Pulse Oximeter Accuracy: Multiple-Laboratory Evaluation.
Mannheimer P.D., St. John R.E., Batchelder R., Gehring H. *Anesth Analg* 2007; 105(6).

Introduction

Validating pulse oximeter saturation (SpO2) accuracy requires comparison to the true arterial oxygen saturation (SaO2) across the full specification range. Although methods are standardized 1, information comparing results among multiple laboratories is limited. We evaluated accuracy of five sensor/ monitor configurations at three facilities.

Methods

Independent data collection was performed at Nellcor's performance testing laboratory in Pleasanton, CA; Clinimark Laboratories, Golden CO; and at the University of Lubeck, Lubeck, Germany. Each facility used the same general methodology for testing as described yb ISO 9919 1. After their respective IRB approved protocols and informed consent, pulse oximeter SpO2 readings obtained on healthy adult volunteers were compared to arterial blood SaO2 assessed by CO-Oximetry over the range <70% - 100% SaO2 during normal perfusion and non-motion conditions. Two manufacturers systems were included (Nellcor OxiMax N-600, Nellcor Pleasanton, CA; Masimo Radical, Masimo Corp., Irvine CA). Each system was tested with multiple sensor designs (Table 1) Digit sensor placement was rotated among subjects in a balanced design. Each subject had an indwelling arterial catheter for periodic sampling and was exposed to progressive stepwise hypoxic air/nitrogen mixtures to attain the specified saturation range. Stable SpO2 levels were maintained to ensure tissues at the pulse oximetry sensor site were at the same SaO2 as found at the radial artery sampling site. *Data Analysis:* Computation spans were adjusted for data inclusion to provide a comparable data density over the lower (≤85% and upper (≥85%) SaO2 ranges as suggested in ISO 9919. Accuracy (root mean square of the SpO2 to SaO2 differences, ARMS was determined for each system. Occurrence of SpO2-SaO2 >4% over the range 70%-100% SaO2 was compared with an expected 95% count of observations (consistent with ARMS =2%) using the Fisher's exact test to determine significance ($P<0.05$).

Results

Thirty-seven subjects spanning a range of age, gender, weight and skin pigmentation were studied, with 1259 data pairs collected (Table 2). A, B and C, performed at the ≤2% ARMS level in each California and Germany; Colorado results imply ARMS>2% with C data statistically significant ($P<0.001$). System D and E exceeded 2% ARMS at all three facilities ($P<0.001$) California/Germany; Colorado: P=0.012 (d), E= NS. Accuracy for each system was better in the upper saturations than the lower 70%-80% SaO2 span. System differences were greatest in the lower span, particularly as observed in California and Germany.

Discussion

ARMS differences between systems A, B, C, D and E appear due primarily to bias in the lower span, through magnitude was laboratory dependent. Possible residual differences between laboratory procedures may affect the local biases at the lower saturations. Relative system bias curves were consistent across labs, suggesting ARMS differences may relate to the subjects, subject management and/or blood sampling and analysis. Further investigation is indicated.

Table 1. Tested Oximetry Systems

Monitor	Sensor	System
OxiMax N-600 (v1.1.2.0)	MAX-A adhesive digit sensor	A
	MAX-N adhesive digit sensor	B
	MAX-FAST® forehead sensor	C
Masimo Radical (v4.3.2.1)	LNOP®-Adt adhesive digit sensor	D
	LNOP®-Blue™ adhesive sensor	E

Table 2. Data Summary

	California	Germany	Colorado
Subjects	12	13	12
Computation Span	68%–99%	70%–97%	70%–100%
Stable Samples	227	793	239
≤85% / >85%	113/114	395/398	120/119
System	A_{RMS} (%SpO_2): Overall and ≤85%/>85%		
A	1.9	1.5	2.3
	2.2/1.15	1.6/ 1.3	2.8/1.7
B	1.8	1.6	2.2
	2.1/1.5	1.7/1.5	2.6/1.8
C	1.3	2.0	2.8*
	1.6/1.0	2.5/1.4	3.4/1.9
D	4.6*	3.0*	2.8*
	6.4/1.6	4.0/1.7	3.6/1.7
E	4.1*	2.7*	2.1
	5.6/1.4	3.5/1.4	2.6/1.5

*Observed A_{RMS} greater than 2%, $P < 0.05$.