Presented at the American Thoracic Society International Conference, May 2005, San Diego, CA

Accuracy of Pulse Oximetry During Breath Holding

Scott Sasse[1], Philip Westbrook[2], Daniel Levendowski[2], Tim Zavora[2], Vladislav Velimirovic[2], Cindy Vincent[1], Miodrag Popovic[3]

VA Medical Center Long Beach and UC Irvine[1], Advanced Brain Monitoring[2], University of Belgrade, Serbia and Montenegro[3]

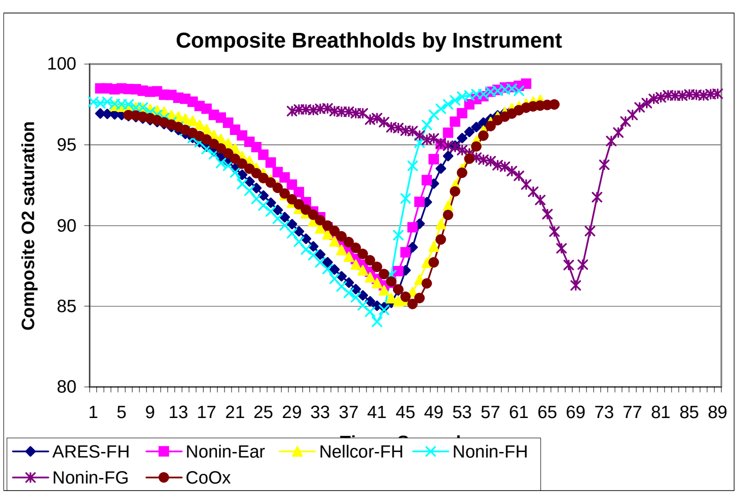
Supported by NIH NHLBI grant HL68463

Rationale: The accuracy of pulse-oximetry is conventionally measured by comparing pulse oximetry to CoOximetry (CoOx) acquired under hypoxic steady-state conditions. This study was designed to measure the accuracy of five different pulse-oximeters during breath holding from functional residual capacity (FRC), a condition that results in dynamic changes more consistent with obstructive breathing during sleep.

Methods: After obtaining approval from the IRB at the Long Beach VAMC, 20 healthy subjects were consented to complete three breath hold (BH) trials from FRC. Supine lung volumes were measured using the nitrogen wash-out technique. Arterial blood was sampled at 5 sec intervals from a radial artery line during and for 20 secs after the end of each breath hold. End-tidal CO2 and respiratory effort were monitored to ensure adequate breath holds. Arterial blood CoOx was measured using a Bayer Model 845 instrument.

Data were simultaneously acquired and synchronized from five oximeters: the ARES Unicorder with its forehead sensor (ARES-FH), Nellcor 595 with its forehead sensor (Nell-FH), the Nonin OEM system with a Nonin forehead sensor (Nonin-FH), the Nonin 8500 with a Nonin ear clip sensor (Nonin-Ear), and the Nonin IPOD with a finger sensor (Nonin-FG).

A cubic spline interpolation was used to predict the time and depth of the nadir in the CoOx data and to estimate data points representative of each one-second period during the breath holds. Once the nadirs from CoOx and the five oximeters were identified, the 40 seconds preceding and the 20 seconds subsequent to the nadirs were extracted and aligned for a total of 61 seconds per breath hold. Second by second comparisons of the data points were made, with summary data submitted to repeated measures ANOVA and t-tests.



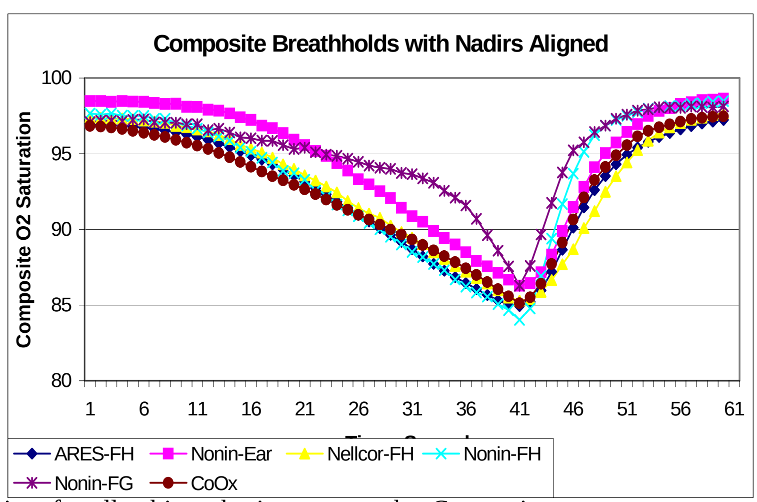


Figure 1.a. Composite oxygen saturation values versus time for all subjects by instrument. b. Composite oxygen saturation values versus time across all subjects by instrument after time alignment, based on the oxygen saturation nadirs.

One subject completed only two (of three) breath hold trials. For the Nonin-FG oximeter, two subjects had missing data for their entire session. The oxygen saturation data were averaged by subject across the available trials to form a composite oxygen saturation curve for each type of oximeter and the CoOx values (see Figures 1 and 2).

Results: To evaluate the goodness of fit of the oximeter breath hold values compared to CoOx, Pearson correlations, the absolute error (negative bias indicates oximeter measured lower O2 saturation compared to CoOx) and root-mean-square errors between CoOx and each of the oximeter measures were computed for each subject across the 61 data points. The results from each subject were then averaged across all subjects and are presented in Table 1. The mean circulation delay (relative to CoOx data) and depth of desaturation are provided to assist with interpretation of Figures 1 and 2.

Table 1: Mean Results Across Subjects

	ARES-FH	Nellcor-FH	Nonin-FH	Nonin-Ear	Nonin-FG	CoOx
Pearson r vs. CoOx	97.2 ± 2.8	96.1 ± 3.1	95.4 ± 3.2	95.0 ± 3.9	79.7 ± 22.3	N/A
Mean Abs. Error (%O2)	0.1 ± 1.7	-0.2 ± 1.5	-0.6 ± 1.0	-1.7 ± 1.8	-2.4 ± 1.7	N/A
RMS Error (% O2)	1.9 ± 0.9	1.8 ± 0.9	2.1 ± 0.7	2.6 ± 1.3	3.4 ± 1.5	N/A
Circulation Delay (secs)	-1.7 ± 2.0	-1.0 ± 2.3	-3.7 ± 2.7	-2.0 ± 2.8	14.5 ± 6.1	N/A
Depth of Desat (% O2)	12.0 ± 2.9	12.1 ± 3.3	13.6 ± 4.6	12.2 ± 4.0	9.8 ± 4.8	11.7 ± 3.5

The root mean square (RMS) error (relative to the CoOx data) for the Nonin-Finger was significantly greater than the RMS errors for the ARES-FH ($p < 0.001$), the Nellcor-FH ($p < 0.001$), and the Nonin-FH ($p < 0.001$) oximeters. The RMS error for the Nonin-Ear was significantly greater than for the ARES-FH ($p < 0.011$) and Nellcor-FH ($p < 0.02$). There were no other significant differences between RMS errors.

Bland-Altman plots were used to evaluate differences between the magnitude of the desaturations measured by CoOx and each of the oximeters (Figure 3).

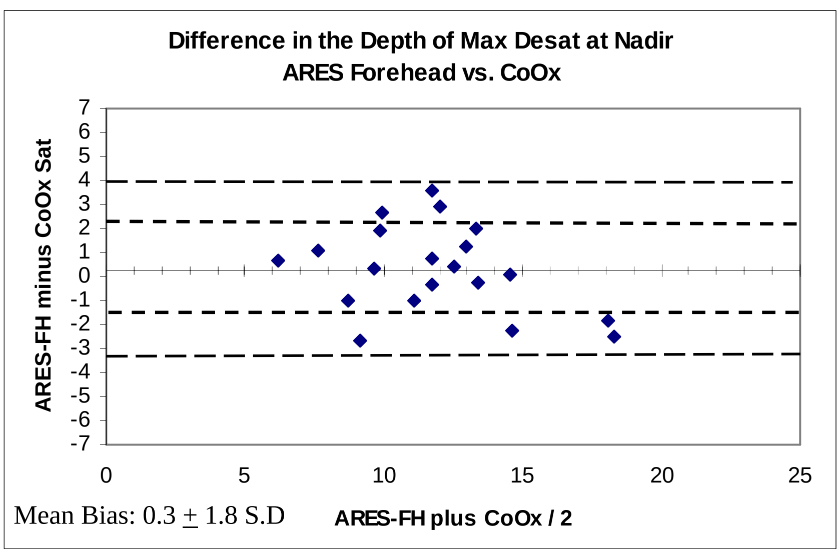


Figure 3: Bland-Altman plots between depth of desaturation measured with CoOx and the five oximeters.

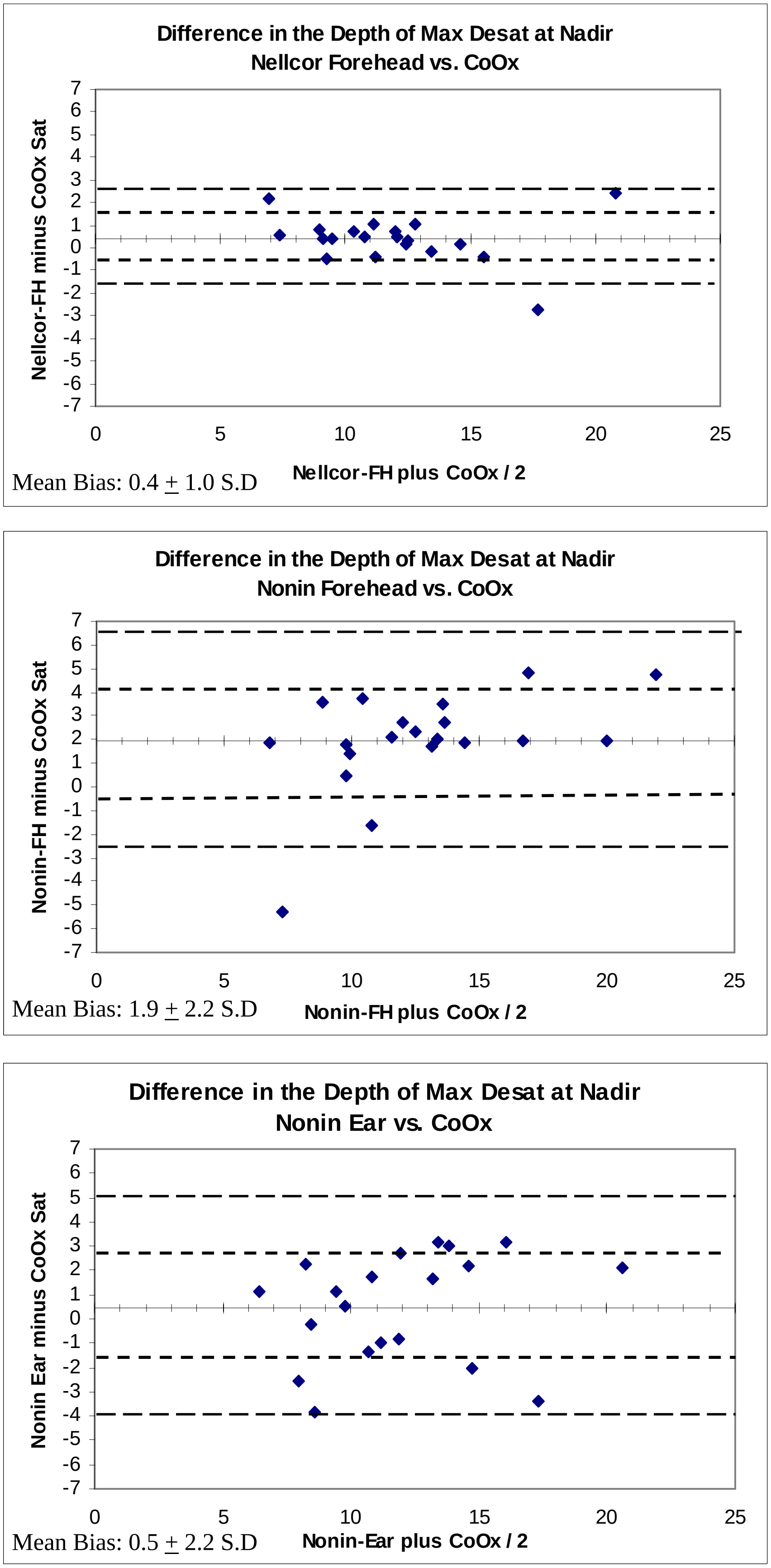


Figure 3 (cont.): Bland-Altman plots between depth of desaturation measured with CoOx and the five oximeters.

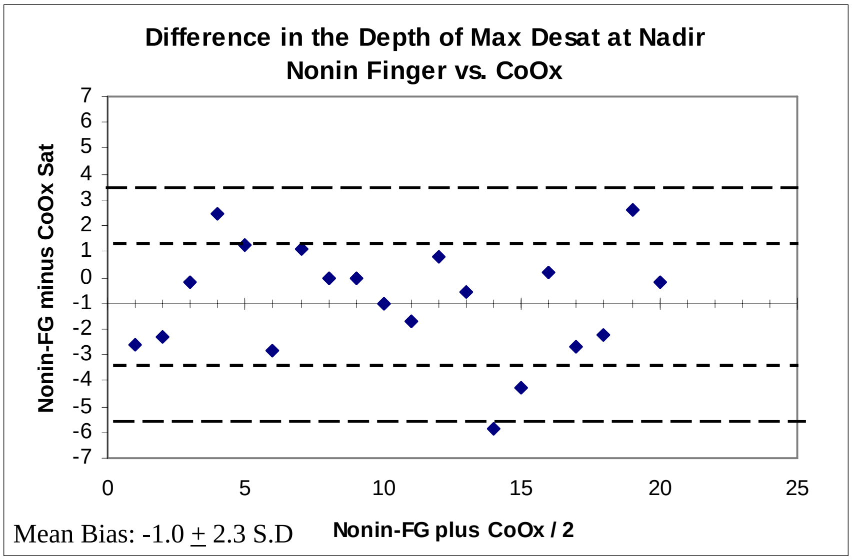


Figure 3 (cont.): Bland-Altman plots between depth of desaturation measured with CoOx and the five oximeters.

A negative correlation existed between the FRC (liters) of each subject and the depth of oxygen desaturation using all oximeters and the CoOx values (Table 2).

Table 2: Correlation between Depth of Desaturation and FRC

	ARES-FH	Nellcor-FH	Nonin-FH	Nonin-Ear	Nonin-FG	CoOx
Pearson r	-0.47	-0.54	-0.63	-0.49	-0.40	-0.55

The breath hold data were stratified into three groups based on the percentage predicted of supine FRC (liters) (% FRC): High FRC = %FRC ≥ 70%, n = 6; Middle FRC = % FRC > 50% and <70%, n = 7; and Low FRC = %FRC < 50%, n = 5.

Table 3: Mean Results Across Subjects

Group	ARES-FH	Nellcor-FH	Nonin-FH	Nonin-Ear	Nonin-FG	CoOx
			Depth of Desaturation			
High FRC	11.2 ± 3.1	11.1 ± 2.0	11.2 ± 4.4	11.2 ± 3.7	8.7 ± 2.4	10.2 ± 2.5
Middle FRC	11.7 ± 2.0	10.7 ± 2.1	12.8 ± 1.9	11.5 ± 2.8	10.7 ± 1.8	10.4 ± 2.2
Low FRC	15.6 ± 1.8	17.1 ± 3.3	20.6 ± 2.8	17.2 ± 3.4	15.9 ± 2.0	17.2 ± 2.5
			RMS Error			**BMI**
High FRC	1.4 ± 0.6	1.6 ± 0.4	1.9 ± 0.7	2.1 ± 1.0	3.2 ± 1.1	24.9 ± 2.7
Middle FRC	1.7 ± 0.6	1.5 ± 0.9	2.0 ± 0.6	2.4 ± 1.0	2.7 ± 1.5	26.3 ± 4.3
Low FRC	2.4 ± 0.5	2.7 ± 0.9	2.8 ± 0.6	2.9 ± 1.1	4.5 ± 1.7	30.6 ± 5.8

Plots of the composite O2 saturation across subjects in each of the three FRC groups versus time are presented in Figures 4, 5 and 6. Given the small sample size, trends were identified using p values < 0.08.

The depth of desaturation for individuals with a low FRC showed a trend towards being greater as measured by ARES-FH (p = 0.056, p = 0.05), Nellcor-FH (p = 0.03, p = 0.05), Nonin-FH (p = 0.02, p = 0.02), and Nonin-Ear (p = 0.07, p = 0.07) relative to the depth of desaturation in subjects with middle range FRCs or high FRCs , respectively (See Table 3).

The BMI was significantly greater for the subjects with a low FRC compared to those with a large FRC ($p = 0.05$). A trend towards a smaller depth of desaturation occurred in the middle FRC group relative to the low FRC group using the Nonin-FG ($p = 0.07$).

The RMS error for the ARES-FH ($p = 0.08$) and Nellcor-FH ($p = 0.07$) showed a trend towards less RMS error in the high FRC group compared to the low FRC group (Table 3). This finding suggests that the RMS error increased in conjunction with the depth of the desaturation. The RMS errors for the other oximeters were similarly distributed among the three groups.

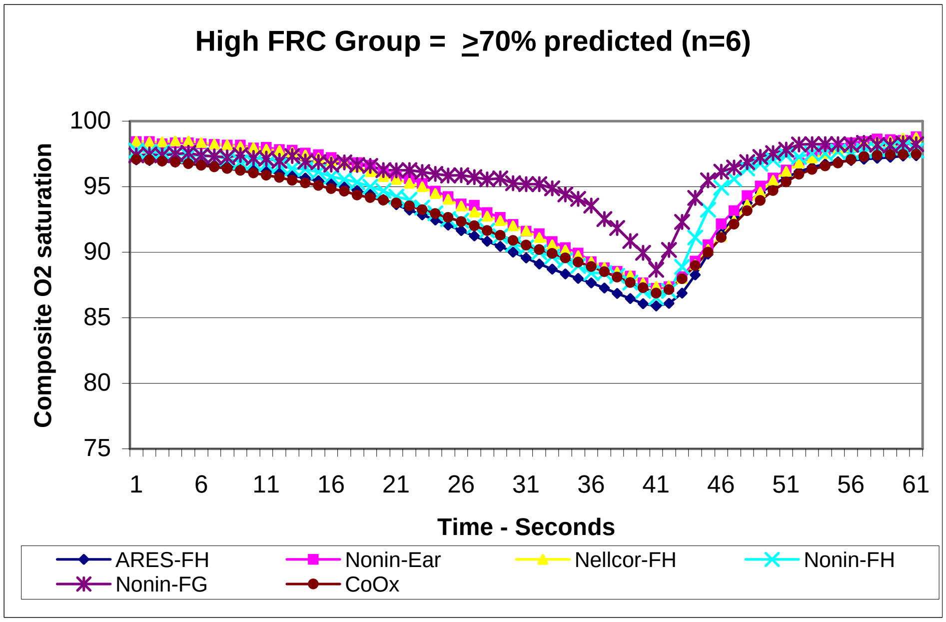


Figure 4: Composite oxygen saturation versus time for Group 1subjects (FRC ≥ 70% predicted).

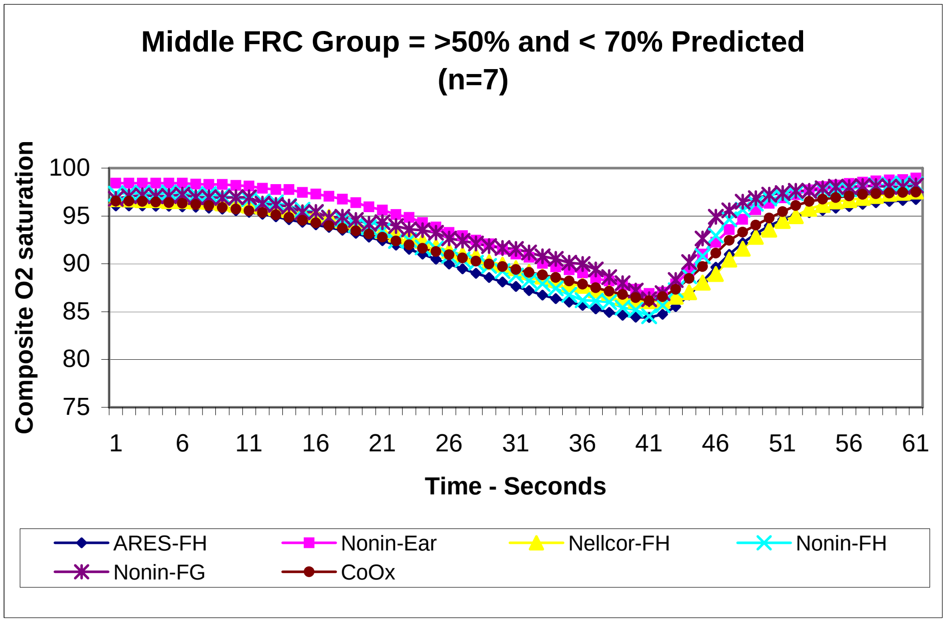


Figure 5: Composite oxygen saturation versus time for Group 2 subjects (50% < FRC <70% predicted).

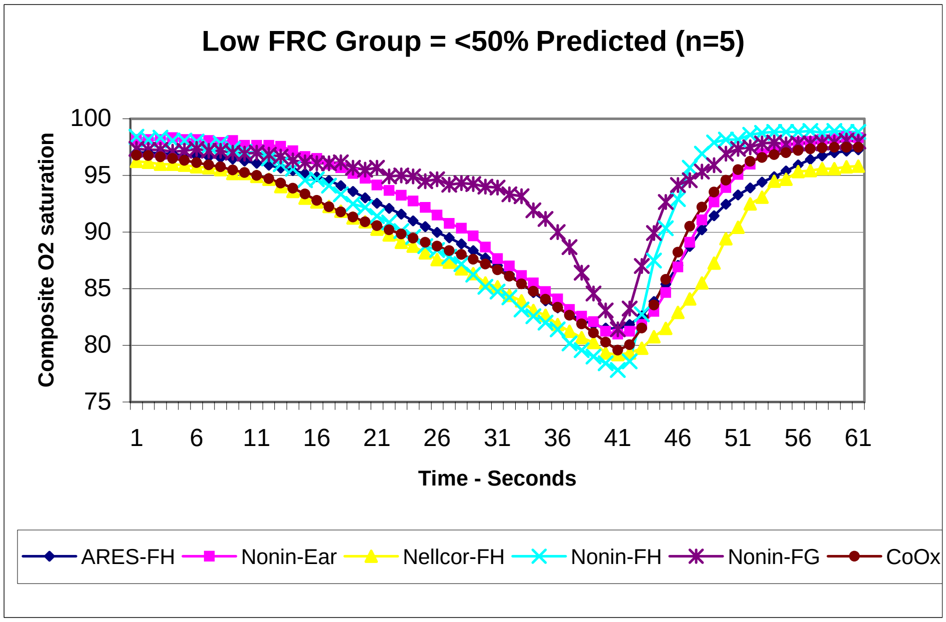


Figure 6: Composite oxygen saturation versus time for Group 3 subjects (FRC ≤ 50% predicted).

Conclusions: In this study in which dynamic changes in oxygen saturation were measured during breath holding, forehead oximetry was more accurate than finger oximetry both in terms of RMS error and depth of desaturation. The ARES and Nellcor forehead oximeters demonstrated lower RMS error relative to ear oximetry. The oxygen saturation nadir point during breath holding was significantly delayed using finger oximetry relative to CoOx. The depth of oxygen desaturation during breath holding was greater in subjects with smaller FRCs. The ARES and Nellcor forehead oximeters appeared to be more accurate when the lowest oxygen saturation did not drop below 85% (high FRC group), and less accurate when the oxygen saturation dropped below 85% (low FRC group). All of the other oximeters exhibited similar accuracy across desaturation ranges.