Original Investigation

Mobile, Remote, and Individual Focused: Comparing Breath Carbon Monoxide Readings and Abstinence Between Smartphone-Enabled and Stand-Alone Monitors

Breanna M. Tuck BA¹, Joshua L. Karelitz PhD^{2,3,*}, Rachel L. Tomko PhD^{4,*}, Jennifer Dahne PhD^{4,5}, Patrick Cato BS⁴, Erin A. McClure PhD^{4,5,*}

¹Steve Hick School of Social Work, University of Texas at Austin, Austin, TX; ²Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA; ³Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA; ⁴Department of Psychiatry and Behavioral Sciences, College of Medicine, Medical University of South Carolina, Charleston, SC; ⁵Hollings Cancer Center, Medical University of South Carolina, Charleston, SC

Corresponding Author: Erin A. McClure, PhD, Medical University of South Carolina, 67 President St, MSC 861, Charleston, SC 29425, USA. Telephone: 843-792-7192; Fax: 843-792-8206; E-mail: mccluree@musc.edu

Abstract

Introduction: Newly available, smartphone-enabled carbon monoxide (CO) monitors are lower in cost than traditional stand-alone monitors and represent a marked advancement for smoking research. New products are promising, but data are needed to compare breath CO readings between smartphone-enabled and stand-alone monitors. The purpose of this study was to (1) determine the agreement between the mobile iCO (Bedfont Scientific Ltd) with two other monitors from the same manufacturer (Micro⁺ pro and Micro⁺ basic) and (2) determine optimal, monitor-specific, cotinine-confirmed abstinence cutoff values.

Methods: Adult (≥ 18) smokers ($n = 26$) and nonsmokers ($n = 21$) provided three breath CO samples (using three different monitors) in each of 10 sessions, and urine cotinine was measured for gold standard determination of abstinence. CO comparisons ($N = 437$) were analyzed using regression-based Bland–Altman Analysis of Agreement; receiver operating characteristics curves were used to determine optimal abstinence cutoffs.

Results: Bland–Altman analyses indicated that the iCO monitor provided higher CO results than both Micro⁺ monitors. Sensitivity and specificity analyses showed that the optimal CO cutoff for determining abstinence was <3 ppm for the Micro⁺ pro (88% sensitivity, 93% specificity) and Micro⁺ basic (83% sensitivity, 98% specificity), but was higher for the iCO (<6 ppm; 73% sensitivity, 100% specificity).

Conclusions: Relative to both Micro⁺ monitors, the smartphone-enabled iCO provided systematically higher CO values and required a higher cutoff to reliably determine smoking abstinence. This does not indicate that CO values obtained using the iCO are not valid; instead, these results suggest that monitor-specific abstinence cutoffs are needed to ensure accurate bioverification of smoking status.

Implications: Results from this study indicate that CO values from the smartphone-enabled iCO should not be used interchangeably with the stand-alone Micro⁺ pro and Micro⁺ basic, particularly when lower CO values (<10 ppm) are critical (ie, determination of abstinence vs confirming smoking status for study inclusion). Optimal CO cutoffs recommended for determining abstinence on Micro⁺ and iCO monitors are at <3 and <6 ppm, respectively.

Introduction

Breath carbon monoxide (CO) monitoring is an objective, biochemical measurement that accurately reflects an individual's smoking status.^{1,2} Breath CO is frequently used as an outcome measure in smoking cessation intervention studies (abstinent vs not abstinent) and as a way to confirm self-reported smoking. The collection of breath CO as a biochemical marker of smoking is appealing as it is noninvasive and yields reliable readings instantly when compared with collecting other biological materials.³ It is also a cost-effective option compared with more expensive laboratory assays. Most commercially available CO monitors cost approximately \$700–1200 USD, but can be used across multiple participants or patients and, with proper calibration, can be used for years.

Though breath CO has merit, there are challenges to its use. The half-life of CO is around 2–8 hours (depending on activity level),^{1,2} making multiple breath samples per day across consecutive days necessary to accurately characterize smoking and confirm abstinence. This may not be feasible for studies or clinics that only conduct in-person CO tests. Some studies have implemented CO collection through computer and mobile devices, but this requires that stand-alone CO monitors be sent home with study participants and are often paired with video verification, which may require separate hardware or software.^{4–6} This is also problematic because monitors may be damaged or not returned. Additionally, the size and mobility of monitors is a barrier to remote CO collection. A recent study attempted to circumvent issues of cost and the need for in-person visits by mailing personal CO monitors to study participants.⁷ That study found a low rate of CO monitor use and sample submission, which was due, in part, to practical considerations, such as the necessity to download software (via computer) and issues in receiving packages.

Although there are challenges introduced with any method of biochemical verification of smoking in research studies, the added rigor of such confirmation is critical and particularly important for remote studies. Therefore, further investigation and integration of novel and innovative biochemical verification of smoking is warranted. To improve the feasibility of remote CO collection specifically, Bedfont Scientific Ltd now markets the iCO, which is an inexpensive (~\$60 USD), smartphone-enabled breath CO monitor designed for individual sample collection (no mouthpieces or accessories for proper use or calibration) for a limited number of tests (200 breath tests or 3 years). As more emphasis is being placed on remote monitoring and mobile health strategies in both research and clinical settings,^{8,9} this device allows for frequent data collection while minimizing or eliminating the need for in-person visits. However, previous work has shown that CO readings across monitors are not necessarily consistent and it may not be appropriate to use values interchangeably. Indeed, one study found that comparisons between two CO monitors from different companies showed variations in CO readings across monitors.¹⁰ One recent report compared the iCO with the piCO⁺ Smokerlyzer (also manufactured by Bedfont Scientific Ltd) and found that values on the two monitors were highly correlated, but the iCO produced lower values.¹¹ However, that study did not determine agreement across the full continuum of CO values (ie, did not assess raw CO values) and also did not obtain samples from nonsmokers or light smokers to determine optimal smoking abstinence cutoff values. Additionally, that study used custom mouthpieces to allow the use of a single monitor across participants and alterations to the device may have affected CO readings. Furthermore, recommendations for appropriate abstinence cutoffs for breath CO have varied in the literature and included values from 3–4^{12,13} to

5–6,² to 7–10 parts per million (ppm).^{1,14,15} The cutoff for abstinence may depend largely on the type of CO monitor being used. Therefore, the aims of this study were to (1) assess the level of agreement in CO readings between the iCO monitor (smartphone-enabled device) with two frequently used, commercially available, traditional, stand-alone monitors from Bedfont Scientific Ltd (the Micro⁺ basic [previously marketed as the piCO Smokerlyzer] and the Micro⁺ pro) and (2) determine optimal, monitor-specific, quantitative urinary cotinine-confirmed cutoff CO values to indicate abstinence from smoking for all three breath CO monitors.

Methods

Participants

Adult (ages ≥18) cigarette smokers and nonsmokers were recruited from the community in Charleston, SC (US) from May 2019 through February 2020. Participants were considered smokers if they had smoked combustible cigarettes at least weekly over the past month (ie, at least some smoking every week during the past 4 weeks). Otherwise, participants were considered nonsmokers. Self-reported smoking status was verified by breath CO and instant-read urinary cotinine tests at screening. Participants were excluded if they had used any nicotine replacement therapy products in the 2 weeks prior to screening, if they had used electronic cigarettes or other tobacco products on 15 or more days out of the last 30, if they had used cannabis on 20 or more days out of the last 30, or had a positive cannabis urine drug screen at screening. Regular/frequent cannabis or other nicotine/tobacco use in the past 30 days was exclusionary as these could potentially affect the concordance between breath CO and cotinine measures. The past 30 days was used to assess regularity and recency of use of these products for exclusionary purposes. Co-use exclusion criteria were crafted to reduce the likelihood of product use that would affect the agreement between cotinine and CO values at the beginning and end of the study, while still allowing for some amount of infrequent use. Participants were also excluded if their medical and psychiatric health was poor, as determined by elevated scores on the Patient Health Questionnaire (PHQ-4),¹⁶ and a brief quality of life assessment.¹⁷ All study procedures were approved by the Institutional Review Board at the Medical University of South Carolina (MUSC).

Procedures

Following screening and determination of eligibility, participants attended 10 study visits over a 3-week period. During each visit, participants provided three CO samples, one from each monitor—iCO, Micro⁺ pro, and Micro⁺ basic. Breath samples were separated by approximately 5 minutes (no less than 3 minutes and no more than 10 minutes apart). Monitor order was counterbalanced across participants and study visits. All CO samples were collected with a trained research staff present. Participants were instructed to take a deep breath, hold for 15 seconds, and then exhale slowly into the monitor, attempting to exhale for 15–30 seconds and completely emptying their lungs. This process was also guided by instructions presented through the smartphone app for the iCO.

At Visit 1 (Day 1), each participant was assigned an iCO monitor, which was not altered (no custom mouthpieces used or modifications made to the device) and was only used by that participant throughout the study. The iCO connects to a mobile phone through the headphone jack and works with the Smokerlyzer app

available on iOS and Android devices. All CO sample instructions and prompts are displayed on the mobile phone. The resulting CO value that is visible is a range, but results can be emailed to any recipient, and that email provides a single value. CO values were recorded in real-time by research staff and results were emailed to research staff for verification and documentation. To create consistency between iCO usage, three study iPhones (iPhone 5C, operating system 10.3.3; Apple, Cupertino, CA) were used to conduct the CO tests. Five participants enrolled early in the study used their personal mobile phones (four iPhones and one Android device), prior to using the same study phones for CO collection. There were a total of three Micro⁺ pro and Micro⁺ basic monitors used for this study exclusively. All monitors were tested and calibrated (using a 20 ppm CO canister) regularly following manufacturer recommendations.

Measures

For all participants, demographics, tobacco/nicotine history, and current tobacco/nicotine use were assessed, in addition to general health and functioning (brief health questionnaire, the PHQ-4¹⁶ and a brief quality of life assessment¹⁷). All participants were asked about past 30-day cigarette, electronic cigarette, nicotine replacement therapy, and cannabis use via Timeline Follow-back procedures.¹⁸ For current smokers, nicotine dependence was assessed with the Fagerström Test for Nicotine Dependence (FTND).¹⁹

Exhaled breath CO (using the Micro⁺ pro) was collected at screening to confirm smoking versus nonsmoking status for participants (in addition to urinary cotinine). Participants also provided a urine sample at screening for the instant-read analysis of cotinine (a metabolite of nicotine) and a urine drug screen (Δ^9 -tetrahydrocannabinol detection only to confirm eligibility). Participants also provided urine samples on Days 1 and 10 for the quantitative testing of cotinine (ng/mL).

Statistical Analyses

Study data were managed using REDCap²⁰ hosted by the South Carolina Clinical and Translational Research Institute at MUSC. All analyses were conducted using IBM SPSS version 26. Differences in participant demographics between smoking groups were assessed using *t*-tests and chi-square tests. Preliminary CO comparisons used generalized estimating equations (GEE),²¹ with an independent working correlation matrix and identity link to analyze the difference in CO readings between monitor pairs (ie, iCO – Micro⁺ pro; iCO – Micro⁺ basic; Micro⁺ basic – Micro⁺ pro). Additional preliminary GEEs included time between samples (minutes) as a time-varying covariate (ie, the amount of time between monitors differed across visits due to monitor order randomization) to test for differences between monitors as a function of time elapsing between data collection.

Aim 1: Analysis of Agreement

Consistent with a previous publication comparing CO monitors,¹⁰ regression-based Bland–Altman analysis of agreement²² was used to determine the level of agreement between monitors to address Aim 1. Full details for this analysis have been described elsewhere.^{23,24} In short, this method plots the mean difference (ie, bias) between monitors (iCO – Micro⁺ pro) against the mean value between monitors ((iCO + Micro⁺ pro)/2), for each pair of measurements. The mean value between monitors is used rather than measures from one device or another because the use of a single “gold standard” measure

would increase the risk for Type I error.²⁵ The mean bias is compared against an a priori threshold of clinical significance ($\pm 20\%$ of the mean CO reading)¹⁰ to determine whether the monitors can be used interchangeably. Also computed were the 95% limits of agreement (LoA), which act as 95% confidence intervals (CIs) for the mean bias. When narrow, the LoA band indicates good agreement between methods, whereas a wide LoA band indicates poor agreement. To account for the within-subject correlation due to repeated measures, GEEs with an AR1 working correlation matrix and identity link were used in place of traditional regression to compute the mean bias and 95% LoA.^{26,27}

Aim 2: Optimal Abstinence Cutoffs

Receiver operating characteristic curves—plotting sensitivity by specificity—were used to determine the optimal abstinence cutoffs for each monitor verified by urinary cotinine (≤ 50 ng/mL).² Sensitivity is the percentage of CO values above a designated CO value when cotinine was >50 ng/mL (ie, the proportion of smokers correctly identified as smokers); specificity is the percentage of CO values below the abstinence criterion when cotinine was ≤ 50 ng/mL (ie, the proportion of nonsmokers or abstainers correctly identified as nonsmokers/abstainers). Sensitivity relates to the detection of smoking and specificity is verification of urinary cotinine status. CO monitor-specific optimal abstinence cutoffs were identified as the CO value, which maximized both sensitivity and specificity. We also assessed the area under the curve (AUC) for each monitor to compare levels of diagnostic accuracy. Ranging from 0.5 (no discrimination) to 1.0 (perfect discrimination), AUC values 0.5–0.7 are considered low, 0.7–0.9 moderate, and ≥ 0.9 are high.²⁸

Results

A total of 52 participants were screened for inclusion in study procedures. Of those, five participants were not enrolled (four were not eligible and one did not complete a Day 1 visit). Forty-seven participants were enrolled in the study and completed a Day 1 study visit (26 smokers, 21 nonsmokers). Demographics and smoking characteristics of the sample ($N = 47$), separated by smoking categorization, are shown in Table 1. Smokers and nonsmokers did not differ in demographics ($p > .08$). As expected, smokers had higher breath CO and urinary cotinine values than nonsmokers ($p < .001$).

Of the 47 participants enrolled in the study, 42 participants completed all ten study visits and provided 10 CO samples per monitor (30 CO samples total per participant). The remaining five participants completed an average of 3.2 study visits (range 1–6). One participant had a persistently high CO value (≥ 80 ppm) during 8 of 10 study visits. As these elevated CO values were consistent across monitors, we determined that data from this participant should be retained in analyses. Overall, there were a combined total of 437 breath CO samples across all three monitors. Smokers provided significantly fewer CO samples than nonsmokers overall (231 samples from smokers, 8.9 mean comparisons provided vs 206 from nonsmokers, 9.8 mean comparisons provided [$F = 11.5$], $p = .001$).

Preliminary CO Comparisons

Mean differences in CO did not vary due to the time between sample collection for any of the monitor pairs (Wald $\chi^2(1)$'s < 2.9 , p 's $> .09$), and this variable was dropped from all further analyses. Across the entire range of measurement, the iCO gave mean ($\pm$ SEM) CO values 0.27 ± 0.56 ppm higher than the Micro⁺ pro (Wald $\chi^2(1) = 0.23$,

$p = .63$) and 1.14 ± 0.62 ppm higher than the Micro⁺ basic (Wald $\chi^2(1) = 3.41$, $p = .07$). Values from the Micro⁺ pro were an average of 0.87 ± 0.22 ppm higher than those from the Micro⁺ basic (Wald $\chi^2(1) = 15.75$, $p = .00007$). Notably, these mean differences do not convey how differences between monitors increased with larger CO values. The Bland–Altman analyses illustrate how the difference in measurement varied as a function of the mean CO between monitors and address whether these monitors may be used interchangeably.

Aim 1: Bland–Altman Analysis of Agreement

Using GEE to regress the difference between monitors (ie, iCO – Micro⁺ pro, iCO – Micro⁺ basic, Micro⁺ basic – Micro⁺ pro) on the average CO value between respective monitor pairs provided the mean bias across the entire range of CO samples for use in the Bland–Altman plots (Figure 1). For the comparison between the iCO and Micro⁺ pro monitors, the mean bias was $\beta = 0.18$, Wald $\chi^2(1) = 8.67$, $p = .003$, with an intercept of -1.90 . This indicates that the iCO started off approximately 2 ppm lower than the Micro⁺ pro, and the difference between monitors increased by about 2 ppm for each additional ten-unit increase in mean CO. Shown in Figure 1A, for mean CO values ≤ 10 ppm, the iCO-obtained values were lower than those from the Micro⁺ pro, but iCO values were higher than the Micro⁺ pro for CO values > 10 ppm. Further inspection of the Bland–Altman plot (Figure 1A) shows that the GEE-obtained mean bias falls outside of the a priori limits of clinical significance ($\pm 20\%$ of the mean CO between monitors) when CO values are below 10 ppm.

Comparing the iCO to the Micro⁺ basic, the GEE-obtained mean bias was $\beta = 0.17$, Wald $\chi^2(1) = 17.37$, $p = .00003$, with an intercept of -0.95 . Similar to the above comparison, the iCO started off approximately 1 ppm lower than the Micro⁺ basic, and the difference between monitors increased by about 2 ppm for each additional 10-unit increase in mean CO. Readings from the iCO were below those from the Micro⁺ basic when their mean CO was < 6 ppm, after which the iCO values continued to run higher than the Micro⁺ basic (Figure 1B). Although the mean bias was within the a priori limits of clinical significance for CO values > 2 , the 95% LoA (acting as a 95% CI for the mean bias) are consistently outside of the limits of clinical significance.

As shown in Supplementary Figure 1, the mean bias in measurement between the Micro⁺ pro and Micro⁺ basic monitors was found to be minor and did not significantly change across the range of measurement, $\beta = -0.02$, Wald $\chi^2(1) = 0.69$, $p = .41$, with an intercept of -0.63 . In other words, CO values from the Micro⁺ basic started off about 1 ppm lower than that of the Micro⁺ pro monitor and, on average, remained at that level. The mean bias fell within the a priori limits of clinical significance for all CO values > 3 ppm and the relatively narrower 95% LoA bands, suggesting agreement between the Micro⁺ pro and basic.

Aim 2: Monitor-Specific Optimal Abstinence Cutoffs

The diagnostic accuracy for predicting cotinine status was similar between the Micro⁺ pro (AUC = 0.94, 95% CI 0.89, 0.99) and Micro⁺ basic (AUC = 0.91, 95% CI 0.85, 0.98), each falling in the high accuracy range (ie, AUC's > 0.90). The iCO was found to be moderately accurate, with an AUC of 0.87 (95% CI 0.82, 0.96). As all 95% CIs overlapped between monitors, there were no significant differences in diagnostic accuracy.

Sensitivity and specificity for a range of CO values collected during study visits when cotinine was assessed, across all three monitors, are shown in Table 2. The optimal CO criterion for determining abstinence (ie, the CO cutoff maximizing both sensitivity to detect smoking and specificity to verify cotinine status; bolded values in Table 2) was < 3 ppm for the Micro⁺ pro (88% sensitivity, 93% specificity) and Micro⁺ basic (83% sensitivity, 98% specificity), but higher for the iCO at < 6 ppm (73% sensitivity, 100% specificity). If the Micro⁺ pro and Micro⁺ basic cutoff of < 3 ppm were to be used for the iCO, sensitivity would increase to 81%, but specificity would drop to 81%. Therefore, the correct classification of smokers would increase as would the number of nonsmokers/abstainers incorrectly classified as smokers.

Discussion

When comparing the smartphone-enabled iCO with two commercially available, stand-alone CO monitors (Micro⁺ pro and Micro⁺ basic), the Bland–Altman Analysis of Agreement indicated that the

Table 1. Demographics and Smoking Characteristics of the Study Sample ($N = 47$) Separated by Smoking ($n = 26$) or Nonsmoking ($n = 21$) Status

	Smokers ($n = 26$)	Nonsmokers ($n = 21$)
Demographics		
Age, mean (SD)	43.8 (14.2)	35.4 (12.1)
Female % (n)	34.6 (9)	52.4 (11)
Race % (n)		
White	53.8 (14)	71.4 (15)
African American	42.3 (11)	14.3 (3)
More than 1 race/ other/unknown	3.8 (1)	14.3 (3)
Smoking characteristics, mean (SD)		
Cigarettes/d (screening)	13.6 (9.4)	—
Months at current smoking rate	13.1 (11.7)	—
FTND Total Score	4.5 (2.5)	—
Breath CO (ppm; screening)	19.6 (17.7)	1.9 (0.6)
Urinary cotinine (ng/mL; Day 1)	8070 (8517)	2.4 (2.9)
Study outcomes, mean (SD)		
CO comparisons completed (out of 10)	8.9 (2.7)	9.8 (0.9)

Nonsmokers had not smoked in the past 30 d (5/21 were former smokers). One smoker was initially categorized as a light smoker (2/30 past day smoking, negative cotinine sample). This participant was categorized as a nonsmoker for analyses. Breath CO at screening measured using Micro⁺ pro. CO = carbon monoxide; FTND = Fagerström Test of Nicotine Dependence.

Figure 1. Bland-Altman plots comparing iCO and Micro⁺ pro (A) and iCO and Micro⁺ basic (B) monitors. In both plots, dashed black lines “A” and “E” are 95% limits of agreement, dash-dot gray lines “B” and “F” are the limits of clinical significance ($\pm 20\%$ of the average CO reading), solid black line “C” is the GEE-obtained mean difference between monitors, and the solid gray line “D” at 0 illustrates no difference between monitors.

iCO monitor provided systematically different CO results than both the Micro⁺ pro and the Micro⁺ basic monitors. This does not indicate that CO values obtained using the iCO monitor are not valid; instead, these results suggest that monitor-specific abstinence criteria are needed to ensure accurate biochemical verification of abstinence from smoking. The results from this study indicate that it is not

advisable that values from the iCO be used interchangeably with the Micro⁺ pro and Micro⁺ basic, particularly when lower CO values (<10 ppm) are critical (ie, determination of abstinence vs confirming smoking status for study inclusion). However, results did find comparability between the Micro⁺ monitors (Supplementary Figure 1), indicating that these two monitors could be used interchangeably.

Sensitivity and specificity analyses for determining the optimal CO criterion to indicate smoking abstinence was shown to be <3 ppm for the Micro⁺ pro and Micro⁺ basic, but higher for the iCO at <6 ppm. Use of a higher CO cutoff commonly used in smoking cessation studies (ie, <10 ppm)^{1,15,29} with the iCO or Micro⁺ pro would lead to the perfect correct classification of cotinine-confirmed abstinence (100% specificity), but poor detection of actual smoking (63% and 69% sensitivity, respectively). That is, all urinary cotinine values ≤ 50 ng/mL resulted in CO readings < 10 ppm for both monitors, but 37% and 31% of those with cotinine values > 50 ng/mL also had CO readings < 10 ppm for the iCO and Micro⁺ pro monitors, respectively.

The current study extended upon a report that compared CO monitors from different companies and found that agreement was not within 20% of average values (a priori level of acceptable variation) and resulting CO values could not be used interchangeably.¹⁰ Here, we demonstrated that the iCO and two other monitors had systematic variation in CO values that must be taken into account when reporting CO and determining abstinence. Notably, Karelitz et al.¹⁰ found that among published smoking studies that disclosed the brand of CO monitor used, 61% were Bedfont products. Therefore, these products are being used frequently, and these results will aid in comparisons of abstinence outcomes and biochemical readings across monitors and studies. Considering the variability between different CO monitors, these data are important in setting appropriate and data-driven cutoff points to classify abstinence. This also suggests that, depending on the type, CO monitors should not necessarily be used interchangeably in research and clinical settings, which is a practice that is likely common in the field, though rarely reported in manuscripts.

This study has several limitations. First, although research staff observed breath CO samples and instructed participants on how to submit proper samples, there are factors that may affect readings, such as the speed of exhalation,³⁰ that were not controlled in this study. Second, to avoid mobile connection issues and potential

Table 2. Sensitivity to Smoking and Specificity for Identifying Urinary Cotinine-Verified Abstinence by Carbon Monoxide (CO) Cutoff Level and Monitor

	iCO			Micro ⁺ pro			Micro ⁺ basic		
CO cutoff (ppm)	Sensitivity	Specificity	Sensitivity + specificity	Sensitivity	Specificity	Sensitivity + specificity	Sensitivity	Specificity	Sensitivity + specificity
1	1.000	0.122	1.122	1.000	0.000	1.000	1.000	0.024	1.024
2	0.875	0.537	1.412	0.979	0.341	1.320	0.896	0.585	1.481
3	0.813	0.805	1.618	0.875	0.927	1.802	0.833	0.976	1.809
4	0.750	0.927	1.677	0.813	0.976	1.789	0.792	1.000	1.792
5	0.729	0.976	1.705	0.792	0.976	1.768	0.729	1.000	1.729
6	0.729	1.000	1.729	0.771	1.000	1.771	—	—	—
7	0.688	1.000	1.688	0.729	1.000	1.729	0.708	1.000	1.708
8	—	—	—	0.708	1.000	1.708	—	—	—
9	0.667	1.000	1.667	—	—	—	0.688	1.000	1.688
10	0.625	1.000	1.625	0.688	1.000	1.688	—	—	—

Bold values represent optimal CO cutoff values for each monitor (ie, the CO value maximizing both sensitivity and specificity); CO cutoff is the CO level below which would indicate nonsmoker/abstinence. “—” indicates zero CO values at that cutoff from the respective monitor when urine cotinine was assessed at study visits 1 and 10. ppm = parts per million.

delays in submitting samples during the study visit, study devices (iPhones) were used to obtain CO readings from the iCO for participants. We cannot be certain that the mobile device used with the iCO would not affect readings or that readings would not be affected by user error, and as such, this is a limitation and an aspect of this study that would not be replicated when these devices are used remotely with the participant's own phone, as intended. Third, we did not recruit light or intermittent smokers or ask current smokers to engage in a quit attempt during the study, which would have yielded a complete distribution of CO values (abstinent and not abstinent) both between and within participants. Finally, variability within monitors was not in the scope of this study, and we cannot provide an assessment of agreement within the same monitor. Though with proper calibration and testing, we would expect within-monitor variability to be low.

The smartphone-enabled iCO has been recently integrated into mobile apps and used as part of smoking cessation interventions,³¹ with more studies ongoing and greater interest in remote biochemical verification of abstinence. Given recent restrictions on in-person research and the need to limit exposure due to the COVID-19 pandemic, more researchers and clinicians are adopting remote procedures into their work. This change may be likely to continue even after in-person contact resumes, which illustrates the importance of evaluating and validating smartphone-enabled, low-cost CO devices that lend themselves well to remote biochemical collection and verification. Results of the current study provided an optimal, urinary cotinine-verified iCO-specific CO cutoff for determining abstinence, which may be used in remote smoking cessation research.

Supplementary Material

A Contributorship Form detailing each author's specific involvement with this content, as well as any supplementary data, are available online at <https://academic.oup.com/ntr>.

Funding

Funding for this study was provided in part by pilot research funding from the Hollings Cancer Center's Cancer Center Support Grant P30 CA138313 at the Medical University of South Carolina and the South Carolina Clinical and Translational Research Institute at MUSC (NIH/NCATS UL11 TR001450). Other support to complete this work was provided by grants from NIDA K23 DA045766 (PI, Dahne), NCI R21 CA241842 (PI, Dahne), pilot research funding from an American Cancer Society Institutional Research Grant awarded to the Hollings Cancer Center at the Medical University of South Carolina (ACS IRG 97-219-14, awarded to Dahne), the National Cancer Institute T32 CA186873 at the University of Pittsburgh (PI, Yuan), and NICHD K12 HD055885 (PI, McGinty to Tomko).

Acknowledgments

We wish to thank Benjamin Laprise, Elizabeth Bradley, Jenny Nankoua, and Lori Ann Ueberroth (members of the Project Quit Team) at the Medical University of South Carolina, who assisted in the collection and management of data and the successful execution of this study. We also thank Kevin M. Gray and Matthew J. Carpenter for helpful feedback and comments on the experimental design. Finally, we wish to sincerely thank the participants who contributed data to this study.

Declaration of Interests

None declared.

References

1. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine Tob Res.* 2002;4(2):149–159.
2. Benowitz NL, Bernert JT, Foulds J, et al. Biochemical verification of tobacco use and abstinence: 2019 update. *Nicotine Tob Res.* 2020;22(7):1086–1097.
3. Jarvis MJ, Tunstall-Pedoe H, Feyerabend C, Vesey C, Saloojee Y. Comparison of tests used to distinguish smokers from nonsmokers. *Am J Public Health.* 1987;77(11):1435–1438.
4. Dallery J, Raiff BR, Kim SJ, Marsch LA, Stitzer M, Grabinski MJ. Nationwide access to an internet-based contingency management intervention to promote smoking cessation: a randomized controlled trial. *Addiction.* 2017;112(5):875–883.
5. Hertzberg JS, Carpenter VL, Kirby AC, et al. Mobile contingency management as an adjunctive smoking cessation treatment for smokers with posttraumatic stress disorder. *Nicotine Tob Res.* 2013;15(11):1934–1938.
6. Karelitz JL, Michael VC, Boldry M, Perkins KA. Validating use of internet-submitted carbon monoxide values by video to determine quit status. *Nicotine Tob Res.* 2017;19(8):990–993.
7. Herbec A, Brown J, Shahab L, West R. Lessons learned from unsuccessful use of personal carbon monoxide monitors to remotely assess abstinence in a pragmatic trial of a smartphone stop smoking app – a secondary analysis. *Addict Behav Rep.* 2019;9:100122.
8. Johnson K, Isham A, Shah DV, Gustafson DH. Potential roles for new communication technologies in treatment of addiction. *Curr Psychiatry Rep.* 2011;13(5):390–397.
9. Riley WT, Oh A, Aklin WM, Wolff-Hughes DL. National Institutes of Health support of digital health behavior research. *Health Educ Behav.* 2019;46(2 suppl):12–19.
10. Karelitz JL, Michael VC, Perkins KA. Analysis of agreement between expired-air carbon monoxide monitors. *J Smok Cessat.* 2017;12(2):105–112.
11. Wong HY, Subramaniam M, Bullen C, Siddiq A, Danaee M, Yee A. The mobile-phone-based iCO Smokerlyzer: comparison with the piCO+ Smokerlyzer among smokers undergoing methadone-maintained therapy. *Tob Induc Dis.* 2019;17:1–5.
12. Perkins KA, Karelitz JL, Jao NC. Optimal carbon monoxide criteria to confirm 24-hr smoking abstinence. *Nicotine Tob Res.* 2013;15(5):978–982.
13. Cropsey KL, Trent LR, Clark CB, Stevens EN, Lahti AC, Hendricks PS. How low should you go? Determining the optimal cutoff for exhaled carbon monoxide to confirm smoking abstinence when using cotinine as reference. *Nicotine Tob Res.* 2014;16(10):1348–1355.
14. Erb P, Raiff BR, Meredith SE, Dallery J. The accuracy of a lower-cost breath carbon monoxide meter in distinguishing smokers from non-smokers. *J Smok Cessat.* 2015;10(1):59–64.
15. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction.* 2005;100(3):299–303.
16. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50(6):613–621.
17. Centers for Disease Control and Prevention. *Measuring Healthy Days.* Atlanta, GA: CDC; 2000.
18. Sobell LC, Sobell MB, Leo GI, Cancelli A. Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *Br J Addict.* 1988;83(4):393–402.
19. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict.* 1991;86(9):1119–1127.
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377–381.
21. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics.* 1988;44(4):1049–1060.
22. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8(2):135–160.

23. Altman DG. Construction of age-related reference centiles using absolute residuals. *Stat Med.* 1993;12(10):917–924.
24. Brown R, Richmond S. An update on the analysis of agreement for orthodontic indices. *Eur J Orthod.* 2005;27(3):286–291.
25. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet.* 1995;346(8982):1085–1087.
26. Myles PS, Cui J. Using the Bland–Altman method to measure agreement with repeated measures. *Br J Anaesth.* 2007;99(3):309–311.
27. Tomko RL, McClure EA, Cato PA, et al. An electronic, smart lighter to measure cigarette smoking: a pilot study to assess feasibility and initial validity. *Addict Behav.* 2019;98:106052.
28. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med.* 2003;29(7):1043–1051.
29. Department of Health. *Local Stop Smoking Services: Service Delivery and Monitoring Guidance 2011/12.* London, UK: United Kingdom National Health Service; 2011.
30. Raiff BR, Faix C, Turturici M, Dallery J. Breath carbon monoxide output is affected by speed of emptying the lungs: implications for laboratory and smoking cessation research. *Nicotine Tob Res.* 2010;12(8):834–838.
31. Krishnan N, Elf JL, Chon S, Golub JE. COach2Quit: a pilot randomized controlled trial of a personal carbon monoxide monitor for smoking cessation. *Nicotine Tob Res.* 2019;21(11):1573–1577.