Transcutaneous Oxygen Saturation Accuracy in Critically Ill Children

Joshua Cushing Brooks (jcushingb@gmail.com)
Pediatric Critical Care Research Group https://orcid.org/0000-0001-5701-8082
Sainath Raman
Lady Cilento Children's Hospital
Kristen Gibbons
Paediatric Critical Care Research Group
Tara Williams
Paediatric Critical Care Research Group
Kimble Robert Dunster
Lady Cilento Children's Hospital
Deborah Pegg
Lady Cilento Children's Hospital
Maria Harkins
Lady Cilento Children's Hospital
Alphonso Nelson
Paediatric Critical Care Research Group
Trang Pham
Paediatric Critical Care Research Group
Andreas Schibler
Lady Cilento Children's Hospital

Research

Keywords: Paediatrics, Critical Care, Pulse oximetry, Oxygen saturation, Cyanotic congenital heart disease, Congenital heart disease, Acute respiratory failure, Septic shock

Posted Date: January 27th, 2020

DOI: https://doi.org/10.21203/rs.2.21938/v1

Abstract

Background

Pulse oximetry (SpO 2) is used to monitor oxygen saturation levels to avoid hypoxaemia in children. Sensor manufacturers claim high sensitivity, specificity and accuracy. Few studies have evaluated accuracy and precision of SpO 2 in children.

Methods

This prospective, observational study was conducted in a 36-bed mixed medical/surgical paediatric intensive care unit. All children <16 years old with an arterial line were eligible. Paired SpO 2 readings obtained with a Masimo and a Nellcor sensor were prospectively matched and validated to the arterial haemoglobin oxygen saturation (SaO 2). Bias between SpO 2 and SaO 2 (SpO 2 -SaO 2), accuracy root mean square (A rms), sensitivity, specificity and kappa agreement were calculated for sensors. Multivariable regression analysis was conducted to determine the relationship between clinical variables and bias in paired sensor readings.

Findings

There were 929 participants with 16,839 readings (9,382 simultaneous Masimo and Nellcor). Nineteen percent of paired values had SaO 2 <88%. Bias increased with decreasing SaO 2 . Both sensors failed to achieve FDA's A rms requirement in all ranges. Of the 15.5% patients with 'true hypoxaemia' (SaO 2 <88%), 28.6% (n=1165) were not correctly identified by pulse oximetry. Variables associated with higher odds of bias included sepsis, respiratory distress and post-cardiac arrest; increasing lactate; vasoconstrictor use; lower SaO 2 and low admission weight. Interpretation Both tested sensors, with current algorithms, are not precise enough for a PICU setting. Sensor readings in patients with respiratory disease, sepsis and cardiac arrest should be used with caution.

Introduction

Pulse oximetry (SpO_2) is a commonly used monitoring tool to assess patient stability, guide emergency airway management, titrate vasoactive medications and fluid management, set fraction of inspired oxygen (FiO_2), positive end-expiratory pressure (PEEP) and supplemental oxygen delivery in hypoxaemic children. Significant clinical decisions in the patient's care are made based on readings of the pulse oximeter, such as oxygen administration, transfer to higher level of care or escalation of therapy. Although arterial haemoglobin oxygen saturation (SaO_2) is considered the 'true gold standard' measure of oxygenation, it is not always available in clinical settings, particularly in infants and children. Therefore, pulse oximetry readings—as the commonly used surrogate measure to arterial oxygen tension (PaO_2)—preferentially need to be accurate. The detection of a true hypoxaemia (defined as $SaO_2 < 88\%$) by pulse oximeter, with a high sensitivity and specificity, is one of the most clinically relevant measures to assess the severity of a critically ill child[1].

As with any recorded variable, an element of bias is expected when comparing to the true value. Bias has been reported between SpO_2 and SaO_2[2,3]. The acceptable bias depends on the clinical setting. As such, a smaller bias (and a higher precision) is required in the critically ill child. Unfortunately, movement artifacts, record-lag, and low perfusion states can reduce measurement reliability.

Numerous commercial pulse oximetry sensors with variable bias and precision are available on the market. The Masimo (Masimo Corporation, Irvine, CA) and Nellcor (Covidien-Nellcor Pty Ltd, Boulder, CO) sensors are the most commonly used in paediatric intensive care. The United States Food and Drug Administration (FDA) requires documented proof of accuracy root mean square (A_{rms}) < 3% with an equal number of samples in the decile ranges of 70–100%[4]. Both manufacturers claim effective sensitivity, specificity, and A_{rms} in clinical settings[5 – 7].

Only a few studies have evaluated the accuracy and precision of SpO_2 in critically ill paediatric patients. Earlier studies indicated that SpO_2 systematically overestimates SaO_2 in paediatric patients, but these studies were limited by small quantities of samples in the hypoxaemic range[3,8,9]. Data obtained in children with cyanotic heart diseases (baseline saturations < 90%) showed that pulse oximetry performed less well in this group [2,3,10].

The primary objective of this study was to assess the performance of pulse oximetry in critically ill paediatric patients in the SpO_2 range of 70–100%. With transcutaneous oxygen saturation measurement by Masimo and Nellcor sensors matched to arterial oxygen saturation measurements, we aimed to quantify the bias between SpO_2 and SaO_2, calculate A_{rms}, and determine the sensitivity and specificity to detect true hypoxaemia. The secondary objective was to perform an exploratory analysis to identify clinical factors that result in higher levels of bias.

Methods

Study Design

This prospective, observational study was conducted in a 36-bed mixed medical and surgical tertiary paediatric intensive care unit (PICU) with approximately 2000 admissions annually. All infants and children (0–16 years of age) with an arterial line in situ were eligible. Patients were excluded if receiving extracorporeal life support. Since pulse oximetry (SpO_2) and arterial blood gas analysis (SaO_2) are standard practice, waiver of formal consent was approved by the hospital institutional review board (HREC/15/QRCH/165 and RD005952). Additionally, assents from parents or patients (where appropriate) were sought before recording data.

Covidien Pty. Ltd. provided the Nellcor modules. The Masimo modules and all the sensors for both modules were sourced from the standard equipment of the hospital. Neither Masimo Corporation nor Covidien Pty Ltd were involved in the design or conduct of the study.

Measurements

Blood gas analysis samples were taken at the clinical discretion of the clinician. The bedside nurses were instructed to place a Masimo and a Nellcor SpO_2 sensor on the same limb as the arterial line. If placed on earlobe, sensor was placed on closest upper extremity. If sensor was placed on the umbilicus, either lower extremity was used. Nurses were trained on sensor placement. Bedside nurses validated the SpO_2 values in the electronic medical record at the time the arterial blood gas (SaO_2) was taken. Only samples processed within five minutes of SaO_2 were accepted for data analysis. Peripheral oxygen saturation for both sensors was recorded with an average over 20 seconds. No measurement was taken if the quality of the SpO_2 trace was low quality based either on the visual signal or, where available, by presence of a low perfusion index (< 2). SaO_2 was measured using co-oximetry (ABL800, Radiometer Medical Aps, Brønshøj, Denmark).

Data Recording and Statistical Analysis

In order to analyse five subgroups (sepsis, respiratory disease, post-cardiac arrest, post-operative, and other conditions), 95 participants per group were required in a multivariable regression analysis to achieve a power of 90%, sample difference of 1% and sigma effect of 6%. The intended patient sample size was approximately 500 patients.

Baseline characteristics are presented using standard descriptive statistics (number [%], mean [SD] or median [interquartile range, IQR]). Fisher's exact test, Student's t-test or the Kruskal-Wallis test were used to compare characteristics between groups.

The accuracy root mean square (A_{rms}) was calculated using the formula:

$$A_{rms} = \sqrt{\sum_{i=1}^{n} \frac{(SpO2i - SaO2i)^2}{n}}.$$

As per FDA guidance, readings were categorized as A_{rms}<3% versus A_{rms}≥3%[4]. Ross et al. reported an overall A_{rms} of 6.5% in the SpO_2 range 65–97% using a sample size of 1980 samples[3].

The bias between SpO_2 and SaO_2 was calculated as (SpO_2-SaO_2). True hypoxaemia (defined as SaO_2 < 88) was cross-tabulated with hypoxaemia using SpO_2 readings; sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and kappa agreement statistic were calculated, along with their 95% confidence intervals (CIs). When presenting summary data, SaO_2 is used as the grouping variable to prevent differing number of observations in Masimo and Nellcor sensors due to different SpO_2 readings.

Bland-Altman plots were created to assess the agreement between SaO_2 and SpO_2 data. Bivariable analysis was conducted (adjusting for related observations for each participant) to determine the relationship between clinical variables and bias (dichotomised as ≤ 3% and > 3%).

Clinical variables investigated were: diagnosis (cardiovascular [CVS]-cyanotic, CVS-acyanotic, sepsis, respiratory, cardiac arrest, post-operative, other), age at ICU admission (years), weight at admission (kg), sex, ethnicity (Aboriginal and/or Torres Strait Islander [ATSI], not ATSI), sensor type (Massimo, Nellcor), SaO_2 category (< 80, 80-<88, 88-<92, ≥ 92%), lactate (continuous variable), total haemoglobin (Hb, continuous), pH (categorical: <7·3, 7·3–7·45, > 7·45), oxygen saturation index ([OSI] = [MAP x FiO_2 x 100]/SpO_2[11,12]) categorised as none (< 5), mild (5-<7.5), moderate (7.5-<12.3) and severe (≥ 12.3), ventilation (yes/no based on PEEP or mean airway pressure [MAP] in patient respiratory record), inotropes (yes/no), vasodilators (yes/no) and vasoconstrictors (yes/no). Inotropes used included phenylephrine, methoxamine, vasopressin, epinephrine and norepinephrine. Vasodilators include clonidine, nitroglycerin, hydralazine and alprostadil. Inotropes include epinephrine, dobutamine, dopamine, isoproterenol, prostaglandins and digoxin.

Multivariable analysis was then undertaken. The confounders were included in the multivariable model if the bivariable p-value was less than 0·25 or they were biologically plausible influencers on the relationship between SpO_2 and SaO_2. Where variables were related, only one was included to meet the model assumptions. Regression estimates or odds ratios are presented along with 95% CIs; the type I error was set at 0·05. Analyses were undertaken in StataSE version 14 (StataCorp Pty. Ltd., College Station, Texas). All parameters recorded are available in the supplemental file.

Results

Baseline Characteristics

Of the 948 recruited participants, 1 (0·1%) had no usable oximetry data and 18 (1·9%) participants were aged 17 or over and therefore met the exclusion criteria. Following exclusion of these participants, data from 929 (97·3%) children remained available for analysis. Of these, 91 (9·8%) had a cyanotic heart disease (right-to-left shunt; pre- and post-operative), 235 (25·3%) acyanotic heart disease (pre- and post-operative), and 603 (64·9%) children were admitted to intensive care for non-cardiac conditions, including 97 (10·4%) with the diagnosis of sepsis, 91 (9·8%) respiratory disease, 38 (4·1%) post-cardiac arrest, 285 (30·7%) post-operative for non-cardiac conditions, and 92 (9·9%) other conditions. The number of measurements differed per patient from a minimum 1 reading to a maximum 681 readings, with a median 11 readings (IQR: 21). There were significant differences in age at ICU admission and weight between these groups (Table 1).

For the 929 participants, there were 25,352 SaO_2 readings; of these, 16,839 (66·4%) had at least one SpO_2 reading (Masimo and/or Nellcor) matched to an arterial SaO_2 measurement. 9,382 (37·0%) of the 25,352 SaO_2 readings simultaneously had both Masimo and Nellcor sensors. The resultant reporting focuses

on the 9,382 SaO_2 paired readings with both a Masimo and Nellcor sensor.

Table 1
Participant demographics compared between diagnostic groups

Variable		All Children (N = 929)		Group												
				CVS – cyanotic (N = 91)		CVS – acyanotic (N = 235)		Sepsis (N = 97)		Respiratory (N = 91)		Post-cardiac arrest (N = 38)		Post-operative (N = 285)		Ot (N
		Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	M
Age at ICU admission (years)		1·9	8·6	0·9	0·8	6·2	6·0	9·8	9·5	2·4	2·3	13·3	12·3	11·6	10·6	12
Weight (kg)		12·0	20·7	4·7	4·7	7·7	15·7	11·8	20·0	7·0	9·1	17·0	40·0	18·0	25·1	20
Gender (n, %)	Male	505	54·4	51	56·0	128	54·5	51	52·6	52	57·1	25	65·8	146	51·2	52
	Female	424	45·6	40	44·0	107	45·5	46	47·4	39	42·9	13	34·2	139	48·8	40
Ethnicity (n, %)	Not ATSI	837	90·1	79	86·8	214	91·1	88	90·7	82	90·1	36	94·7	259	90·9	79
	ATSI	92	9·9	12	13·2	21	8·9	9	9·3	9	9·9	2	5·3	26	9·1	13

Demographics for all children included in the study. CVS, cardiovascular system; ICU intensive care unit; IQR interquartile range; ATSI Aboriginal and/or Torres

Bias

Of the 9,382 paired readings, 19·0% of the values were below $SaO_2 < 88\%$. Masimo and Nellcor median bias varied with SaO_2 category and were inversely proportional to increasing SaO_2 categories (Table 2). There was significant difference between Masimo and Nellcor median bias in paired readings, in the entire range and categories of 88–92% and ≥92% (Table 2). The Bland-Altman plots (Fig. 1) confirmed the inverse agreement between range of saturation and bias. With the exclusion of outliers (where clinician is likely to ignore values as erroneous), the IQR extended out greater than 3% with wide ranges of variability in all categories of SpO_2 readings with the exception of the 92–100% range (Fig. 2). Intra-individual bias for repeated measurements varied (precision), with a decrease with improved saturations (Appendix, Figure A2a-d).

Table 2
Descriptive statistics for bias ($SpO_2 - SaO_2$) for paired readings

SaO_2 Category	Massimo					Nellcor						All				
	N	Min	Max	Med	IQR	N	Min	Max	Med	IQR	p	N	Min	Max	Med	IQR
Entire range	9382	-70	59	1	3·8	9382	-77	59	0·2	3·6	< 0·001	18764	-77	59	1	3·3
$SaO_2 < 70$	261	-19	59	11	12	261	-4·9	59	11·6	13·7	0·129	522	-9	59	11	13·1
$70 \leq SaO_2 < 75$	259	-41	29	5·5	8	259	-10·6	29	5·9	6·1	0·186	518	-41	29	5·9	6·6
$75 \leq SaO_2 < 80$	450	-30	25	4	6	450	-6·1	25	4·4	6·1	0·092	900	-30	25	4·1	6·9
$80 \leq SaO_2 < 88$	817	-35	19	3	7	817	-16·9	20	3	6·6	0·389	1634	-35	20	3	7
$88 \leq SaO_2 < 92$	668	-18	12	4	5	668	-27·8	12	2	5·2	< 0·001	1336	-27·8	12	2·9	5·1
$SaO_2 \geq 92$	6927	-70	8	1	3	6927	-77	8	-0·2	3	< 0·001	13854	-77	8	0	2·5

Accuracy Root Mean Square (A_{rms})

Both sensors failed to achieve FDA's A_{rms} requirement in any of the saturation decile ranges (Table 3) particularly in the range of 90–100% (A_{rms} = 4·5 when both sensors included).

Table 3
A_{rms} measures by SaO_2 category for paired readings (N = 18,764)

SaO_2 Category	Massimo		Nellcor	
	N	A_{rms}	N	A_{rms}
< 70	261	19.0	261	19.3
70–<80	709	8.3	709	8.4
80–<90	1074	6.3	1074	6.5
90–100	7338	3.4	7338	3.6
Total	9382	5·4	9382	5·5

Diagnosis of 'true' hypoxaemia

When comparing hypoxaemia using SaO_2 (gold standard) and SpO_2, true hypoxaemic (SaO_2 < 88%) readings constituted 15.5% (n = 4076) of all paired readings. Amongst this group, 28.6% of readings (n = 1165) were not correctly identified as hypoxaemic by both sensors. As seen in Table 4, sensitivity was 71.4% (95% CI 70.0–72.8%), specificity was 97.6% (95% CI 97.4–97.8%), positive predictive value was 84.5% (95% CI 83.3–85.7%) and negative predictive value was 94.9% (95% CI 94.6–95.2%). Agreement between SaO_2 and SpO_2 was strong (κ = 0·737, p < 0·001; Table 4).

Although both Masimo and Nellcor sensor reliabilities were strong, in the true hypoxic range of SaO_2 < 88%, the sensors missed 30·5% and 26·4% of values in hypoxic patients, respectively (Appendix, Tables A3-A4). Sensitivity expectedly worsened with cut-offs of SaO_2 < 70% and < 80% as the definition of true hypoxaemia (Appendix, Tables A5-A10).

Table 4
False positive rate – SpO_2 hypoxaemia vs. true hypoxaemia

Hypoxaemia based on SpO_2	True Hypoxaemia (SaO_2 < 88)				Total	
	No		Yes			
	n	%	N	%	n	%
No	21702	82·5	1165	4·4	22867	86·9
Yes	533	2·0	2908	11·1	3441	13·1
Total	22256	84·5	4076	15·5	26223	100·0
Fisher's exact test p < 0·001; kappa = 0·737, p < 0·001						

Relationship between bias and clinical characteristics

Bivariable analysis of absolute bias (≤ 3% and > 3%) demonstrated key associations with a high number of clinical variables (Table 5). Of note, cyanotic patients had a higher likelihood of absolute bias over 3% than acyanotic patients (odds ratio [OR] 2·93, 95% CI 2·17 – 3·96, p < 0·001). Children of less than two years of age were more likely to have greater bias than children over two years of age (OR 0·32, 95% CI 0·24–0·43, p < 0·001). Poor peripheral perfusion indicated by increased lactate (elevated lactate: OR 1·33, 95% CI 1·23–1·45, p < 0·001) was significantly associated with greater bias. In patients with respiratory distress, a higher bias was observed (OR 1·97, 95% CI 1·36–2·85, p < 0·001).

In the multivariable model, the variables included were: diagnostic group; weight at admission; sex; oximeter type; SaO_2 category; lactate level; use of ventilation; and use of vasoconstrictors (Table 5). Disease categories of cyanotic heart disease (CVS-cyanotic), post-operative, and other, as well as sex, ethnicity and ventilation were not significant in the model. Variables associated with higher odds of bias > 3% included patients with sepsis, respiratory distress and post-cardiac arrest; lower weight at admission; use of a Nellcor sensor; increasing lactate levels; vasoconstrictor use and lower SaO_2 values of < 80, 80–88 and 88–92%. For completeness, the data for bivariable and multivariable analysis of non-paired measurements are shown in the Appendix (Table A11).

Table 5
Bivariable and multivariable analysis of clinical characteristics and absolute bias categorised as ≤ 3% and > 3% (paired readings only)

Variable		N	Absolute Bias ≤ 3% (N = 13488)		Absolute Bias > 3% (N = 5276)		Bivariable Analysis			Multivariable Analysis (N = 18650)		
			n	%	n	%	OR	95% CI	p	aOR	95% CI	p
Diagnosis	CVS – cyanotic	4380	2801	20.8	1579	29.9	2.93	2.17, 3.96	< 0.001	0.87	0.71, 1.06	0.172
	CVS – acyanotic (ref)	7108	5421	40.2	1687	32.0	1	-	-	1	-	-
	Sepsis	1602	1239	9.2	363	6.9	1.16	0.74, 1.81	0.509	1.40	1.05, 1.86	0.022
	Respiratory	3806	2505	18.6	1301	24.7	1.97	1.36, 2.85	< 0.001	1.45	1.16, 1.82	0.001
	Post-cardiac arrest	460	388	2.9	72	1.4	0.89	0.49, 1.60	0.698	1.62	1.08, 2.43	0.019
	Post-operative	680	472	3.5	208	3.9	0.83	0.49, 1.42	0.503	1.14	0.81, 1.62	0.449
	Other	728	662	4.9	66	1.3	0.34	0.17, 0.68	0.002	1.52	1.05, 2.21	0.027
Age at ICU admission (n, %)	0–2 years (ref)	15114	10442	77.4	4672	88.6	1	-	-	-		
	2–17 years	3650	3046	22.6	604	11.5	0.32	0.24, 0.43	< 0.001	-		
Weight at admission (kg)*		18764	4.8	6.5	4.1	3.8	0.51	0.43, 0.59	< 0.001	0.77	0.70, 0.85	< 0.001
Sex (n, %)	Male (ref)	9144	6362	47.2	2782	52.7	1	-	-	1	-	-
	Female	9620	7126	52.9	2494	47.3	1.00	0.78, 1.30	0.966	1.13	0.98, 1.31	0.095
Ethnicity (n, %)	Not ATSI (ref)	16874	12173	90.3	4701	89.1	1	-	-	1	-	-
	ATSI	1890	1315	9·8	575	10·9	0·94	0·60, 1·48	0·790	1·29	0·99, 1·68	0·055
Oximeter type	Massimo (ref)	9382	7054	52·3	2328	44·1	1	-	-	1	-	-
	Nellcor	9382	6434	47·7	2948	55·9	1·48	1·38, 1·59	< 0·001	3·81	3·42, 4·24	< 0·001
SaO_2 Category	$SaO_2 < 80$	1940	530	3·9	1410	26·7	21·20	17·83, 25·21	< 0·001	4·81	3·60, 6·43	< 0·001
	$80 \leq SaO_2 < 88$	1634	735	5·5	899	17·0	7·68	6·66, 8·87	< 0·001	3·71	2·83, 4·86	< 0·001
	$88 \leq SaO_2 < 92$	1336	674	5·0	662	12·6	4·48	3·93, 5·11	< 0·001	3·20	2·40, 4·28	< 0·001
	$SaO_2 \geq 92$ (ref)	13854	11549	85·6	2305	43·7	1	-	-	1	-	-
Lactate*		18650	1.1	0·7	1.2	0·8	1·33	1·23, 1·45	< 0·001	1·15	1·03, 1·28	0·011
Total Hb^		18742	119·0	21·0	125·4	20·5	1·009	1·007, 1·01	< 0·001	-		
Oxygen saturation index	None (ref)	6087	5100	64·9	987	40·1	1	-	-	-		
	Mild	3042	2145	27·3	897	36·4	2·01	1·75, 2·29	< 0·001	-		
	Moderate	1004	543	6·9	461	18·7	3·33	2·76, 4·02	< 0·001	-		
	Severe	191	72	0·9	119	4·8	6·19	4·25, 9·02	< 0·001	-		
Ventilated	No (ref)	6626	4607	34·2	2019	38·3	1	-	-	1	-	-

OR odds ratio; aOR adjusted odds ratio CI confidence interval; * median (interquartile range), log-transformation used for bivariable analysis; ^ mean (standard deviation)

Variable		N	Absolute Bias ≤ 3% (N = 13488)		Absolute Bias > 3% (N = 5276)		Bivariable Analysis			Multivariable Analysis (N = 18650)		
			n	%	n	%	OR	95% CI	p	aOR	95% CI	p
	Yes	12138	8881	65·8	3257	61·7	0·90	0·82, 0·98	0·013	1·02	0·91, 1·14	0·748
Inotropes	No (ref)	11826	8834	65·5	2992	56·7	1	-	-	-		
	Yes	6938	4654	34·5	2284	43·3	1·23	1·13, 1·34	< 0·001	-		
Vasodilators	No (ref)	10128	7231	53·6	2897	54·9	1	-	-	-		
	Yes	8636	6257	46·4	2379	45·1	0·90	0·83, 0·99	0·028	-		
Vasoconstrictors	No (ref)	15090	11098	82·3	3992	75·7	1	-	-	1	-	-
	Yes	3674	2390	17·7	1284	24·3	1·28	1·15, 1·42	< 0·001	1·23	1·05, 1·44	0·008

OR odds ratio; aOR adjusted odds ratio CI confidence interval; * median (interquartile range), log-transformation used for bivariable analysis; ^ mean (standard deviation)

Discussion

Our study has the largest dataset of prospectively sampled and validated paired samples to date. With an a priori exploratory analysis plan, the study provides insights into the parameters that dictate bias. However, the results cause concern. Both Masimo and Nellcor sensors are not precise enough for the requirements of a paediatric intensive care setting. The low sensitivity of both sensors (71·4%) is alarming since 28·6% of the patient readings (n = 1165) in the true hypoxaemic group (SaO_2 < 88%) were not detected as being hypoxaemia by the saturation sensors. These results corroborate with recent paediatric studies, showing the tendency of saturation sensors to overestimate, particularly at SaO_2 < 88%[2,3]. Given that paediatric patients in our ICU have 15·5% of their total readings of SaO_2 under 80%, this low sensitivity is sub-optimal.

Harris et al. attempted to address sensor precision and bias in paediatric patients with cyanotic congenital heart disease (CCHD). They compared a sensor intended for low saturation scenarios (Masimo Blue) with standard models (Nellcor and Masimo standard) in patients with peripheral saturations under 90%. Simultaneous sensor saturations were recorded with arterial blood gas measurements. They demonstrated that, although the Masimo Blue sensor had better precision and lower bias (especially < 85%), it had limited reliability in patients with CCHD.[2]

In bivariable analysis, the children with cyanotic congenital heart conditions in our study had a higher odds (OR 2·93, 95% CI 2·17–3·96, p < 0·001) of an absolute bias > 3% compared to those with acyanotic cardiac conditions. In a small paediatric CCHD cohort of 19 children (515 paired measurements), Scrimgeour et al. demonstrated a negative correlation between mean bias and SaO_2, wherein pulse oximetry (Masimo) overestimated saturations in 82% of the measurements[13]. Griksaitis et al. investigated a more inclusive cohort of all congenital heart disease patients, recording 527 paired SpO_2 (Masimo) and SaO_2 measurements from 25 patients[10]. They observed poor precision and a large bias, which increased in the < 75% range of SaO_2[10].

We observed in our study—as a measure for accuracy—A_{rms} values greater than 3% in all SaO_2 categories (even SaO_2 > 92%), indicate that Masimo and Nellcor sensors, as they are currently being used in paediatric patients, may require changes to the industrial algorithm to achieve the FDA's bias criteria of A_{rms} <3% in a paediatric population. Achieving these improvements requires further studies of saturation sensors in critical care paediatric populations. We also showed that intra-individual bias (precision) was not constant and the bias decreased with increasing saturation values.

Our study also aimed to describe the disease states and criteria that contribute to sensor bias. Disease categories, such as respiratory disease, sepsis and post-cardiac arrest were shown to be significantly associated with absolute bias > 3% in our model. There is a strong association between increasing severity of respiratory diseases and poor pulse oximetry accuracy[3]. Some researchers suggest that SpO_2/FiO_2 ratios could be employed instead of the PaO_2/FiO_2 ratio in severity prediction of acute respiratory distress syndrome (ARDS) and acute respiratory failure[11,14,15]. The recent paediatric ARDS definitions incorporate SpO_2/FiO_2 ratios in severity stratification. There are two issues with this approach due to the inaccuracy of pulse oximetry found here. First, patients may get incorrectly stratified into a low-risk group and receive less scrutiny compared to if they were in the high-risk group. Furthermore, research studies based on this stratification will yield erroneous results that might then change practice and disadvantage future patients.

The diagnosis of sepsis was associated with high bias, likely explained with the expected low-perfusion state. Similarly, Hummler et al. showed diminished accuracy of saturation sensors in trials on septic animals[16]. Wilson et al. showed that pulse oximetry overestimated SaO_2 compared to SpO_2 by 2·75% in adult patients with severe sepsis and septic shock[17].

Young age (< 2 years) was associated with a greater bias. However, in this age group cyanotic conditions were more prevalent. Within our model, vasoconstrictors led to a noticeable association with absolute bias > 3%, which is likely due to poor prefusion[18].

Clinicians rarely base clinical decisions solely on saturation sensors, but rather the whole clinical picture. However, the sensor bias seen in our study raises concerns over future use of SpO_2 in particular patient groups and clinical scenarios as well as the sensors' ability to estimate SaO_2 for use in the development of artificial intelligence, predictive analytics and machine learning.

Limitations

We aimed to time match the SpO_2 and SaO_2 reading with a high standard of study protocol education, however there are limitations to how precisely the bedside staff recorded and validated the SpO_2 values. The SpO_2 were averaged over 20 seconds. Motion artefacts could have been a systematic error, but we instructed to eliminate these readings from the data recording. Disease categories such as 'other' were used to encompass a large array of clinical diagnoses besides cyanotic and acyanotic cardiovascular patients, septic patients, respiratory patients, post-cardiac arrest patients and post-operative patients. Hence it is difficult to differentiate in which of these 'other' disease states pulse oximetry bias was most affected. Our study aimed to look at five disease subgroups, of which, not all groups were fully filled, thus not achieving the intended power of 90%. In addition, as further analysis of our disease groups occurred, two additional groups were analysed.

Conclusions

Within the limitations of available technology, both Masimo and Nellcor sensors exhibit suboptimal accuracy, particularly in patients with saturations < 88%. Nellcor has slightly better sensitivity when compared to Masimo but is more biased in multivariable modelling. In disease states such as sepsis, respiratory diseases, post-cardiac arrest and vasoconstrictor use, the clinician may have to be circumspect when basing decisions on peripheral oxygen saturations. Ultimately, further improvements need to be made in industrial algorithms to improve sensor accuracy.

Declarations

Ethical Approval and Consent to Participate

This study was reviewed and approved by the Queensland Children's Hospital IRB under HREC/15/QRCH/165 and RD005952. Waiver of formal consent was approved by the IRB, given that pulse oximetry (SpO_2) and arterial blood gas analysis (SaO_2) are standard practice in the PICU. Additionally, assents from parents or patients (where appropriate) were sought before recording data.

Consent for Publication

Not applicable.

Availability of Supporting Data

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests

There are no conflicts of interest to report.

Funding

This study has been supported by a grant from the Intensive Care Foundation, Australia. The funding body contributed financial support for analysis and interpretation of the data.

Author Contributions

JB and AS conceived and designed this study, acquired and interpreted data, and drafted and revised this work. JB, TW and DP helped instruct nurses on validation and recording of sensor readings. KG and SR made significant contributions to interpretation of the data, as well as drafting and revising this work. NA, KRD, TW, MH and DP conceived and designed this study and revised the manuscript. TP contributed to interpretation of the data and revision of the manuscript. All of the authors have given final approval and accountability for the work herein.

Acknowledgements

The authors would like to thank the nurses at Queensland Children's Hospital that supported the Saturation Accuracy in Paediatrics Study. We would like to thank Andrew Barlow, Senior Data Manager at Queensland Children's Hospital for extracting data from the PICU electronic health record.

Author's Information

Not Applicable

References

1. World Health Organisation (WHO). Oxygen therapy for children. 2016 (accessed 9 Dec 2018).

2. Harris BU, Char DS, Feinstein JA, Verma A, Shiboski SC, Ramamoorthy C. Accuracy of Pulse Oximeters Intended for Hypoxemic Pediatric Patients. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2016; **17**(4): 315-20.
3. Ross PA, Newth CJ, Khemani RG. Accuracy of pulse oximetry in children. *Pediatrics* 2014; **133**(1): 22-9.
4. International Organisation for Standardisation (ISO). Particular requirements for basic safety and essential performance of pulse oximeter equipment (ISO 80601-2-562-2011)2013. (accessed 5 March 2019).
5. de-Wahl Granelli A, Wennergren M, Sandberg K, et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39,821 newborns. *BMJ* 2009; **338**: a3037.
6. Masimo. Newborn Sensors. 2017. http://www.masimo.com/siteassets/us/documents/pdf/plm-10199a_product_information_newborn_sensors_us.pdf (accessed 5 Mar 2019).
7. Covidien. NellcorTM Pulse Oximeters with OxiMaxTM Technology from Covidien. 2011. https://smhttp-ssl-51929.nexcesscdn.net/media/attachment/file/1/_/1.1.2.2_pulse_oximeters_with_oximax_technology_brochure.pdf (accessed 5 Mar 2019).
8. Torres A, Jr., Skender KM, Wohrley JD, et al. Pulse oximetry in children with congenital heart disease: effects of cardiopulmonary bypass and cyanosis. *Journal of intensive care medicine* 2004; **19**(4): 229-34.
9. Sedaghat-Yazdi F, Torres A, Jr., Fortuna R, Geiss DM. Pulse oximeter accuracy and precision affected by sensor location in cyanotic children. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2008; **9**(4): 393-7.
10. Griksaitis MJ, Scrimgeour GE, Pappachan JV, Baldock AJ. Accuracy of the Masimo SET(R) LNCS neo peripheral pulse oximeter in cyanotic congenital heart disease. *Cardiology in the young* 2016; **26**(6): 1183-6.
11. Force ADT, Ranieri VM, Rubenfeld GD, et al. Acute respiratory distress syndrome: the Berlin Definition. *Jama* 2012; **307**(23): 2526-33.
12. Thomas NJ, Jouvet P, Willson D. Acute lung injury in children–kids really aren't just "little adults". *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2013; **14**(4): 429-32.
13. Scrimgeour GE, Griksaitis MJ, Pappachan JV, Baldock AJ. The Accuracy of Noninvasive Peripheral Pulse Oximetry After Palliative Cardiac Surgery in Patients With Cyanotic Congenital Heart Disease. *World journal for pediatric & congenital heart surgery* 2017; **8**(1): 32-8.
14. Ray S, Rogers L, Pagel C, Raman S, Peters MJ, Ramnarayan P. PaO2/FIO2 Ratio Derived From the SpO2/FIO2 Ratio to Improve Mortality Prediction Using the Pediatric Index of Mortality-3 Score in Transported Intensive Care Admissions. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2017; **18**(3): e131-e6.
15. Bilan N, Dastranji A, Ghalehgolab Behbahani A. Comparison of the spo2/fio2 ratio and the pao2/fio2 ratio in patients with acute lung injury or acute respiratory distress syndrome. *Journal of cardiovascular and thoracic research* 2015; **7**(1): 28-31.
16. Hummler HD, Engelmann A, Pohlandt F, Hogel J, Franz AR. Decreased accuracy of pulse oximetry measurements during low perfusion caused by sepsis: Is the perfusion index of any value? *Intensive care medicine* 2006; **32**(9): 1428-31.
17. Wilson BJ, Cowan HJ, Lord JA, Zuege DJ, Zygun DA. The accuracy of pulse oximetry in emergency department patients with severe sepsis and septic shock: a retrospective cohort study. *BMC emergency medicine* 2010; **10**: 9.
18. Chan ED, Chan MM, Chan MM. Pulse oximetry: understanding its basic principles facilitates appreciation of its limitations. *Respiratory medicine* 2013; **107**(6): 789-99.

Figures

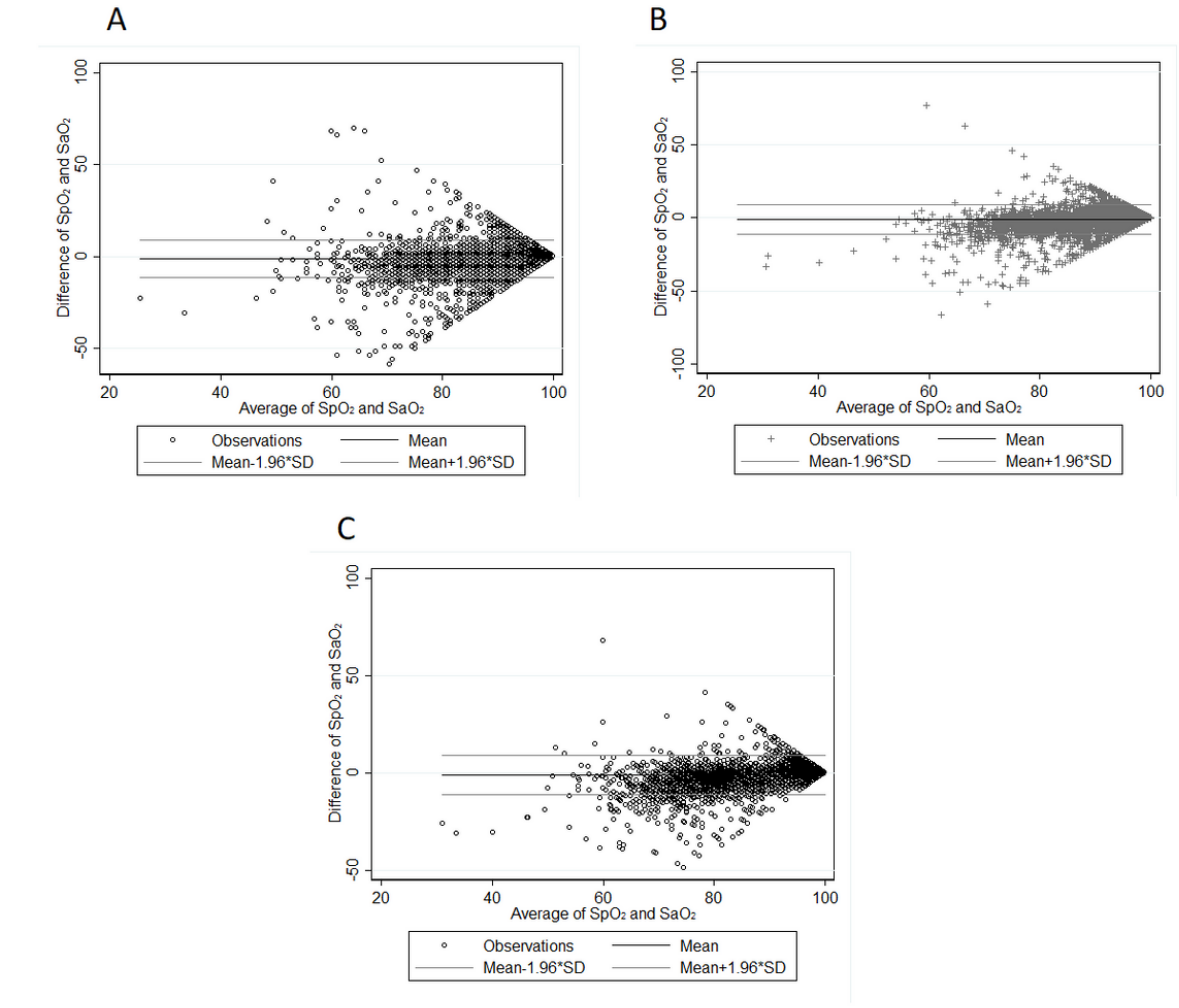

Figure 1

Bland-Altman (BA) Plots paired readings for a) Masimo only, b) Nellcor only, and c) cyanotic children only.

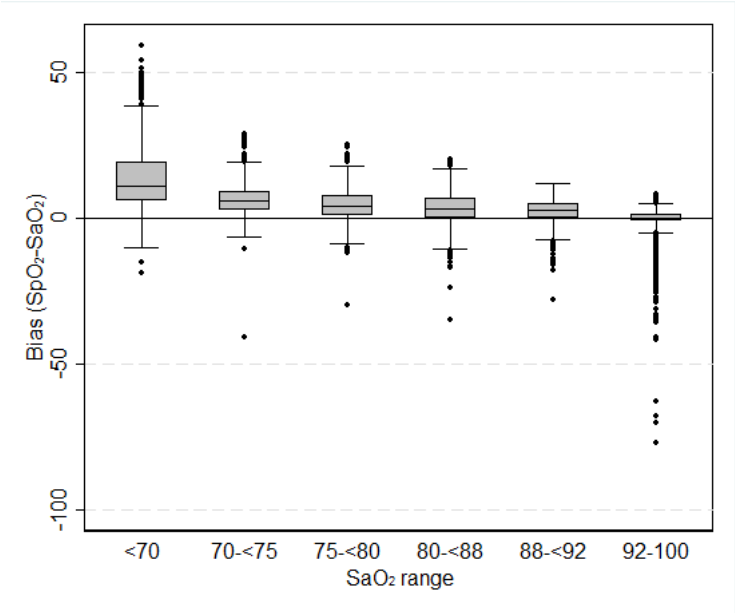

Figure 2

Boxplot of bias versus SaO2 reading for paired measuirements (N=18,764)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SAPTable1supdigcontent.docx
- SAPTable3Supdigcontent.docx
- SAPAppendix.docx
- SAPTable2supdigcontent.docx