ANALYSIS OF AGREEMENT BETWEEN EXPIRED-AIR CARBON MONOXIDE MONITORS

Joshua L. Karelitz, **Valerie C. Michael**, and **Kenneth A. Perkins**

Department of Psychiatry of the University of Pittsburgh, Pittsburgh, Pennsylvania

Abstract

Introduction—The current study examined the level of agreement in expired-air carbon monoxide (CO) values, focusing especially on those confirming abstinence, between the two most commonly used CO monitors, the Vitalograph BreathCO and the Bedfont piCO+ Smokerlyzer.

Methods—Expired-air samples were collected via both monitors from adult dependent smokers (44 M, 34 F) participating in studies using CO values to confirm abstinence durations of: 24 hours, 12 hours, or no abstinence. All met DSM-IV nicotine dependence criteria and had a mean (SD) Fagerström Test of Cigarette Dependence score of 5.1 (1.8). Paired data collected across multiple visits were analyzed by regression-based Bland-Altman method of Limits of Agreement.

Findings—Analysis indicated a lack of agreement in CO measurement between monitors. Overall, the Bedfont monitor gave mean (±SEM) readings 3.83 (±.23) ppm higher than the Vitalograph monitor. Mean differences between monitors were larger for those ad lib smoking (5.65±.38 ppm) than those abstaining 12–24 hours (1.71±.13 ppm). Yet, there also was not consistent agreement in classification of 24 hour abstinence between monitors.

Conclusions—Systematic differences in CO readings demonstrate these two very common monitors may not result in interchangeable values, and reported outcomes in smoking research based on CO values may depend on the monitor used.

INTRODUCTION

Expired-air carbon monoxide (CO) measurement is an objective, noninvasive method to biochemically assess recent smoking exposure. With an average half-life of 4-hrs, expired-air CO has been shown to be sensitive to recent smoking exposure (i.e., <24 hrs; Benowitz et al., 2002). Clinical research uses CO values as a method to confirm abstinence in treatment studies (typically for >24 hrs) or brief deprivation (e.g. overnight) prior to laboratory testing of symptoms such as craving or withdrawal (Addicott et al., 2012; Perkins, Jao, & Karelitz,

Address for correspondence: JL Karelitz, Phone: (412) 605-0552; FAX: (412) 586-9838; karelitzjl@upmc.edu.
J.L. Karelitz, M.A., V.C. Michael, B.S., K.A. Perkins, PhD, WPIC, University of Pittsburgh School of Medicine, 3811 O'Hara Street, Pittsburgh, PA 15213 USA;

CONFLICT OF INTERESTS

The authors do not have any financial interests in either of the CO monitor companies discussed in this paper. None of the authors have any other potential conflicts of interest to report.

ETHICAL STANDARDS

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helinski Declaration of 1975, as revised in 2008.

2013; Perkins, Karelitz, Conklin, Sayette, & Giedgowd, 2010). Still others have used CO level as an inclusion criterion to ensure current status as a smoker (Froeliger, Modlin, Wang, Kozink, & McClernon, 2012; VanderVeen, Cohen, Cukrowica, & Trotter, 2008) or to assess current amount of smoke exposure (Perkins, Karelitz, & Jao, 2013). Unlike the primary biochemical method to assess exposure to nicotine per se, salivary or urinary cotinine, CO measurement has the benefit of not being affected by nicotine replacement therapies used in some cessation interventions (Benowitz et al., 2002).

Although a variety of CO monitors are available (European Respiratory Society, n.d.), the most widely used monitors in smoking research are from Bedfont Scientific (Kent, United Kingdom) and Vitalograph (Lenexa, KS). A Google Scholar (www.scholar.google.com) search over a ten year period (2004 to 2014; performed August, 2015) using the keywords "CO monitor" and "smoking" returned 745 publications. Surprisingly, 434 (58%) of the articles did not disclose the monitor type. Of the 311 that did, the vast majority used Bedfont (189, or 61%) or Vitalograph (109, 35%) monitors, while just 11 used Micro Medical (Kent, United Kingdom) and 2 used CO Sleuth (Breathe E-Z Systems, Leawood, KS).

Possible variation in measurement between monitors may have implications for the results reported in studies using CO values to confirm current smoking or recent duration of abstinence. Using Bedfont monitors, Erb, Raiff, Meredith, and Dallery (2014) identified CO values <8 ppm as maximizing both sensitivity and specificity for classifying smokers from nonsmokers. By contrast, Perkins et al. (2013) found < 5 ppm to be the optimal cutoff for 24 hour abstinence when using the Vitalograph BreathCO monitor, similar to <4 ppm found by Javors, Hatch, and Lamb (2005), Cropsey et al. (2014) and Emery and Levine (2015) when also using the Vitalograph monitor. This discrepancy suggests there may be variation between these brands of CO monitors in validating abstinence and raises concern over whether they can be used interchangeably.

The aim of this study was to examine the level of agreement in measurement between two very commonly used CO monitors, the Vitalograph BreathCO and the Bedfont piCO+ Smokerlyzer, through Bland-Altman analysis of agreement (Bland & Altman, 1986) of continuous values and comparison of dichotomous smoking abstinence classifications. Bland-Altman analysis determines the level of agreement in measurement between monitors rather than their accuracy (i.e., closeness to true value) or precision (i.e., closeness between repeated measures) (Mantha, Roizen, Fleisher, Thisted, & Foss, 2000), both of which have been reported elsewhere per each monitor (CoVita, 2010; Javors, et al., 2005; Moscato, Poscia, Gargaruti, Capelli, & Cavaliere, 2014; Vitalograph, n.d.). Knowing whether these monitors agree (e.g., 4 ppm for one is 4 ppm for the other) will allow for the monitors and diagnostic criteria to be used interchangeably. If the monitors do not agree in measurement, monitor-specific abstinence criteria may be necessary. Additionally, there are implications for replication of studies between researchers using different monitors.

METHODS

Participants

To provide a range of recent smoking exposure, participants (n=78) were those from studies using CO values to confirm 24 hour abstinence (verified by <5ppm; Perkins, Karelitz, & Jao, 2013), 12-hour (overnight) abstinence (verified by 10ppm, Javors, et al., 2005), or no abstinence (i.e. current ad lib smoking, verified by >10ppm). Participants agreed to provide additional expired-air samples for this comparison between monitors, secondary to the main study in which they were enrolled. Those studies included an initial evaluation of a novel smoking cessation medication in those wanting help to quit (Perkins et al., 2015), and laboratory studies in those not wanting to quit that examined responses during nonabstinence and overnight abstinence conditions (Perkins, Karelitz, & Michael, 2015), including one that is ongoing. Values were assessed between April 2014 and January 2015. Participants were recruited through print and online notices for the respective studies, placed in the surrounding University of Pittsburgh community. All were healthy adults (44 M, 34 F) who met inclusion criteria for each study, both involving smoking 10 cigarettes per day for 1 year and meeting DSM-IV nicotine dependence criteria (American Psychiatric Association, 1994), those very typically recruited for studies of smoking behavior and abstinence (e.g. Addicott et al., 2012; Benowitz et al. 2002, Perkins, Jao, & Karelitz, 2013; Perkins, Karelitz, Conklin, Sayette, & Giedgowd, 2010). Mean (SD) sample characteristics were 16.2 (5.3) cigarettes per day, 28.8 (8.9) years old, and 5.1 (1.8) Fagerström Test of Cigarette Dependence score (previously known as Fagerström Test of Nicotine Dependence, Fagerström, 2012; Heatherton, Kozlowski, Frecker, & Fagerström, 1991).

CO Monitors

Expired-air samples were obtained using Vitalograph BreathCO monitors and Bedfont pico+ Smokerlyzer monitors. The Bedfont pico+ Smokerlyzer has an LCD screen to guide the participant through the expired-air CO process. The Vitalograph BreathCO monitor requires an administrator to guide the participant through the entire process (described in detail below). Both monitors have a range of measurement of 0–100 ppm and display CO values in 1 ppm increments. The monitors were calibrated to brand specific standards (Vitalograph 50 ppm, Bedfont 20 ppm) prior to testing and then as recommended by each respective manufacturer (Vitalograph every 12 months, Bedfont every 6 months). Three Vitalograph BreathCO monitors and three Bedfont pico+ Smokerlyzer monitors were used for this study. In 2014, before the study began, the Vitalograph monitors had new sensors installed by the manufacturer and all three Bedfont monitors were purchased new. According to each respective CO monitor's manual, the BreathCO monitor is accurate ±3 ppm (Vitalograph, n.d.) and the pico+ Smokerlyzer is accurate ±2% (CoVita, 2010). Previous studies have found both monitors to have high internal consistency, with intraclass correlation coefficients for the Vitalograph ranging from 0.935 to 0.994 (Javors, et al., 2005) and 0.985 for the Bedfont monitor (Moscato et al., 2014)

Procedure

Participants were observed as they provided air samples following the instructions specific to each brand. Because observed CO values can be affected by speed of exhalation (Raiff,

Faix, Turturici, & Dallery, 2010), participants were instructed to exhale slowly when providing all samples across CO monitors. For the Vitalograph monitor, participants were told to take a deep breath and hold for 20 sec (measured by experimenter via digital timer) then exhale fully into the mouthpiece attached the monitor. For the Bedfont monitor, the participant double-clicked a button on the monitor, initiating the breath sample collection. The Bedfont monitor then guided the participant through the entire procedure using pictures on an LED screen. Participants were instructed to hold their breath for 15 sec (measured by the Bedfont monitor) and then exhale fully through the mouthpiece attached to the monitor. All pairs of air samples were collected with at least 5 min, but no more than 10 min, between samples, with the Vitalograph monitor administered before the Bedfont monitor in all cases. The number of visits varied between studies, resulting in a varying number of paired observations from each participant.

Statistical Analysis

As noted, the outcome results for expired-air CO values were compared at clinical criteria for 24 hour abstinence (<5ppm), overnight abstinence (12 hours,   10ppm), and ad lib smoking (>10ppm). Of primary interest was the agreement in classification between monitors, particularly determining how many participants would meet abstinent or ad lib criteria for one monitor when the criterion is met under the other monitor.

A Bland-Altman analysis of agreement (Bland & Altman, 1986) was performed to measure the level of agreement between CO monitors. This analysis relies on a plot of the difference between monitors (Bedfont-Vitalograph) against the mean reading for both combined ((Bedfont + Vitalograph)/2), determined for each pair of observations. Typically, a new scale or measurement type is compared against an existing or "gold standard" method. Bland and Altman (1995) explicitly stated that comparing the difference between measurements to a standard method would inflate the risk for Type I error. Following their suggestion, and not having an established standard to compare against, the mean reading between monitors was used as a proxy for the "true" CO value. The Bland-Altman plot shows how the agreement in measurement between monitors varies as a function of an increase in CO level due to recent smoking exposure. The mean difference between monitors (bias) is computed, along with the 95% limits of agreement. The 95% limits of agreement (LoA) form a band around the mean bias and act as 95% confidence intervals. A narrow LoA band width indicates good agreement between methods and a wide LoA band width indicates poor agreement. A two-step regression-based analysis was used to calculate the 95% LoA, due to data with nonconstant variance in the differences, repeated measures, and varying number of observations from each participant. Details of this method have been described in previous articles (Altman, 1993; Bland & Altman, 1999; Brown & Richmond, 2005). All analyses were performed using SPSS version 22.0.

The Bland-Altman analysis does not provide a level of statistical significance; however, observed differences are compared against an a priori threshold of clinical significance. If the observed differences fall within the threshold of clinical significance, it is presumed that the methods can be used interchangeably. This analysis is preferred over more familiar statistics (i.e., correlation), due to its ability to evaluate the difference in measurement rather

than assess the degree of association between methods of measurement. The threshold of clinical significance for this study was defined as ±20% of the average reading between monitors. This value was chosen as it would allow for ±1 ppm variation at an average reading between monitors of 5 ppm, a 24-hr abstinence cutoff level shown to have high levels of sensitivity and specificity (Perkins, Karelitz, & Jao, 2013). It also would allow for ±2 ppm variation for confirming 12-hr abstinence (at CO<10 ppm). For those ad lib smoking, ±20% of the average CO reading for those ad lib smoking allowed for an acceptable range of variation that was in proportion to the CO value obtained across all levels of measurement. At these higher levels of CO, a proportionally larger level of variation was considered acceptable as these specific values are typically not used to guide diagnostic or clinical decisions.

RESULTS

A scatterplot of paired CO values from each monitor is shown in Figure 1, with a regression line (dashed) showing best fit and a diagonal line (solid) illustrating perfect agreement. Across the entire sample (654 pairs of observations), the Bedfont monitor gave a mean (±SEM) reading 3.83 (±.23) ppm higher than the Vitalograph monitor.

Comparisons of Dichotomous Abstinence Duration Criteria

In comparisons of the monitors for verifying 24-hour abstinence, Figure 2 contains a pair of bar graphs displaying frequencies of CO values from each monitor when the opposite monitor is <5 ppm. For the set of breath measurements when CO was <5 ppm using the Bedfont, 87.8% of the values obtained with the Vitalograph were also <5 ppm (79 out of 90). Conversely, when Vitalograph values were <5 ppm, the Bedfont values ranged from 1 ppm to 13 ppm, with only 48.5% of values also <5 ppm (79 out of 163).

When looking at those meeting the 12 hour abstinence criterion (10 ppm) using the Bedfont monitor, Vitalograph CO values ranged from 1 ppm to 16 ppm, with 96.7% of values 10 ppm. When the same criterion was met using the Vitalograph monitor, Bedfont CO values ranged from 1 ppm to 17 ppm, with 88.4% 10 ppm.

For confirming no abstinence, those classified as ad lib smoking (CO >10 ppm) using the Bedfont monitor had Vitalograph values ranging from 2 ppm to 55 ppm, with 90.8% >10 ppm. When the ad lib smoking criterion was met using the Vitalograph monitor, Bedfont CO values ranged from 8 ppm to 78 ppm, with 97.5% >10 ppm.

Bland-Altman Comparison of Continuous Values

Combining all CO values and regressing the differences between monitors (Bedfont-Vitalograph) on the average reading ((Bedfont + Vitalograph)/2) provided the mean bias across the range of samples, β=.26, t(653)=20.21, p<.00001, with an intercept of −.65. Such bias suggests for each ten-unit increase in CO, the Bedfont monitor will provide CO readings about 3 ppm higher than the average between the two monitors. To calculate the 95% LoA, the absolute values of the residuals from the first regression were regressed on the average CO readings between monitors, resulting in a significant linear relationship, β=.09,

t(653)=9.58, p<.00001, with an intercept of 1.57. This equation was combined with the earlier equation to form the 95% LoA bands.

The Bland-Altman graph is displayed in Figure 3, plotting the difference between monitors (Bedfont-Vitalograph) against their mean reading. The regression-obtained mean bias is shown (line 'B'), along with the 95% LoA (lines 'A' and 'C'). If the monitors were in complete agreement across the range of measurement, the mean bias would be a flat (statistically non-significant) line with an intercept close to zero. Inspection of Figure 3 shows widening LoA bands as the average CO values increase with a similarly increasing mean bias showing that agreement between monitors decreases as CO values increase.

Figure 4 is a Bland-Altman plot with a horizontal line at zero (line 'E') illustrating complete agreement (e.g., no difference) between monitors, with bands forming the upper and lower limits of clinical significance (±20% of the average CO value; lines 'D' and 'F'). These limits of clinical significance show the acceptable amount of variation in measurement of CO between monitors. Also shown in Figure 4 is the regression-obtained mean bias for comparison to the limits of clinical significance. The regression-obtained mean bias falls outside of the a priori limits of clinical significance, indicating a lack of agreement in measurement of CO between monitors.

DISCUSSION

We compared agreement in abstinence duration classification (24 hour, 12 hour, none--ad lib smoking) and overall expired-air CO measurement between two commonly used CO monitors, the Vitalograph BreathCO and the Bedfont pico+ Smokerlyzer. Values obtained from the Bedfont monitor were consistently higher than those from the Vitalograph monitor. The elevated CO values obtained from the Bedfont monitors are consistent with a previous study finding overestimation of CO by roughly 20% when measuring known concentrations of carbon monoxide of 5 ppm and 9.9 ppm (Moscato et al., 2014).

When confirming 24 hour abstinence (<5 ppm) in dichotomous fashion, as typically reported in smoking cessation research (e.g. Javors et al. 2005; Perkins et al. 2013), the Vitalograph monitors agreed with the Bedfont classifications 87.8% of the time. However, when the same criterion was met using the Vitalograph monitors, the Bedfont monitors agreed less than half of the time (48.5%), indicating reported results for abstinence rates in clinical trials could be influenced by the monitor used. Yet, there was greater agreement in determining 12 hour abstinence and ad lib smoking. The Bedfont monitors agreed with Vitalograph readings of 10 ppm 88.4% of the time and 97.5% of the time when Vitalograph values >10 ppm, while the Vitalograph values agreed with 96.7% of the Bedfont values 10 ppm and with 90.8% of the Bedfont values >10 ppm.

In order to test whether the CO monitors could be used interchangeably in studies of recent smoking exposure, we performed a regression-based Bland-Altman analysis of agreement, which compares CO values in continuous fashion. Values obtained from the Bedfont monitors were consistently higher than those from the Vitalograph monitors, with larger absolute differences as the CO values increased. Thus, along with less agreement when

confirming abstinence dichotomously, the monitors also show less agreement in comparisons of continuous CO values with greater or more recent smoking exposure. The Bland-Altman plot illustrated this variation of differences along the range of CO measurement. The limits of agreement (LoA) calculated using the Bland-Altman procedure provided upper and lower limits between which 95% of future measurements are estimated to fall, similar to confidence intervals. This band of limits was more narrow at the lower range of measurement (e.g., closer to CO cutoff values for abstinence) than at the higher range of measurement. Comparison of the regression-based mean bias to the a priori level clinical significance (±20% of the average CO reading) showed the mean difference in measurement between monitors was outside of the range of accepted difference, indicating a lack of agreement in CO measurement between monitors. In other words, these two CO monitors may not be appropriate to use interchangeably.

Our study had a few limitations. First, we did not counterbalance the order of monitors, always administering the Vitalograph monitors before the Bedfont monitors. The Vitalograph monitor was administered first to standardize the primary measure of CO in all studies, whereas the Bedfont monitors were administered solely for the comparisons described in this paper. However, any effect of this order is uncertain, as repeated testing of expired-air CO levels has been shown to yield both higher (Jarvis, Belcher, Vesey, & Hutchison, 1986) and lower values (Middleton & Morice, 2000) between samples within the same monitor. A slight decline with the second testing might be expected, given that CO half-life is influenced by respiration (Benowitz et al., 2002), but the second test was always done here with the Bedfont, which tended to read higher rather than lower. Second, the duration of breath holding prior to exhaling into each monitor was different between monitor types, as per each respective set of brand specific instructions. Previous research has found that shorter breath hold duration, as with the Bedfont instructions, was associated with lower levels of expired-air CO (West, 1984), which again is the opposite of what was found in this study. Future replications of this study design should test the effect of breath holding duration to eliminate this as a possible source of difference between monitors.

There were multiple strengths in our study and its design. The within-subjects design allowed our data to withstand individual differences that have been shown to influence expired-air CO measurement, such as speed of exhalation (Raiff et al., 2010), time of last cigarette (Benowitz et al., 2002), and other health issues (i.e., lactose intolerance, liver sepsis, etc.; McNeill, Owen, Belcher, Sutherland, & Fleming, 1990; Morimatsu et al., 2010). Additionally, we collected air samples from smokers across a spectrum of smoking abstinence duration levels to determine if the agreement between monitors might be relevant only for studies of acute smoking behavior or of cessation. Further, the regression-based Bland-Altman statistic was robust enough to handle repeated measurements with varying numbers of observations per subject (Bland & Altman, 1999). This statistic is widely regarded as the ideal way to compare methods of measurement, employing a practical and clinically relevant way to determine whether methods are able to be used interchangeably (Bland & Altman, 2003; Hanneman, 2008; Ludbrook, 2002).

The lack of agreement in classification of abstinence status and measurement of CO between these monitors has several implications. First, it is important to report the specific brand and

model of CO monitor used in any smoking study, especially when comparing results across different studies. As the current work demonstrated, these monitors can differ significantly, making interpretation of clinical results and future replication difficult without knowing how the data were collected. Second, the brand of CO monitor used to verify recent abstinence may need to be considered when evaluating the numbers of participants above versus below a study's dichotomous cutoff criterion for 24-hour abstinence (i.e. quit rate) as a result of different treatment conditions, as the observed CO values may depend on the monitor. These monitor specific values should be taken into consideration when determining abstinence standards for future studies. Similarly, studies of acute smoking exposure that rely on absolute CO values in continuous fashion may also need to report the monitor used, as the difference in CO values between these monitors increased with greater recent smoking. Third, reasons for why these monitors systematically vary in CO readings may warrant research attention, such as the instructions to participants or the actual components in the monitors.

Additional research is suggested by the findings of the current study. First, replication of these results in studies counter-balancing the order of the monitors being compared would be important to fully document variation in measurement between CO monitors. Second, it may be useful to identify monitor-specific cutoffs to determine smoking abstinence. The optimal CO value that maximizes sensitivity and specificity for identifying those quit versus not quit should be determined for each monitor type and compared between types, especially within the same participants. Such research may benefit by using other biomarkers of recent nicotine exposure (e.g., cotinine) as a reference point, similar to methods used by Cropsey et al. (2014). The use of additional biomarkers as reference would strengthen future studies as these measures would not be affected by other combustibles (e.g., marijuana) and would be sensitive to noncombustible sources of nicotine (e.g., electronic cigarettes, NRT, etc.).

Aside from absolute CO values to confirm abstinence durations, as in this study, percent reduction in CO values has also been used to verify minimum durations of cigarette smoking abstinence (Dallery & Raiff, 2007; Lamb, Morral, Galbicka, Kirby, & Iguchi, 2005). This method requires participants to demonstrate some duration of abstinence (usually <24 hr) through a reduction in CO relative to a previously obtained baseline value. Because the current study examined differences in absolute CO values between monitors, we cannot directly determine if magnitude of CO decline from baseline may vary as a function of monitor brand.

Finally, CO values from new CO monitors should be compared with existing monitors using the Bland and Altman procedure to ensure the values obtained between monitors varying in make and/or age are similar and the monitors can be used interchangeably. For example, as Bland and Altman (1986) have pointed out, significant correlations between monitors only indicate the strength of the linear relationship, not the agreement or difference between them. Additional research is needed to further elucidate the differences in CO measurement between monitors. Comparing agreements in measurement of known CO values across a range of concentrations would be the most direct way to assess this and could be another step in this line of research.

Acknowledgments

References

Addicott MA, Baranger DA, Kozink RV, Smoski MJ, Dichter GS, McClernon FJ. Smoking withdrawal is associated with increases in brain activation during decision making and reward anticipation: a preliminary study. Psychopharmacology. 2012; 219(2):563–573. DOI: 10.1007/s00213-011-2404-3 [PubMed: 21766170]

Altman DG. Construction of age-related reference centiles using absolute residuals. Statistics in Medicine. 1993; 12(10):917–924. DOI: 10.1002/sim.4780121003 [PubMed: 8337548]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 1994.

Benowitz NL, Jacob P, Hall S, Tsoh J, Ahijevych K, Jarvis MJ, LeHouezec J. Biochemical verification of tobacco use and cessation. Nicotine and Tobacco Research. 2002; 4(2):149–159. DOI: 10.1080/14622200210123581 [PubMed: 12028847]

Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986; 327(8476):307–310. DOI: 10.1016/S0140-6736(86)90837-8

Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. The Lancet. 1995; 346(8982):1085–1087. DOI: 10.1016/S0140-6736(95)91748-9

Bland JM, Altman DG. Measuring agreement in method comparison studies. Statistical Methods in Medical Research. 1999; 8(2):135–160. DOI: 10.1177/096228029900800204 [PubMed: 10501650]

Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound in Obstetrics and Gynecology. 2003; 22(1):85–93. DOI: 10.1002/uog.122 [PubMed: 12858311]

Brown R, Richmond S. An update on the analysis of agreement for orthodontic indices. The European Journal of Orthodontics. 2005; 27(3):286–291. DOI: 10.1093/ejo/cjh078 [PubMed: 15947229]

CoVita. piCO+ Smokerlyzer operating manual. 2010. Retrieved from http://www.covita.net/pdfs/piCO +%20Manual.pdf

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 & Tobacco Research. 2014; 16(10):1348–1355. DOI: 10.1093/ntr/ntu085 [PubMed: 24891552]

Dallery J, Raiff BR. Delay discounting predicts cigarette smoking in a laboratory model of abstinence reinforcement. Psychopharmacology. 2007; 190(4):485–496. DOI: 10.1007/s00213-006-0627-5 [PubMed: 17205320]

Emery RL, Levine MD. Optimal carbon monoxide criteria to confirm smoking status among postpartum women. Nicotine & Tobacco Research. 2015; doi: 10.1093/ntr/ntv196

Erb P, Raiff BR, Meredith SE, Dallery J. The Accuracy of a Lower-Cost Breath Carbon Monoxide Meter in Distinguishing Smokers from Non-smokers. Journal of Smoking Cessation. 2015; 10(01): 59–64. DOI: 10.1017/jsc.2013.37

European Respiratory Society. Exhaled CO measurement. n.d. Retrieved January 6, 2015, from www.ersbuyersguide.org/features/category/exhaled-co-measurement

Fagerström K. Determinants of tobacco use and renaming the FTND to the Fagerström Test for Cigarette Dependence. Nicotine and Tobacco Research. 2012; 14(1):75–78. DOI: 10.1093/ntr/ntr137 [PubMed: 22025545]

Froeliger B, Modlin L, Wang L, Kozink RV, McClernon FJ. Nicotine withdrawal modulates frontal brain function during an affective Stroop task. Psychopharmacology. 2012; 220(4):707–718. DOI: 10.1007/s00213-011-2522-y [PubMed: 21989805]

Hanneman SK. Design, analysis and interpretation of method-comparison studies. AACN Advanced Critical Care. 2008; 19(2):223.doi: 10.1097/01.AACN.0000318125.41512.a3 [PubMed: 18560291]

Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for nicotine dependence: a revision of the Fagerström Tolerance Questionnaire. British Journal of Addiction. 1991; 86(9):1119–1127. DOI: 10.1111/j.1360-0443.1991.tb01879.x [PubMed: 1932883]

Jarvis MJ, Belcher M, Vesey C, Hutchison DC. Low cost carbon monoxide monitors in smoking assessment. Thorax. 1986; 41(11):886. [PubMed: 3824275]

Javors MA, Hatch JP, Lamb RJ. Cut-off levels for breath carbon monoxide as a marker for cigarette smoking. Addiction. 2005; 100(2):159–167. DOI: 10.1111/j.1360-0443.2004.00957.x [PubMed: 15679745]

Lamb RJ, Morral AR, Galbicka G, Kirby KC, Iguchi MY. Shaping reduced smoking in smokers without cessation plans. Experimental and Clinical Psychopharmacology. 2005; 13(2):83.doi: 10.1037/1064-1297.13.2.83 [PubMed: 15943541]

Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. Clinical and Experimental Pharmacology and Physiology. 2002; 29(7):527–536. DOI: 10.1046/j.1440-1681.2002.03686.x [PubMed: 12060093]

Mantha S, Roizen MF, Fleisher LA, Thisted R, Foss J. Comparing methods of clinical measurement: reporting standards for Bland and Altman analysis. Anesthesia & Analgesia. 2000; 90(3):593–602. DOI: 10.1097/00000539-200003000-00018 [PubMed: 10702443]

McNeill AD, Owen LA, Belcher M, Sutherland G, Fleming S. Abstinence from smoking and expired-air carbon monoxide levels: lactose intolerance as a possible source of error. American Journal of Public Health. 1990; 80(9):1114–1115. DOI: 10.2105/AJPH.80.9.1114 [PubMed: 2382751]

Middleton ET, Morice AH. Breath carbon monoxide as an indication of smoking habit. CHEST Journal. 2000; 117(3):758–763. DOI: 10.1378/chest.117.3.758

Morimatsu H, Takahashi T, Matsusaki T, Hayashi M, Matsumi J, Shimizu H, … Morita K. An increase in exhaled CO concentration in systemic inflammation/sepsis. Journal of Breath Research. 2010; 4(4):047103.doi: 10.1088/1752-7155/4/4/047103 [PubMed: 21383490]

Moscato U, Poscia A, Gargaruti R, Capelli G, Cavaliere F. Normal values of exhaled carbon monoxide in healthy subjects: comparison between two methods of assessment. BMC Pulmonary Medicine. 2014; 14(1):204.doi: 10.1186/1471-2466-14-204 [PubMed: 25515007]

Perkins KA, Jao NC, Karelitz JL. Consistency of daily cigarette smoking amount in dependent adults. Psychology of Addictive Behaviors. 2013; 27(3):723.doi: 10.1037/a0030287 [PubMed: 23088404]

Perkins KA, Karelitz JL, Conklin CA, Sayette MA, Giedgowd GE. Acute negative affect relief from smoking depends on the affect situation and measure but not on nicotine. Biological Psychiatry. 2010; 67(8):707–14. DOI: 10.1016/j.biopsych.2009.12.017 [PubMed: 20132927]

Perkins KA, Karelitz JL, Jao NC. Optimal carbon monoxide criteria to confirm 24-hr smoking abstinence. Nicotine and Tobacco Research. 2013; 15(5):978–82. DOI: 10.1093/ntr/nts205 [PubMed: 22990219]

Perkins KA, Karelitz JL, Michael VC. Reinforcement enhancing effects of acute nicotine via electronic cigarettes. Drug and Alcohol Dependence. 2015; 153:104–108. DOI: 10.1016/j.drugalcdep.2015.05.041 [PubMed: 26070455]

Perkins KA, Karelitz JL, Michael VC, Fromuth M, Conklin CA, Chengappa KR, Hope C, Lerman C. Initial evaluation of fenofibrate for efficacy in aiding smoking abstinence. Nicotine & Tobacco Research. 2015; doi: 10.1093/ntr/ntv085

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 and Tobacco Research. 2010; 12(8):834–838. DOI: 10.1093/ntr/ntq090 [PubMed: 20530193]

VanderVeen JW, Cohen LM, Cukrowicz KC, Trotter DR. The role of impulsivity on smoking maintenance. Nicotine & Tobacco Research. 2008; 10(8):1397–1404. [PubMed: 18686188]

Vitalograph. Vitalograph BreathCO quick start user training manual. n.d. Retrieved from http://vitalograph.com/downloads/view/07323

West RJ. The effect of duration of breath-holding on expired air carbon monoxide concentration in cigarette smokers. Addictive Behaviors. 1984; 9(3):307–309. DOI: 10.1016/0306-4603(84)90026-1 [PubMed: 6496208]
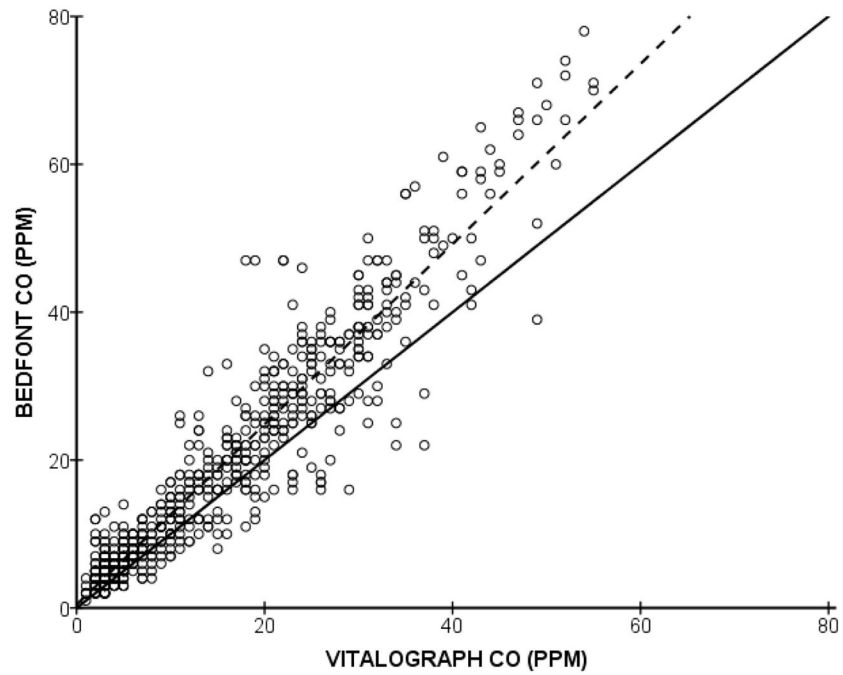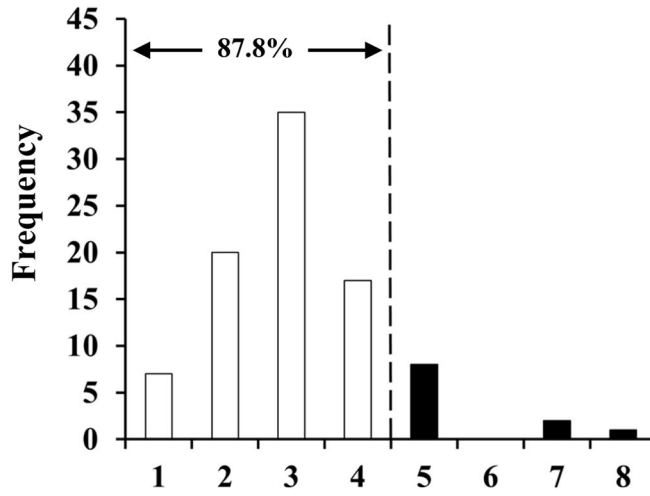
Figure 1.
Scatterplot of paired CO values from each monitor, with the solid line illustrating perfect agreement and dashed line showing regression line of best fit.
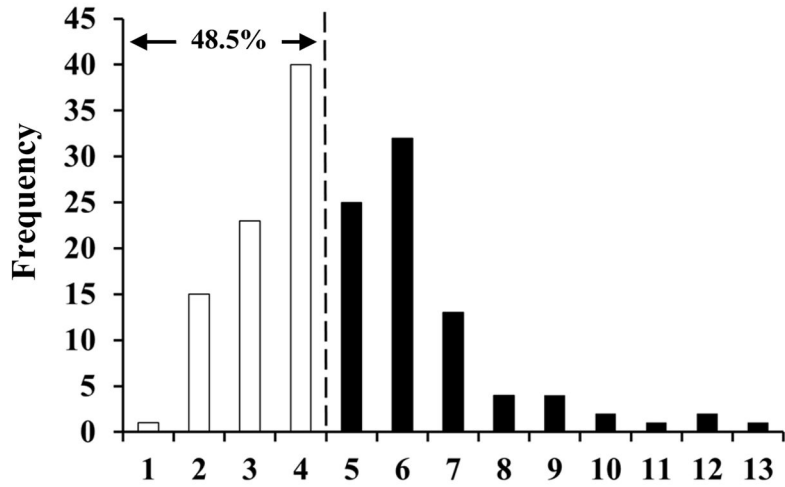
Figure 2.
Frequencies of CO values from each monitor (Vitalograph top, Bedfont bottom) when 24 hour abstinence criterion (<5 ppm) is met under the other monitor.
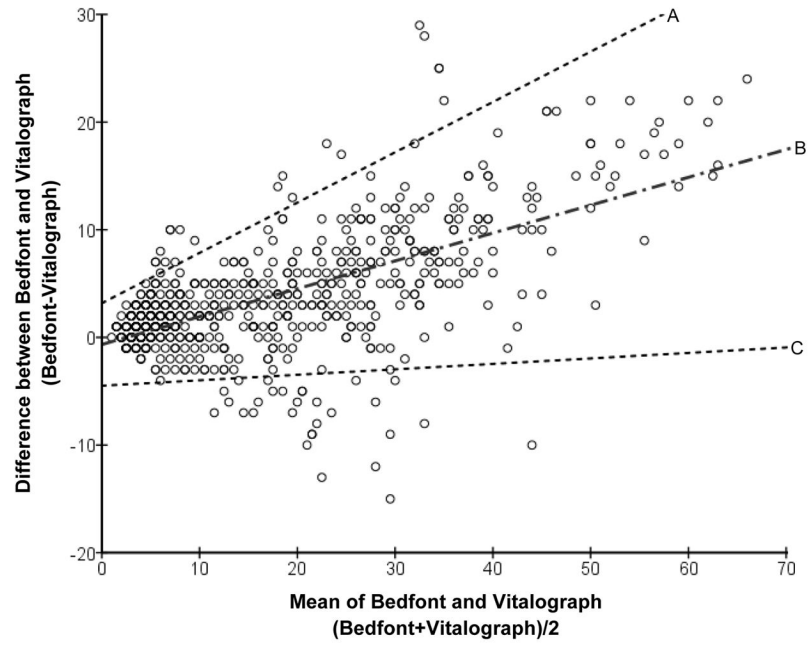
Figure 3.
Bland-Altman plot. Line 'B' is the regression-obtained mean difference between monitors with lines 'A' and 'C' forming the 95% Limits of Agreement.
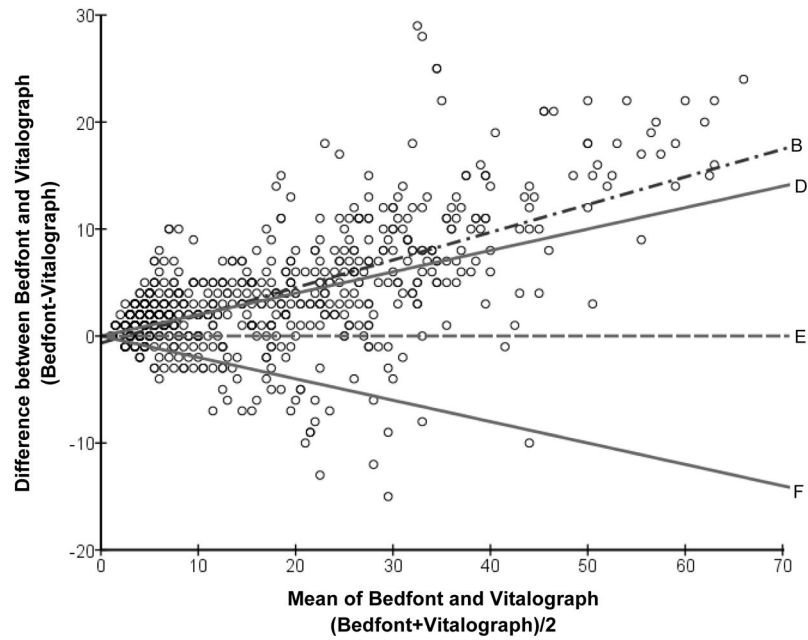
Figure 4.
Bland-Altman plot. Line 'E' illustrates no difference between monitors, with bands forming the upper (line 'D') and lower limits (line 'F') of clinical significance, ±20% of the mean CO value. Line 'B' is the regression-obtained mean difference between monitors.