A Preliminary Study Using Smartphone Accelerometers to Sense Gait Impairments Due to Alcohol Intoxication

BRIAN SUFFOLETTO, M.D.,[a,*] PRITIKA DASGUPTA, M.P.H., M.H.I., M.S.,[b] RAY UYMATIAO, B.S.,[a] JAMES HUBER, B.S.,[a] KATE FLICKINGER, M.S.,[a] & ERVIN SEJDIC, PH.D.[c]

[a]*Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania*
[b]*Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania*
[c]*Department of Electrical and Computer Engineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania*

ABSTRACT. Objective: Sensing the effects of alcohol consumption in real time could offer numerous opportunities to reduce related harms. This study sought to explore accuracy of gait-related features measured by smartphone accelerometer sensors on detecting alcohol intoxication (breath alcohol concentration [BrAC] > .08%). **Method:** In a controlled laboratory study, participants ($N = 17$; 12 male) were asked to walk 10 steps in a straight line, turn, and walk 10 steps back before drinking and each hour, for up to 7 hours after drinking a weight-based dose of alcohol to reach a BrAC of .20%. Smartphones were placed on the lumbar region and 3-axis accelerometer data was recorded at a rate of 100 Hz. Accelerometer data were segmented into task segments (i.e., walk forward, walk backward). Features were generated for each overlapping 1-second windows, and the data set was split into training and testing data sets. Logistic regression models were used to estimate accuracy for classifying BrAC ≤ .08% from BrAC > .08% for each subject. **Results:** Across participants, BrAC > .08% was predicted with a mean accuracy of 92.5% using logistic regression, an improvement from a naive model accuracy of 88.2% (mean sensitivity = .89; specificity = .92; positive predictive value = .77; and negative predictive value = .97). The two most informative accelerometer features were mean signal amplitude and variance of the signal in the *x*-axis (i.e., gait sway). **Conclusions:** We found preliminary evidence supporting use of gait-related features measured by smartphone accelerometer sensors to detect alcohol intoxication. Future research should determine whether these findings replicate in situ. (*J. Stud. Alcohol Drugs, 81,* 505–510, 2020)

SENSING ALCOHOL INTOXICATION in real time could offer opportunities for triggering just-in-time interventions aimed at improving prevention and treatment of alcohol use disorders. For individuals in treatment, it could trigger immediate remote support from a sponsor, which could reduce relapse risk. For an individual with heavy drinking, it could trigger just-in-time resources aimed at reducing further alcohol consumption or other related risks like drinking and driving, potentially mitigating individual and public health harms.

Several methods exist for remote real-time monitoring of alcohol consumption. Transdermal alcohol monitoring provides approximation of blood alcohol concentration (Marques & McKnight, 2009). Portable breath analysis of exhaled alcohol metabolites is also commercially available. Barriers to using these methods include costs of device purchase and stigma associated with use in public. Even if these barriers are overcome, differences in physiology across individuals, especially those with routine alcohol consumption, result in varying physiological effects (and thus associated risks) for a given amount of alcohol consumed.

Received: October 21, 2019. Revision: March 25, 2020.

This work was supported by National Institutes of Health Grants K23 AA023284 and R01 AA023650 and National Library of Medicine Grant 4T15LM007059-30.

*Brian Suffoletto is now with the Department of Emergency Medicine, Stanford University, Palo Alto, CA. Correspondence may be sent to him via email at: suffbp@stanford.edu.

Measuring altered human function due to alcohol consumption may provide a more useful method for longitudinal monitoring. Alcohol, especially at levels greater than .08%, produces psychomotor changes, manifested primarily through impaired speech and gross- and fine-motor function. One measure of psychomotor performance that is particularly sensitive to alcohol is gait, which requires coordination of multiple sensory and motor systems (Jansen et al., 1985; Nieschalk et al., 1999).

Smartphones could offer a convenient and scalable way to measure gait features in the real world. More than 96% of Americans own a smartphone (Pew, 2019), which are almost universally embedded with sensors that allow for inertial measurements of gait. Researchers have begun to model the associations between gait abnormalities detected using smartphone sensors and either real or simulated alcohol consumption (Aiello & Agu, 2016; Arnold et al., 2015). Our group has shown that gait-related phone sensor features correlate strongly with estimated alcohol concentrations when drinks are self-reported (Suffoletto et al., 2018). In this pilot study, we sought to determine accuracy of gait-related features measured by smartphone accelerometer sensors on detecting an objective measure of alcohol intoxication (breath alcohol concentration [BrAC] > .08%). We hypothesized that gait would show evidence of instability when BrAC > .08%. Results from this study are critical to building an evidence base for smartphone-based digital interventions that deliver just-in-time support to reduce risks associated with excessive alcohol consumption.

Method

Participants

From August to December 2018, we recruited 22 adults for a controlled laboratory study. Participants were recruited via word of mouth and locally posted advertisements for a study to examine the effects of alcohol on psychomotor tasks. We conducted an initial screen by telephone to ensure they were at least 21 years old and consumed alcohol at least once per week. Consented participants then made appointments to come to the laboratory for one session that would last at least 7 hours and were instructed to abstain from consuming alcohol or using other psychoactive drugs during the 24 hours preceding the session. They were also told to fast and refrain from caffeine consumption at least 4 hours before the session. On the day of the session, participants were screened in person to verify age at least 21 years using their driver license and a brief health survey. Individuals who reported any positive responses on the CAGE questionnaire (Bush et al., 1987), hepatic/renal impairment, or peptic ulcer disease were excluded. Urine samples were also tested for pregnancy in female participants. Women who were pregnant or breastfeeding were excluded.

Procedures

Participants presented to the Department of Emergency Medicine Applied Physiology Lab at the University of Pittsburgh at 8 A.M. After providing informed consent, participants completed a questionnaire including the 10-question Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993). Body weight and height were measured, and an intravenous line was placed to draw blood alcohol measurements and to administer nausea medicine (ondansetron 4 mg).

Alcohol. Investigators prepared an ethanol oral dosing to achieve a goal peak BrAC .20% using the Widemark formula as follows: 2 g/L × (0.7 L/kg [for men] or 0.6 L/kg [for women] × participant weight in kilograms) = dose of ethanol in grams/0.3156 g ethanol per milliliter = milliliter distilled spirits. Vodka was mixed with lime juice and simple syrup and administered according to standard procedures (Fillmore et al., 2000). Participants were given a maximum of 1 hour to finish alcohol consumption. At baseline and each half-hour hour (for up to 7 hours), we measured BrAC (BACtrac s80 Pro; KHN Solutions, Inc., San Francisco, CA). At baseline and at 2, 4, and 6 hours, we measured blood alcohol concentration. Participants left the lab after 7 hours, when they could ambulate safely and had someone to drive them home.

Walking trial. Participants completed a walking trial at baseline and each hour for up to 7 hours following alcohol administration. Before beginning the walking trial, we placed a smartphone on the lower back using an elastic belt. We then instructed the participant to walk 10 steps in a straight line on a flat, carpeted but noncompliant surface, turn around and walk 10 steps back to the beginning spot. When they indicated that they were ready, we started recording accelerometer data from the phyphox app (www.phyphox.org). When the participant completed the walking trial, the phone was removed from the belt and data downloaded to a secure file.

Measures

Alcohol intoxication. We chose to use a threshold of BrAC > .08% as our classifier of alcohol intoxication because it has been used in prior studies of acute alcohol effects on psychomotor performance (Peacock et al., 2015) and represents the legal limit of blood alcohol in adult drivers in the United States.

Gait feature extraction. Smartphone sensors and app captured linear accelerations (in units of ms^2) at a frequency of 100 Hz from the x, y, and z directions, which correspond to the mediolateral, vertical, and anteroposterior directions. We first labeled accelerometer time-series data into the following segments: walk forward, turn, walk back. Accelerometry data for each segment were further segmented into 1-second windows with a 50% overlap consistent with prior machine learning studies (Mannini et al., 2010; Preece et al., 2009). Features were chosen based on prior research (Dasgupta et al., 2018; Sejdic et al., 2014) and generated for each window (feature selection shown in Table 1).

Statistical analyses

To ensure validity of BrAC values, we compared them with BAC values using correlation coefficients. We first generated a population-based model using leave-one-out methods for detecting BrAC > .08%, which showed poor discrimination. We then chose to generate and test model accuracy for each participant separately. First, we used correlation matrix to identify highly correlated feature pairs (r > .75) and removed features with the highest mean absolute correlation. Each data set was split into a "training" and "testing" data set using an 80/20% split. Logistic regression (LR) models were trained using a repeated 10-fold cross validation, in which 10-fold cross validation was repeated three times. We calculated the range of accuracies across individuals (with 95% confidence intervals [CIs]), sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV). We compared the mean model accuracy with and without gait features using a two-sample t test. We identified accelerometer features with the highest information gain using variable importance function in the caret function in R. We explored association of individual characteristics (i.e., age, sex, AUDIT score) with model ac-

Table 1. Accelerometer gait features

Feature	Formula
Mean of acceleration signal (ML, AP, V)	μ_{ML}, μ_{AP}, μ_V
Variance of acceleration signal (ML, AP, V)	$\sigma^2_{ML}, \sigma^2_{AP}, \sigma^2_V$
Correlation of pairwise acceleration signals	*cor*(*ML,AP*), *cor*(*ML,V*), *cor*(*V,AP*)
Covariance of acceleration signal (ML, AP, V)	*cov*(*ML,AP*), *cov*(*ML,V*), *cov*(*V,AP*)
Maximum difference of acceleration signal (ML, AP, V)	d_{ML}, d_{AP}, d_V
Maximum difference of pairwise acceleration signals	$\sqrt{d^2_{ML}+d^2_{AP}}, \sqrt{d^2_{ML}+d^2_V}, \sqrt{d^2_V+d^2_{AP}}, \sqrt{d^2_V+d^2_{AP}+d^2_{ML}}$
Mean trend of acceleration signal (ML, AP, V) of 0.1 second windows within the window	$\mu T = \sum_{i=2}^{n}(\lvert\mu_i - \mu_{i-1}\rvert)$
Windowed mean trend of acceleration signal (ML, AP, V) of 0.1 second windows within the window	$\mu D = \sum_{i=2}^{n}(\lvert\mu - \mu_i\rvert)$
Variance trend of acceleration signal (ML, AP, V)	$\sigma^2 T = \sum_{i=2}^{n}(\lvert\sigma^2_i - \sigma^2_{i-1}\rvert)$
Windowed variance trend of acceleration signal (ML, AP, V)	$\sigma^2 D = \sum_{i=2}^{n}(\lvert\sigma^2 - \sigma^2_i\rvert)$

Notes: ML = mediolateral; AP = anteroposterior; V = vertical.

curacy using univariate regression models. Analyses were conducted using R Version 3.5.2 and Stata 15.0 (StataCorp LP, College Station, TX).

Results

Participants

Seventeen individuals (77%) provided at least one gait measurement when BrAC was greater than .08% and were included in the analysis. Mean age was 27.5 years (*SD* = 5.5), with a range of ages from 21 to 43 years. The majority (70.6%) of participants were men, and all participants were White and non-Hispanic. Mean AUDIT score was 3.5 (*SD* = 2.8), with four participants meeting criteria for risky drinking based on a score between 7 and 15. Mean weight was 76 kg (range: 51–102) and mean height 68 inches (range: 62–73).

BrAC and BAC values

The BrAC was confirmed at 0% at baseline and increased above .08% in all participants by 1 hour. The BrAC began to decline gradually starting at 2 hours 30 minutes. Correlation between BrAC and BAC values was high (r = .96). The changes in and relationships between BrAC and BAC over time by subject are plotted in Figure 1.

Model output for predicting BrAC based on accelerometry data during gait task

Table 2 shows the number of gait trials and model output by participant. Across 17 participants, BrAC greater than .08% was predicted with a mean accuracy of 92.5%, an improvement from a naive model accuracy of 88.2% (two-sample t-test $p < .0001$; mean sensitivity = .89; specificity = .92; PPV = .77; and NPV = .97). There were no significant differences in prediction accuracy based on using data from "walk forward" versus "walk back" segments. In Table 1, we show the number of 1-second windows of accelerometer data used for LR classification and the variability of predictive metrics for the "walk back" segment by participant. We could not identify any participant characteristics (e.g., age, sex, AUDIT score) associated with model accuracy.

Top accelerometry features for predicting breath alcohol concentration

As shown in Table 3, the two most informative acceler-

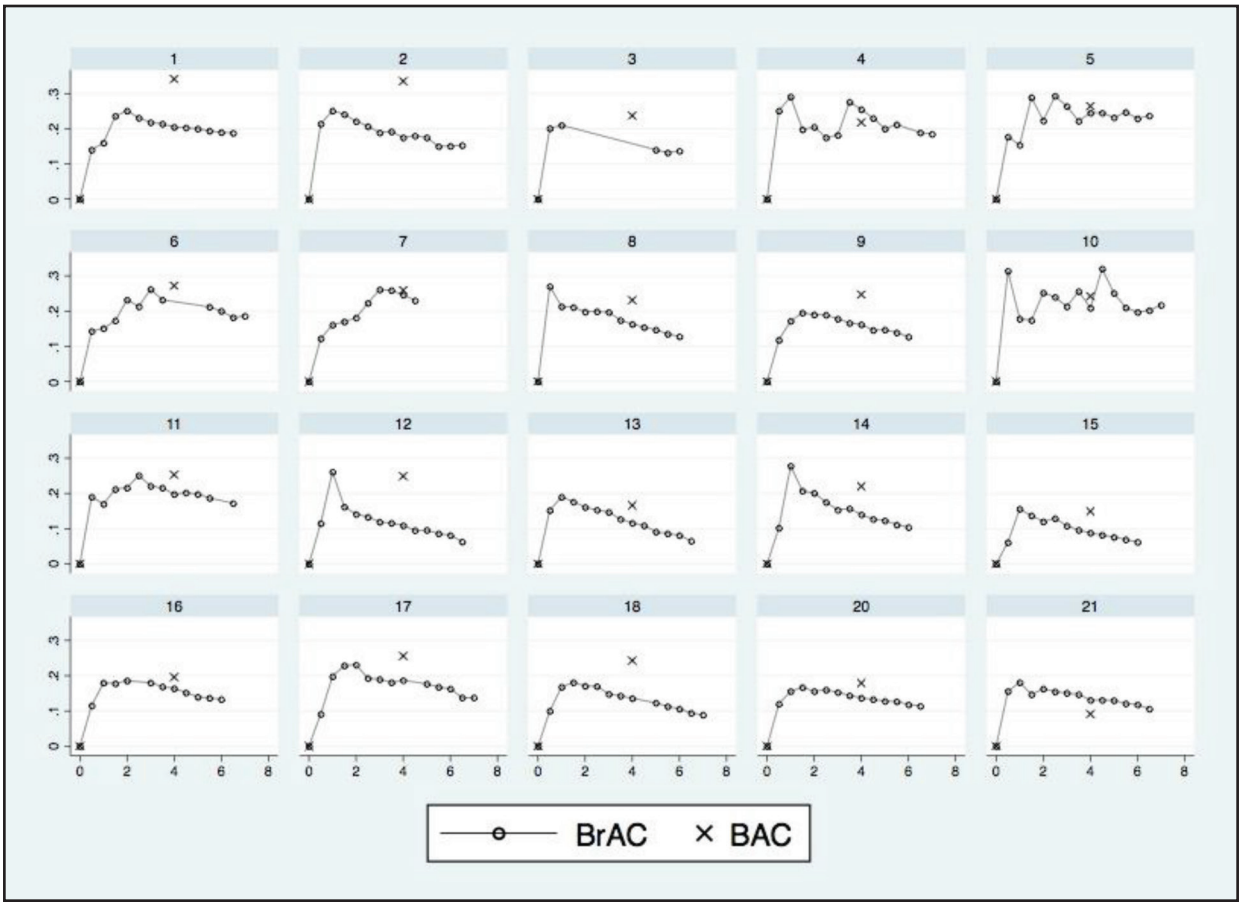

FIGURE 1. Plots of the breath alcohol concentration (BrAC) and blood alcohol concentration (BAC) values over time by subject

TABLE 2. Logistic regression models by participant

	Number of 1-second windows							
Participant	BrAC <.08	BrAC ≥.08	Accuracy	[95% CI]	Sens.	Spec.	PPV	NPV
1	9	79	.88	[.64, .99]	1.00	.87	.50	1.00
2	11	9	1.00	[.40, 1.00]	1.00	1.00	1.00	1.00
3	10	70	1.00	[.79, 1.00]	1.00	1.00	1.00	1.00
4	10	37	.78	[.40, .97]	.67	.83	.67	.83
5	10	20	1.00	[.54, 1.00]	1.00	1.00	1.00	1.00
6	17	25	.88	[.47, 1.00]	1.00	.83	.67	1.00
7	7	50	1.00	[.72, 1.00]	1.00	1.00	1.00	1.00
8	10	62	.93	[.66, 1.00]	1.00	.92	.50	1.00
9	10	58	1.00	[.75, 1.00]	1.00	1.00	1.00	1.00
10	10	80	1.00	[.81, 1.00]	1.00	1.00	1.00	1.00
11	9	45	.90	[.55, 1.00]	1.00	.88	.67	1.00
12	8	50	1.00	[.72, 1.00]	1.00	1.00	1.00	1.00
13	33	11	1.00	[.63, 1.00]	1.00	1.00	1.00	1.00
14	7	55	1.00	[.74, 1.00]	1.00	1.00	1.00	1.00
15	32	24	.82	[.48, .98]	1.00	.71	.67	1.00
16	9	49	.91	[.59, 1.00]	.50	1.00	1.00	.90
17	8	37	.78	[.40, .97]	1.00	.75	.33	1.00
Total	210	761	.93	[.61, .99]	.95	.93	.82	.98

Notes: BrAC = breath alcohol concentration (measured in %); CI = confidence interval; sens. = sensitivity; spec. = specificity; PPV = positive predictive value; NPV = negative predictive value.

Table 3. Top accelerometer features by participant

	Features			
Participant	1	2	3	4
1	mean_x	mean_z	variance_x	variance_y
2	mean_x	mean_y	mean_z	variance_x
3	mean_y	mean_z	variance_x	variance_y
4	mean_x	mean_y	variance_x	correlation_xy
5	mean_y	mean_z	variance_x	variance_z
6	mean_x	mean_y	mean_z	variance_y
7	mean_x	mean_y	mean_z	variance_x
8	mean_x	mean_y	mean_z	variance_x
9	mean_y	mean_z	variance_x	variance_y
10	mean_x	mean_y	mean_z	variance_y
11	mean_x	mean_y	mean_z	variance_x
12	mean_y	variance_x	variance_y	maxdiff_x
13	mean_x	mean_z	variance_x	variance_z
14	mean_y	mean_z	variance_x	variance_y
15	mean_x	mean_y	mean_z	variance_z
16	mean_x	mean_y	variance_y	correlation_yz
17	variance_x	variance_z	correlation_xy	correlation_yz

Notes: Mean = mean of acceleration signal; variance = variance of acceleration signal; correlation = correlation of pairwise acceleration signals; x = mediolateral; y = anteroposterior; z = vertical. Highlighted are the relative importance of the top two features overall.

ometer features were mean signal amplitude in the *x*-axis and variance of the signal in the *x*-axis.

Discussion

In this laboratory study, we found that smartphones can capture unique gait features that are sensitive to alcohol intoxication, classifying alcohol intoxication within individuals with an accuracy of around 90%. These findings extend prior published research examining the use of phone sensors to detect gait changes related to alcohol. Kao et al. (2012) recorded three-axis accelerometry data from three healthy volunteers during a gait task and found that there was larger step-time variance and longer gait stretch measured after alcohol consumption. Arnold et al. (2015) recorded three-axis accelerometry data from naturalistic gait samples from six healthy volunteers and was able to classify 0–2 drinks distinct from 3–6 drinks and more than 6 drinks with an accuracy of 56% in the training set and 70% in the validation set. Aiello and Agu (2016) simulated alcohol intoxication in 34 healthy volunteers and measured accelerometry signals during lab-based gait tasks, finding that they could classify simulated impairment with an accuracy of 89.45% when incorporating gyroscope to accelerometer features. To our knowledge, this is the first study that objectively measured circulating alcohol levels to train detection models.

One significant strength of this study is we found that high accuracy can be achieved using logistic regression models. This allowed us to examine the relative contribution of individual gait features in models (not directly possible using machine learning) where we found that amplitude and variance along the *x*-axis of the phone were key predictors. In this study, based on the position of the phones, the *x*-axis represents side-to-side sway during walking. This finding is consistent with prior research examining the effect of alcohol on balance (Fiorentino, 2018; Marczinski & Mearns, 2020). Another study strength is that we used very brief walking samples (i.e., 10 steps). This suggests that it would be feasible to collect this type of sample in naturalistic settings to deliver just-in-time support. A third strength is the use of objective alcohol concentration (i.e., breath alcohol) to classify legal intoxication instead of drink amount, which can be subject to variability due to reporting biases and will not always accurately represent blood alcohol concentration.

This study's findings are limited by the relatively small sample size, the use of a cohort that largely drinks below risky levels, and controlled setting of data measurement. Given the limited number of data points at which gait tasks were completed with a BrAC of .08% or below, we did not examine whether gait-related features discriminate lower levels of drinking. Our procedures allowed for individuals to drink alcohol over 1 hour; however, there was variability even within this limited period, which likely affected variability in peak BrAC. In addition, we were not able to examine the difference in prediction during the ascending versus descending limb of intoxication. Another limitation to consider is that we placed the smartphone on the lower back, which may not represent where individuals keep their phones in natural environments. We plan to examine how models differ when phones are carried in hand, in the front pocket, or in the side pocket. Last, we did not find that population-based models were accurate in predicting intoxication. We believe that this is because of the variability between individuals in gait patterns and suggests that any application would either need to collect individual gait measures during sober and drinking periods or incorporate some normalization procedures as performed by Arnold et al. (2015).

Despite these limitations, this proof-of-concept study provides a foundation for future research on using smartphones to remotely detect alcohol-related impairments. Current tools to measure alcohol consumption and/or impairment remotely either require the purchase of additional hardware (e.g., SCRAM, breath alcohol analyzers) or the burden of manual recording of alcohol consumption. A mobile application could be built to sense periods of walking (using Google API: "on foot" classification), measure accelerometer signals, and when sway patterns are recognized, trigger either just-in-time support or use further techniques to further improve classification.

References

Aiello, C., & Agu, E. (2016). Investigating postural sway features, normalization and personalization in detecting blood alcohol levels of smartphone users. In *2016 IEEE Wireless Health* (pp. 73–80). Retrieved from https://ieeexplore.ieee.org/document/7764559

Arnold, Z., LaRose, D., & Agu, E. (2015). Smartphone inference of alcohol consumption levels from gait. In *2015 International Conference on Healthcare Informatics* (pp. 417–426). Retrieved from https://web.cs.wpi.edu/~emmanuel/publications/PDFs/C17.pdf

Bush, B., Shaw, S., Cleary, P., Delbanco, T. L., & Aronson, M. D. (1987). Screening for alcohol abuse using the CAGE questionnaire. *American Journal of Medicine, 82,* 231–235. doi:10.1016/0002-9343(87)90061-1

Dasgupta, P., VanSwearingen, J., & Sejdic, E. (2018). "You can tell by the way I use my walk." Predicting the presence of cognitive load with gait measurements. *Biomedical Engineering Online, 17,* 122. doi:10.1186/s12938-018-0555-8

Fillmore, M. T., Dixon, M. J., & Schweizer, T. A. (2000). Differential effects of alcohol on responses to negatively and positively primed stimuli. *Journal of Studies on Alcohol, 61,* 872–880. https://doi.org/10.15288/jsa.2000.61.872

Fiorentino, D. D. (2018). The effects of breath alcohol concentration on postural control. *Traffic Injury Prevention, 19,* 352–357. doi:10.1080/15389588.2018.1423561

Jansen, E. C., Thyssen, H. H., & Brynskov, J. (1985). Gait analysis after intake of increasing amounts of alcohol. *International Journal of Legal Medicine, 94,* 103–107. doi:10.1007/BF00198678

Kao, H. L. C., Ho, B. J., Lin, A. C., & Chu, H. H. (2012). *Phone-based gait analysis to detect alcohol usage.* Retrieved from https://dl.acm.org/doi/pdf/10.1145/2370216.2370354

Mannini, A., & Sabatini, A. M. (2010). Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors (Basel), 10,* 1154–1175. doi:10.3390/s100201154

Marczinski, C. A., & Mearns, C. L. (2020). Automated assessment of alcohol-induced impairment of balance in male and female social drinkers. *Experimental and Clinical Psychopharmacology, 28,* 131–136. doi:10.1037/pha0000309

Marques, P. R., & McKnight, A. S. (2009). Field and laboratory alcohol detection with 2 types of transdermal devices. *Alcoholism: Clinical and Experimental Research, 33,* 703–711. doi:10.1111/j.1530-0277.2008.00887.x

Nieschalk, M., Ortmann, C., West, A., Schmäl, F., Stoll, W., & Fechner, G. (1999). Effects of alcohol on body-sway patterns in human subjects. *International Journal of Legal Medicine, 112,* 253–260. doi:10.1007/s004140050245

Peacock, A., Cash, C., & Bruno, R. (2015). Cognitive impairment following consumption of alcohol with and without energy drinks. *Alcoholism: Clinical and Experimental Research, 39,* 733–742. https://doi.org/10.1111/acer.12680

Pew Research Center. (2019, June 12). *Mobile fact sheet.* Retrieved from https://www.pewresearch.org/internet/fact-sheet/mobile/

Preece, S. J., Goulermas, J. Y., Kenney, L. P., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological Measurement, 30,* R1–R33. doi:10.1088/0967-3334/30/4/R01

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction, 88,* 791–804. doi:10.1111/j.1360-0443.1993.tb02093.x

Sejdic, E., Lowry, K. A., Bellanca, J., Redfern, M. S., & Brach, J. S. (2014). A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22,* 603–612. doi:10.1109/TNSRE.2013.2265887

Suffoletto, B., Gharani, P., Chung, T., & Karimi, H. (2018). Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment. *Gait & Posture, 60,* 116–121. doi:10.1016/j.gaitpost.2017.11.019