-report measures of alcohol consumption (Kaplan & Koffarnus, 2019). Some U.S. states are now promoting and providing these devices at a discount for the purposes of reducing alcohol-impaired driving (Colorado DOT, 2020). Therefore, a study validating the broader accuracy of smartphone-paired alcohol breath testing devices is critically needed.

To fill this knowledge gap, we tested the accuracy of consumer-marketed smartphone breath testing devices by comparing them with the true gold standard of peak BACs obtained via venipuncture as well as the more common reference standard of a police-grade handheld breath testing device. Our secondary objective was to determine the sensitivity of smartphone breath testing devices for detecting breath alcohol concentrations (BrACs) above common legal driving limits as measured by a police-grade device.

METHODS

Study design and participants

Between December 13, 2016, and April 17, 2017, we enrolled in a laboratory validation study a convenience sample of moderate drinkers, aged 21–39 years, who reported ≥4 drinking days and ≥12 drinks per week for the past 2 months. This drinking threshold was chosen to yield a positive Alcohol Use Disorders Identification Test-Consumption (AUDIT-C) score indicative of hazardous drinking in the general population, the priority population for alcohol-related public health interventions (Gomez et al., 2005). Participants were recruited using IRB-approved broadcast email messages to staff at the Hospital of the University of Pennsylvania and study flyer postings placed around the Hospital of the University of Pennsylvania. Participants were excluded if they sought or received alcohol use disorder treatment in the past 6 months or planned to seek such treatment, were non-English-speaking, were pregnant, or had previously been told by a medical professional not to consume alcohol due to a medical condition or had a contraindication to consume alcohol based on their prescription medications. We aimed to recruit a total of 20 participants based on our pilot grant study budget and timeline for data collection.

Procedures

We examined the accuracy of the 6 breath alcohol testing devices that pair with smartphones and that had apps in both the iTunes and Android app stores on December 6, 2016: Alcohoot AHT 101 (Alcohoot, LLC), BACtrack Mobile Pro, (BACtrack), BACtrack Vio (BACtrack), Drinkmate (Edge Tech Labs, LLC), DRIVESAFE Evoc (Alcohol Countermeasure Systems, Inc.), and Floome (2045 Tech S.r.l., Venice, Italy). We purchased 2 devices from the Web site for each brand, except for the Alcohoot AHT 101, which was purchased on Amazon.com. We report identifying numbers on each device in the Supporting Information (Table S1). The 2 commercial test devices were rotated between study visits (e.g., participant #1 used BACTrack Mobile Pro device #1, participant #2 used BACTrack Mobile Pro device #2, participant #3 used BACTrack Mobile Pro device #1, etc.

Participants were given 3 doses of 100-proof vodka (1.0 g per kg for men and 0.9 g per kg for women) mixed with orange juice over the course of 70 min (i.e., 1 every 20 min) to produce a target peak BAC of approximately 0.100%. All participants drank at approximately the same rate. Blood for measurement of the peak BAC was drawn 20 min after the third dose of alcohol. The BrAC was measured 20 min after each of the 3 doses of alcohol and was repeated every 20 min thereafter using a portable police-grade breath testing device (Intoxilyzer 240, CMI, Inc) selected from the list of U.S. Department of Transportation evidentiary devices and 3 of the 6 personal breath testing devices (Department of Transportation, 2017). The Intoxilyzer device was calibrated before every study visit. We monitored a prespecified list of adverse events including nausea and vomiting. The first 10 participants to enroll in the study tested the Alchohoot AHT 101, BACtrack Mobile Pro, and DRIVESAFE Evoc, and the second 10 participants tested the BACtrack Vio, Drinkmate, and Floome. During each testing episode, the order of the 4 breath testing devices (1 police-grade and 3 personal) was randomized using 4x4 Latin squares for each participant to minimize bias due to the order of testing. Breath testing continued until a BrAC of 0.02 was reached on the Intoxilyzer 240 (see Table S2 for more details on the testing protocol).

Data analysis

We compared group differences between series of devices tested with 2 sample t-test for continuous variables and Fisher's exact test for categorical variables.

To determine differences between the police-grade and personal breath testing devices, a 2-way analysis of variance with repeated measures was used where both the device and times of BrAC measurement were repeated measures. To determine differences between BrAC breath testing devices and venous BAC at 115 min, a 1-way analysis of variance in repeated measures was used. We plotted mean paired differences in BrAC between test devices in each series over time relative to the Intoxilyzer to demonstrate how these differences changed over the ascending and descending limbs of the alcohol metabolism curve. All pairwise comparisons were performed using Tukey–Kramer tests to adjust for multiple comparisons. To assess agreement between the police-grade device and each personal breath testing device at increasing BrAC, we produced scatter plots demonstrating within-person correlation between paired Intoxilyzer and test device BrAC readings and calculated intraclass correlation coefficients. Additionally, we measured the sensitivity of devices for detecting driving limit thresholds by tabulating the proportion of measurements in which the personal breath testing device readings reached or exceeded thresholds of ≥0.05% and ≥0.08% when the police-grade breath testing device reported BrAC levels above these thresholds. All analyses were performed using SAS statistical software (version 9.4, SAS Institute, Cary NC).

The University of Pennsylvania Institutional Review Board approved the study, and the study was registered on ClinicalTrials.gov NCT04086576. This manuscript follows the TREND checklist reporting requirements.

RESULTS

Of 32 individuals screened for participation, 20 were consented and enrolled (Figure 1). Of the 20 participants, 11 (55%) were female, 15 were white (75%), 3 were Asian (15%), 2 were Black (10%) and 1 was of Hispanic ethnicity (5%) (Table 1). The mean age was 27 (range 22–32), and mean BMI was 24.6 (range 18.6–35.6). The mean participant peak phlebotomy BAC was 0.10% (range 0.06–0.14%). The mean

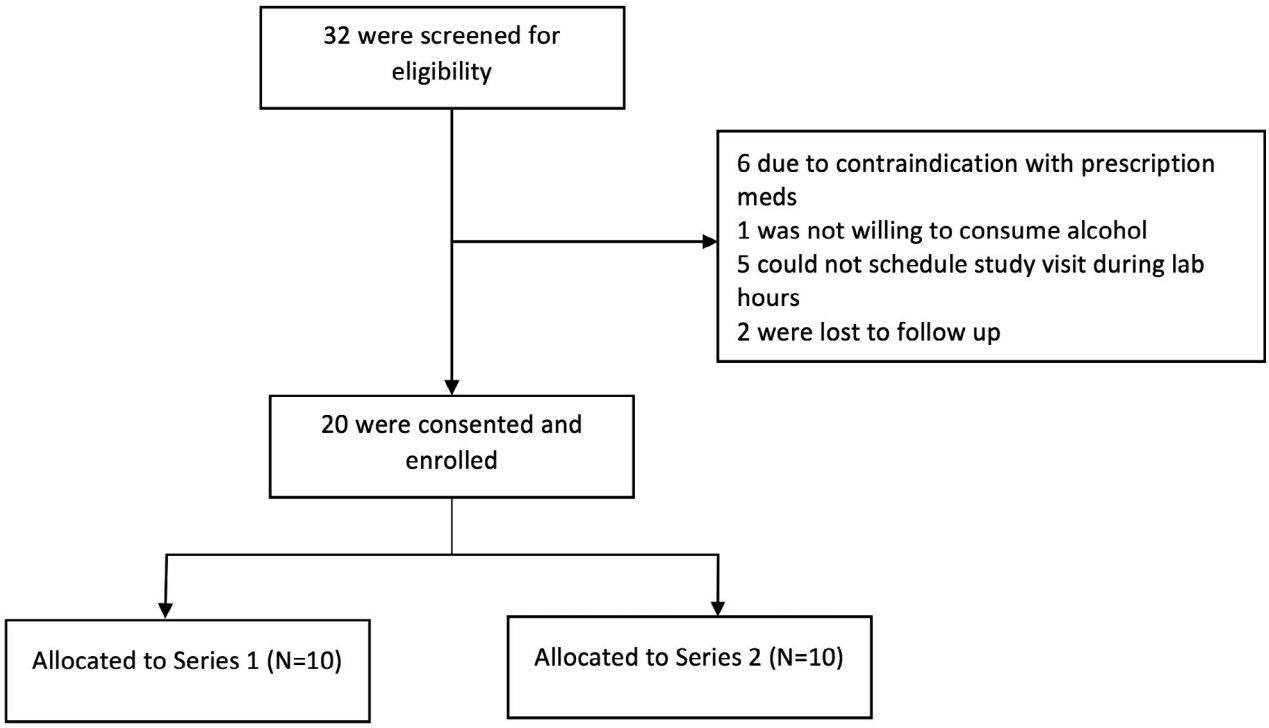


FIGURE 1 Enrollment Diagram. Series 1: Alcohoot, BACtrack Mobile Pro, and DRIVESAFE Evoc. Series 2: BACtrack Vio, Drinkmate, and Floome

TABLE 1 Participant characteristics

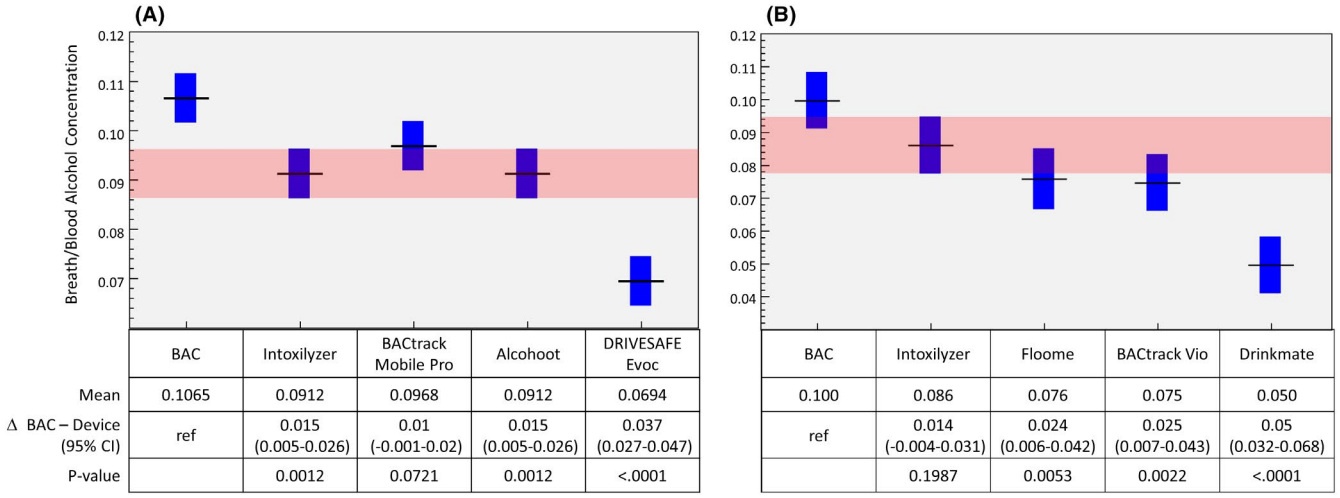
Patient demographic	All participants	Series 1	Series 2	*p*-value
Age (mean ± SD)	26.8 ± 3.2	27.2 ± 3.9	26.3 ± 2.5	0.54
Race				
White	15 (75%)	7 (70%)	8 (80%)	>0.999
Black/African American	2 (10%)	1 (10%)	1 (10%)	
Asian	3 (15%)	2 (20%)	1 (10%)	
Sex				
Female	11 (55%)	5 (50%)	6 (60%)	>0.999
BMI (mean ± SD)	24.6 ± 4.3	24.4 ± 4.6	24.7 ± 4.3	0.386
Peak blood BAC (mean (range))	0.103 (0.061–0.137)	.107 (0.081–0.137)	.100 (0.061 –0.118)	0.386
Peak BrAC police device (mean (range))	0.089 (0.049–0.107)	0.091 (0.079–0.105)	.086 (0.049–0.107)	0.394

Series 1: Alcohoot, BACtrack Mobile Pro, and DRIVESAFE Evoc. Series 2: BACtrack Vio, Drinkmate, and Floome. SD: standard deviation. BMI: body mass index. BAC: blood alcohol concentration drawn from phlebotomy. Group differences between series of devices tested were compared with 2 sample *t*-test for continuous variables and Fisher's exact test for categorical variables.

number of breath testing rounds per participant after receiving the first dose of alcohol until the participant's alcohol level declined to a police-grade breathalyzer BrAC of between 0.03% and 0.02% was 13.3 (SD ±2.97). There were no significant group differences between series (Table 1). There were no adverse events.

All breath testing devices, including the police-grade device, underestimated the phlebotomy BAC by a mean of 0.01% or more (Figure 2a,b). The devices closest to phlebotomy BAC were the BACtrack Mobile Pro and police-grade device with 95% of measurements underestimating the BAC by no more than 0.02%. Measurements using the BACtrack Mobile Pro device were not statistically different from the phlebotomy BAC (p = 0.072) or the police-grade device in series 1 (p = 0.197). In contrast, 95% of the Drinkmate and DRIVESAFE Evoc measurements underestimated the BAC by at least 0.02% or more, with the mean estimates being 0.04% below the peak BAC.

Relative to the police-grade measurements, BACtrack Mobile Pro readings were higher, but significantly so only at 75 (0.0698 vs 0.0597, p =.0003) and 210 min (0.0539 vs 0.0461, p =0.0382). The BACtrack Vio and Alcohoot measurements were similar to the police-grade device where only the BACtrack Vio was significantly lower than the police-grade device at 115 min (0.075 vs 0.086, p =0.03). Floome was



	BAC	Intoxilyzer	BACtrack Mobile Pro	Alcohoot	DRIVESAFE Evoc
Mean	0.1065	0.0912	0.0968	0.0912	0.0694
Δ BAC – Device (95% CI)	ref	0.015 (0.005-0.026)	0.01 (-0.001-0.02)	0.015 (0.005-0.026)	0.037 (0.027-0.047)
P-value		0.0012	0.0721	0.0012	<.0001

BAC	Intoxilyzer	Floome	BACtrack Vio	Drinkmate
0.100	0.086	0.076	0.075	0.050
ref	0.014 (-0.004-0.031)	0.024 (0.006-0.042)	0.025 (0.007-0.043)	0.05 (0.032-0.068)
	0.1987	0.0053	0.0022	<.0001

FIGURE 2 Difference in breath alcohol concentration (BrAC) from police-grade (Intoxilyzer 240) and consumer smartphone-paired breath testing devices relative to Blood Alcohol Concentration (BAC). Horizontal black line represents point estimate; blue bars represent 95% confidence interval. Pink band represents 95% of confidence interval of the police-grade device for comparison to test devices

significantly different than the police-grade device at higher BrAC levels, and both Drinkmate and DRIVESAFE Evoc were consistently significantly lower regardless of BrAC level (Figure 3a,b). Upon inspection of participant-level plots of test device BrAC relative to the police-grade device, there were no significant differences in a given test device BraC relative to police-grade device BrAC across participants (e.g., difference Drinkmate BrAC and Intoxilyzer BrAC in participant 1 vs. participant #2). This suggests that rotating test devices every other participant produced consistent results by test device.

Scatter plots and intraclass correlation coefficients (r_i) demonstrated excellent agreement between the police-grade device and BACtrack Mobile Pro (r_i = 0.916) and good agreement for the Alcohoot (r_i = 0.892) and Floome (r_i = 0.856) (Figure 4) (Koo & Li, 2016). All devices excluding BACtrack Mobile Pro were lower at higher BrAC concentrations (Floome DRIVESAFE Evoc, Drinkmate) or had greater variation from the police-grade device throughout (Alcohoot and BACtrack Vio).

Finally, the BACtrack Mobile Pro and Alcohoot were most sensitive for detecting BAC driving limit thresholds relative to the police-grade device while the Drinkmate and DRIVESAFE Evoc devices failed to detect BAC limit thresholds more than 50% of the time (Figure 5).

DISCUSSION

In this laboratory validation study, we found that the accuracy of smartphone-paired alcohol breath testing devices varied widely. Some devices are suitable for general public use and for clinical and research purposes. For example, BACtrack Mobile Pro came closest to the BAC and was comparable to the police-grade device. On the other hand, devices such as Drinkmate and DRIVESAFE Evoc, which are marketed as safety devices, dangerously underestimated BAC and could mislead users to think that they are fit to drive. These devices failed to detect BrAC levels of 0.08% as measured by a police-grade device more than 50% of the time. Since the completion of the study, one of the devices, Drinkmate, was discontinued and is no longer sold. Other models have been replaced by newer technologies (e.g., BACtrack Vio, Alcohoot).

Given the adverse public health consequences associated with underestimating BAC, our findings suggest that these devices should be more closely regulated governmentally. This is consistent with the findings of a prior study that analyzed the accuracy of an earlier generation of personal alcohol breath testing devices available in the United Kingdom in 2012 (Ashdown et al., 2014; Gornall, 2014). Our study adds to this work by not only demonstrating the variability in accuracy of newer generation smartphone-paired devices relative to a reference standard of a police-grade device, but also by comparing the accuracy relative to the gold standard of BAC obtained by venipuncture. As with previous studies, we also found that BrAC measurements as measured with police-grade device were well correlated BAC, but that BrAC measurements by the police-grade device were also consistently lower by up to 15% than those obtained from venous blood (Jones & Andersson, 2003; Kriikku et al., 2014). We found that BACtrack Mobile Pro measurements had the highest correlation of any test device with the police-grade device and were on average higher and closer to venous BAC than the police-grade device. On the other hand, Drinkmate and DRIVESAFE Evoc were consistently lower than the police-grade device, particularly at peak BrAC levels.

Beyond the devices tested in our study, there are dozens of alcohol breath testing devices sold on online market places such Amazon.com with limited to no information supporting their accuracy or even the origins of where they are produced (Alexandra Berzon, 2019). Additional regulatory strategies may be needed, as many of those devices are manufactured outside the United States and are not subject to FDA oversight. Therefore, at a minimum, we believe that the Food and Drug Administration should enforce existing premarket notification requirements for alcohol breath testing devices, which include registering company address and contact

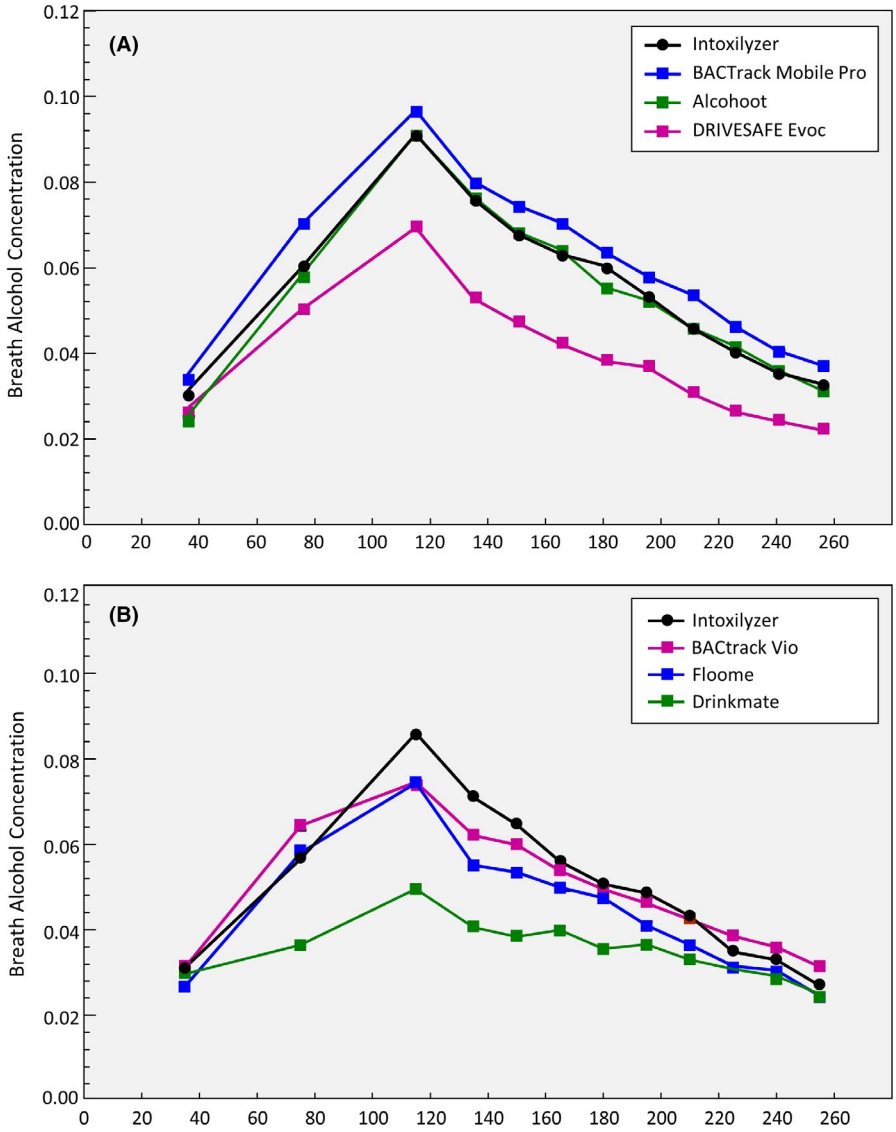


FIGURE 3 Mean paired difference in breath alcohol concentration (BrAC) between smartphone-paired breath testing devices and police-grade device (Intoxilyzer 240) across all time points. Increasingly negative values relative to the Y-axis indicate greater underestimation of BrAC relative to the police-grade device

information and the intent to market the device. The FDA does not require "approval" (obtaining 510[k] clearance based on review of data on accuracy) to market alcohol breath testing devices but should reconsider this in light of our findings. Furthermore, we believe that the Federal Trade Commission should investigate personal smartphone breath testing device companies for deceptive claims of accuracy.

Smartphone-paired breath testing devices that are proven to be accurate are of value for remotely monitoring alcohol consumption, and if paired with smartphone-enabled behavioral interventions could reduce risky drinking behavior and promote abstinence (Gustafson et al., 2014; Luczak & Ramchandani, 2019; Quanbeck et al., 2014). For example, alcohol breath testing measurements prompted by text message and recorded by cellphone video with financial rewards for negative readings were found to decrease drinking frequency among heavy drinkers in a pilot randomized trial (Alessi & Petry, 2013). Newer smartphone apps designed specifically for remote monitoring have been used in several recent studies of interventions to curb hazardous drinking (Hämäläinen et al., 2018; Lauckner et al., 2019; Min et al., 2019). These apps have the advantage of transmitting timestamped and geocoded alcohol breath testing measurements readings automatically to researchers and clinicians. These readings can be verified with automatically captured pictures of the person's face providing the reading. Furthermore, one recent study showed that traditional self-report measures of alcohol consumption underestimate consumption compared with smartphone-prompted breath testing measurements using the BACTrack Mobile Pro (Kaplan & Koffarnus, 2019). Our findings validate the accuracy of the BACTrack Mobile Pro relative to self-report measures.

This study has limitations. The sample was relatively small and of convenience. The results are generalizable only up to doses of

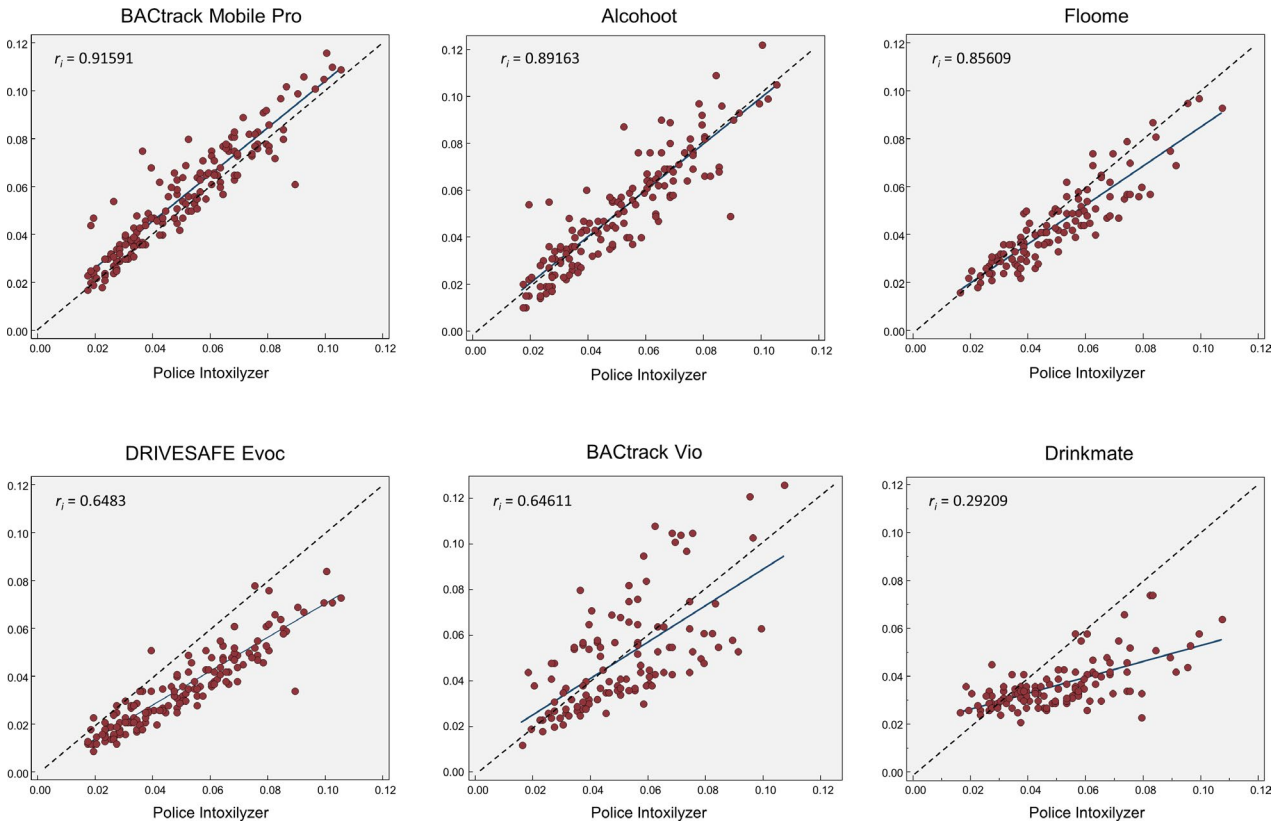


FIGURE 4 Scatter plots demonstrating within-person correlation between police-grade (Intoxilyzer 240) breath test Breath Alcohol Concentration (BrAC) and smartphone-paired breath testing device BrAC. Dotted line represents perfect correlation. Points below dotted line represent measurements in which the smartphone-paired breath testing device was lower than the police-grade device. r_i =intraclass correlation coefficient (1.00 is perfect)

alcohol that yield a BAC of 0.10% among individuals with similar demographic characteristics and BMI. There may also be differences in accuracy between breath testing devices and between devices and phlebotomy-measured blood alcohol levels at higher doses of alcohol. Finally, the study design did not include a formal test of the test–retest reliability of the devices. A strength of this study was that it tested devices against each other within participants and over time using standardized alcohol dosing and testing protocols. The protocol and study design minimized bias in device measurement that could result from participant-level differences. Furthermore, the order in which the devices were tested during each time interval was randomized using 4x4 Latin squares to minimize bias that could result from the order of testing. Finally, while this study on the comparative diagnostic accuracy of devices in the laboratory setting does not meet the formal definition of an NIH clinical trial, we registered the trial on ClinicalTrials.gov for the purposes of public reporting but did so after enrollment. The original study data are available on request from the investigators.

In summary, the accuracy of smartphone-paired devices varied widely in this laboratory study of 2 groups of 10 participants with peak BACs ranging from 0.06 to 0.14%. We identified some consumer-marketed, smartphone-paired alcohol breath testing devices that were accurate and already in use for novel behavioral interventions deployed in clinical and public health settings, such as the BACTrack Mobile Pro. However, we also identified some devices that are likely harmful to public health because they severely underestimated blood alcohol levels and thus potentially provide false reassurance about fitness to drive, such as the Drinkmate. More research is needed to determine how the use of alcohol breath devices

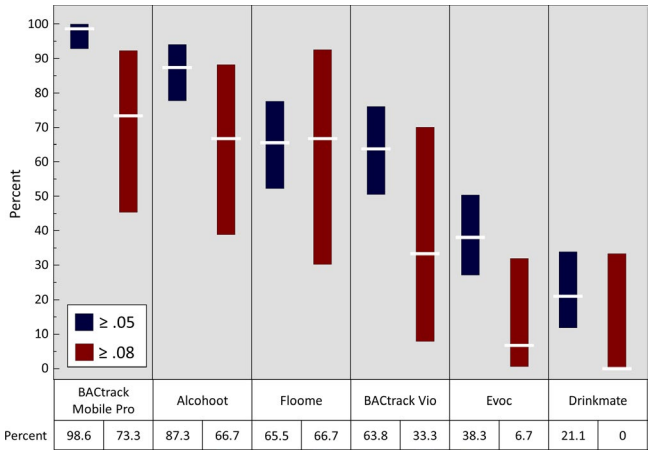


FIGURE 5 Proportion smartphone-paired breath tests that detected a threshold Breath Alcohol Concentration (BrAC) ≥0.05 or ≥0.08 as measured by the police-grade breath testing device (Intoxilyzer 240). White horizontal lines represent mean, and the length of the bars represents 95% confidence intervals

affects the decision to drive after drinking among the public and whether devices that have been verified to be accurate can be used to reduce alcohol-impaired driving.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of Kathryn Wanner, MSW and Kathryn Saulsgiver, PhD for their support in data collection and project management, and the laboratory infrastructure support from Nalaka Gooneratne, MD, MSc and the Perelman School of Medicine Clinical and Translational Research Center.

CONFLICT OF INTEREST

Dr. Kranzler is a member of an advisory board for Dicerna Pharmaceuticals and of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported in the last 3 years by AbbVie, Alkermes, Dicerna, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, Pfizer, Arbor, and Amygdala Neurosciences and is named as an inventor on PCT patent application #15/878,640 entitled: "Genotype-guided dosing of opioid agonists," filed January 24, 2018.

ORCID

Mucio Kit Delgado https://orcid.org/0000-0002-2248-8401
Brenda Curtis https://orcid.org/0000-0002-2511-3322
Henry R. Kranzler https://orcid.org/0000-0002-1018-0450

REFERENCES

Alessi, S.M. & Petry, N.M. (2013) A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. *Addiction*, 108, 900–909.

Alexandra Berzon, S.S. & Scheck, J. (2019) Amazon Has Ceded Control of Its Site. The Result: Thousands of Banned, Unsafe or Mislabeled Products. *The Wall Street Journal*, August 23, 2019. https://www.wsj.com/articles/amazon-has-ceded-control-of-its-site-the-result-thousands-of-banned-unsafe-or-mislabeled-products-11566564990

Ashdown, H.F., Fleming, S., Spencer, E.A., Thompson, M.J. & Stevens, R.J. (2014) Diagnostic accuracy study of three alcohol breathalysers marketed for sale to the public. *British Medical Journal Open*, 4, e005811.

Beirness, D.J. (1987) Self-estimates of blood alcohol concentration in drinking-driving context. *Drug and Alcohol Dependence*, 19, 79–90.

Beirness, D.J., Foss, R.D. & Voas, R.B. (1993) Drinking drivers' estimates of their own blood alcohol concentration. *Journal of Traffic Medicine*, 21, 73–78.

Chambers, L.W., Roberts, R.S. & Voelker, C.A. (1976) The epidemiology of traffic accidents and the effect of the 1969 breathalyser amendment in Canada. *Accident Analysis & Prevention*, 8, 201–206.

Colorado Department of Transportation. Alcohol and Impaired Driving: Buy a Breathalyzer. https://www.codot.gov/safety/alcohol-and-impaired-driving. Accessed: February 1, 2020.

Department of Transporation, National Highway Transporation Safety Administration, Highway Safety Programs. (2017) Conforming Products List of Evidential Breath Alcohol Measurement Devices. November 2, 2017. https://www.govinfo.gov/content/pkg/FR-2017-11-02/pdf/2017-23869.pdf.

FTC (2017) *"Breathometer" Marketers Settle FTC Charges of Misrepresenting Ability to Accurately Measure Users' Blood Alcohol Content*. Washington D.C.: FTC. https://www.ftc.gov/news-events/press-releases/2017/01/breathometer-marketers-settle-ftc-charges-misrepresenting-ability

Gomez, A., Conde, A., Santana, J. & Jorrin, A. (2005) Diagnostic usefulness of brief versions of Alcohol Use Disorders Identification Test (AUDIT) for detecting hazardous drinkers in primary care settings. *Journal of Studies on Alcohol*, 66, 305–308.

Gornall, J. (2014) Personal breathalysers may give false reassurance to drivers, research shows. *BMJ*, 349, g7745.

Gustafson, D.H., McTavish, F.M., Chih, M.-Y., Atwood, A.K., Johnson, R.A., Boyle, M.G. et al. (2014) A smartphone application to support recovery from alcoholism: a randomized clinical trial. *JAMA Psychiatry*, 71, 566–572.

Hämäläinen, M.D., Zetterström, A., Winkvist, M., Söderquist, M., Karlberg, E., Öhagen, P. et al. (2018) Real-time monitoring using a breathalyzer-based eHealth system can identify lapse/relapse patterns in alcohol use disorder patients. *Alcohol and Alcoholism*, 53, 368–375.

Jolly, J. (2015). Turning Your Smartphone into a Breathalyzer. The New York Times. Available at: https://well.blogs.nytimes.com/2015/12/21/turning-your-smartphone-into-a-breathalyzer/. Accessed: January 27, 2017

Jones, A.W. & Andersson, L. (2003) Comparison of ethanol concentrations in venous blood and end-expired breath during a controlled drinking study. *Forensic Science International*, 132, 18–25.

Kaplan, B.A. & Koffarnus, M.N. (2019) Timeline followback self-reports underestimate alcohol use prior to successful contingency management treatment. *Alcohol and Alcoholism*, 54, 258–263.

Koo, T.K. & Li, M.Y. (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163.

Kriikku, P., Wilhelm, L., Jenckel, S., Rintatalo, J., Hurme, J., Kramer, J. et al. (2014) Comparison of breath-alcohol screening test results with venous blood alcohol concentration in suspected drunken drivers. *Forensic Science International*, 239, 57–61.

Lauckner, C., Taylor, E., Patel, D. & Whitmire, A. (2019) The feasibility of using smartphones and mobile breathalyzers to monitor alcohol consumption among people living with HIV/AIDS. *Addiction Science & Clinical Practice*, 14, 43.

Légifrance. *Légifrance: Le Service Public pour la Diffusion du Droit: Décret n° 2012-284 du 28 février 2012 relatif à la possession obligatoire d'un éthylotest par le conducteur d'un véhicule terrestre à moteur*. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000025417826?r=HH-vIMdCdFY (Accessed March 29, 2021).

Luczak, S.E. & Ramchandani, V.A. (2019) Special issue on alcohol biosensors: development, use, and state of the field: Summary, conclusions, and future directions. *Alcohol*, 81, 161.

Martin, R.J., Chaney, B.H., Cremeens-Matthews, J. & Vail-Smith, K. (2016) Perceptions of breath alcohol concentration (BrAC) levels among a sample of bar patrons with BrAC values of 0.08% or higher. *Psychology of Addictive Behaviors*, 30, 680.

Min, A., Lee, D., Gao, G., Jeong, S. & Shih, P.C. (2019) Design and assessment of a personal breathalyzer intervention to support responsible drinking. *International Journal of Human-Computer Studies*, 102382.

National Academies of Sciences, Engineering, and Medicine (2018) *Getting to zero alcohol-impaired driving fatalities: A comprehensive approach to a persistent problem*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24951

National Highway Traffic Safety Administration (2017) *2016 fatal motor vehicle crashes: Overview. Traffic Safety Facts*. Report no. DOT HS 812 456. Washington, DC: National Highway Traffic Safety Administration.

Quanbeck, A., Chih, M.-Y., Isham, A., Johnson, R. & Gustafson, D. (2014) Mobile delivery of treatment for alcohol use disorders: a review of the literature. *Alcohol Research: Current Reviews*, 36, 111.

ALCOHOLISM CLINICAL & EXPERIMENTAL RESEARCH

Schechtman, E. & Shinar, D. (2011) An analysis of alcohol breath tests results with portable and desktop breath testers as surrogates of blood alcohol levels. *Accident Analysis & Prevention*, 43, 2188–2194.

van Tassel, W., Dennis, M. & Parker, M. (2004) *Pocket model, numerical readout breath alcohol measurement devices: A laboratory-and in-vivo based evaluation*. Proceedings of the 17 th International Conference on Alcohol, Drugs and Traffic Safety. Glasgow, Scotland, 2004.

Thombs, D.L., Olds, R.S. & Snyder, B.M. (2003) Field assessment of BAC data to study late-night college drinking. *Journal of Studies on Alcohol*, 64, 322–330.

Tofighi, B., Chemi, C., Ruiz-Valcarcel, J., Hein, P. & Hu, L. (2019) Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. *JMIR mHealth and uHealth*, 7, e11831.

Zaloshnja, E., Miller, T.R. & Blincoe, L.J. (2013). Costs of Alcohol-Involved Crashes, United States, 2010. *Association for the Advancement of Automotive Medicine*, 57, 3–12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3861831/

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Delgado MK, Shofer F, Wetherill R, et al. Accuracy of Consumer-marketed smartphone-paired alcohol breath testing devices: A laboratory validation study. *Alcohol Clin Exp Res*. 2021;45:1091–1099. https://doi.org/10.1111/acer.14597