Original Article

Pilot Study on the Comparison of Seven Commonly Used Over-the-Counter Pulse Oximeter Finger Devices

Sara Kristel P. Sungahid, MD[1] and Albert L. Rafanan, MD[1]

ABSTRACT

Background: Over-the-counter pulse oximeters are commonly used in patient care but have yet to be evaluated for their accuracy and precision. The objective of the study was to compare the accuracy of seven over-the-counter pulse oximeter finger devices in measuring oxygen saturation (SpO_2), with arterial blood gas oxygen saturation (SaO_2) as the reference value.

Methodology: This was a cross-sectional study from June to November 2023. Pulse oximeter readings of participants aged 18 and above were compared with their SaO_2. Precision was calculated using mean bias and standard deviation while accuracy was assessed through average root mean square error (ARMS). Subgroup analysis of patients with hypoxemia (SaO_2 <95%) was done.

Results: The study included 107 participants. Among the devices, ChoiceMMed MD300C29 had the lowest standard deviation (±1.84) and Yongrow showed the lowest mean bias (0.21), indicating higher precision among the devices studied. ChoiceMMed MD300C29 was the most accurate based on the ARMS score (1.89%). In contrast, IMDK had the highest mean bias (-0.90) and the largest standard deviation (±2.47), indicating lower precision. Hypoxemic subgroup analysis revealed significant biases, lower precision, and lower accuracy for all devices.

Conclusion: ChoiceMMed MD300C29 and Yongrow demonstrate superior precision, with the former emerging as the most accurate. IMDK lags with the lowest precision and accuracy. Subgroup analysis shows less precision and accuracy of readings for hypoxemic patients, cautioning against the reliance on over-the-counter pulse oximeters for accurate SpO_2 measurements. For enhanced accuracy, the study recommends the use of devices with lower mean bias, standard deviation, and ARMS values.

Keywords: pulse oximeter, precision, accuracy

AFFILIATIONS
[1]Section of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chong Hua Hospital, Cebu City

CORRESPONDING AUTHOR
Sara Kristel P. Sungahid, MD
Section of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chong Hua Hospital, Cebu City; sarasungahid44@gmail.com

Paper presented at the 43rd Annual Chest Convention of the Philippine College of Chest Physicians (Poster), EDSA Shangri-La Hotel, Mandaluyong City, March 2024; Chong Hua Hospital Annual Research Contest Presentation (Finalist), 2024

ISSN 3028-1199 (Online)
Printed in the Philippines

DOI: 10.70172/pjcd.v23i1.10201

Received: 2 July 2024
Accepted: 8 April 2025

This article has supplementary materials which are accessible from the article's landing page at www.philippinejournalofchestdiseases.com

INTRODUCTION

A pulse oximeter is a device that uses light beams to estimate the oxygen saturation of the blood and the pulse rate. It is usually placed on the finger. Oxygen saturation gives information about the amount of oxygen carried in the blood, without the need to draw a blood sample.[1] The device provides approximate values for these in vivo parameters and is utilized in many areas of medicine.[2] In clinical practice, monitoring of SpO_2 values (i.e., oxygen saturations measured by pulse oximeter) is required to titrate oxygen therapy to avoid the risks of hypoxemia and hyperoxemia.[3]

Using pulse oximeters is a common practice in hospitals to measure oxygen saturation and pulse rate. With the recent pandemic, its use was not only limited to the hospital but also in individual homes.[1,5] Pulse oximetry has become an indispensable, low-cost, and non-invasive diagnostic tool to assess a patient's oxygen saturation.[4]

Pulse oximetry is utilized from first-patient contact, monitoring, outpatient care and, most importantly, in critical care units. In our hospital, nurses and doctors use their own, personally purchased pulse oximeters to aid in the assessment and monitoring of their patients. These pulse oximeters are widely available and sold in many channels (e.g., drugstores, online sellers) but have yet to be evaluated for their accuracy and validity.

The United States Food and Drug Administration (FDA) distinguishes between prescription use and over-the-counter (OTC) categories of pulse oximeters, with the former undergoing FDA review and clinical testing for accuracy. On the other hand, OTC oximeters, often marketed and used for general wellness or sporting/aviation purposes, may lack FDA clearance for medical use.[1]

In evaluating pulse oximeter accuracy, we must consider factors that may affect measurements such as skin pigmentation, dyshemoglobinemias, severe anemia, low perfusion, and nail polish.[6] A study found that oxygen saturation is often overestimated in individuals with darker skin pigmentation or those identified as Black/African American. Hence, a lower SpO_2 target could lead to occult hypoxemia and worse outcomes for patients with darker skin tones.[7,8]

This validation study compares oxygen saturation measurements from various over-the-counter pulse oximeters (SpO_2), with arterial blood gas oxygen saturation (SaO_2) as reference. The study will use FDA-required performance metrics, including average root mean square error (ARMS; sometimes referred to as accuracy root mean square error) and bias, to evaluate pulse oximeter accuracy.

Objectives

General objective

To compare the accuracy and precision of seven OTC pulse oximeter devices in measuring peripheral oxygen saturation, with arterial oxygen saturation measurement as the standard value.

Specific objectives

1. To determine the mean difference between SpO_2 taken by OTC pulse oximeter finger devices and standardized

SaO_2 measured via arterial blood gas (ABG) analysis
2. To determine the accuracy of OTC pulse oximeter finger devices compared to SaO_2, reported as the average root mean square error (ARMS)
3. To determine the agreement between SpO_2 and standardized SaO_2
4. To evaluate and compare the precision and accuracy of OTC pulse oximeter finger devices in patients with hypoxemia (SaO_2 <95%).

METHODOLOGY

Study design

This was a single-center, cross-sectional study.

Study procedure

Study participants were patients scheduled for ABG testing at Chong Hua Hospital Cebu, Philippines from June 2023 to November 2023.

The inclusion criteria for the study encompassed individuals aged 18 and above, ordered for ABG testing, and possessing an axillary temperature above 35°C, a hemoglobin level greater than 9 g/dL, and a mean arterial pressure exceeding 60 mmHg.

The exclusion criteria included individuals with dyshemoglobinemia, patients with nail polish and/or artificial fingernails, those exhibiting finger clubbing or have peripheral vascular disease, and patients who had received intravenous dyes in the past 7 days.

The sample size was set at 200 data points, in line with the FDA guidelines for accuracy testing.[1] Of the 200 patients scheduled for ABG testing, only 107 met the criteria.

During arterial blood extraction for ABG, the investigators measured the SpO_2 on the patient's finger. An arterial blood volume of 2 cm^3 was obtained from the radial or brachial artery, sent to the laboratory at room temperature, and analyzed within 30 minutes following World Health Organization guidelines. The axillary temperature of the patient was also obtained. During blood extraction, patients were placed in the supine position with the head of the bed elevated to 30 degrees. Other procedures (e.g., suctioning, position change, medication administration) were avoided during this short period.

A single index finger was used to measure SpO_2 values for seven pre-selected pulse oximeters. Each pulse oximeter was placed on the finger for at least 10 seconds until a stable reading was established. All seven readings were taken within 3 minutes from the ABG extraction. A plastic cover was placed on each SpO_2 measurement location to prevent environmental light interference. Measured SpO_2 were recorded in the data sheet. Battery levels of all pulse oximeters were constantly checked; batteries were changed immediately if the indicator showed non-full charge to ensure optimal device performance.

A survey was conducted among hospital staff to identify the top seven pulse oximeter models they frequently use, which were then selected for this study. The pulse oximeters evaluated were: (1) ChoiceMMed MD300C13; (2) ChoiceMMed MD300C29; (3) ChoiceMMed MD300C5; (4) INDOPLAS; (5) Contec CMS50ED; (6) Yongrow YK-83LED; and (7) IMDK C101A3 (Supplementary Material 1). All are readily available for purchase online. Among them, only ChoiceMMed MD300C13 and ChoiceMMed MD300C29 models are FDA-approved. All ChoiceMMed pulse oximeters included in the study also hold ISO (International Organization for Standardization) certifications while the others do not.

Statistical analysis

Comparisons of pulse oximeter readings (SpO_2) with arterial SaO_2 were used to calculate bias, precision (standard deviation [SD] of the bias), mean absolute error, and root mean square error (ARMS). Bias was calculated as the difference between SpO_2 and SaO_2 readings for all patients using each brand of pulse oximeter. The mean bias was determined by averaging these differences across each pulse oximeter brand, with the sign (positive or negative) determining the direction of bias; absolute values were used to determine the magnitude of bias. For each brand, the precision (in %) was computed as the standard deviation of the difference between SpO_2 and SaO_2. The mean absolute error (in %) was calculated as the average absolute difference between SpO_2 and SaO_2. The ARMS (in %) was computed as the square root of the average squared difference between SpO_2 and SaO_2.

Linear regression plots were used to visualize the relationship between SaO_2 and the bias (SpO_2 – SaO_2) of each brand. Bland-Altman plots showed the mean bias and limits of agreement (95% confidence interval). Variability in bias was classified according to Cohen's effect size criteria as low (under 20%), fair (20% to 40%), moderate (40% to 60%), high (60% to 80%), and very high (80% to 100%). To determine the correlation between SaO_2 and bias, multiple R and R-squared tests were used. A multiple R value of less than 0.3, 0.3 to 0.5, and 0.5 to 1 was interpreted as weak, moderate, and strong correlation, respectively.

Hypoxemia was defined as having an SaO_2 of less than 95%. Independent samples t-test was used to compare the bias between patients with and without hypoxemia (<95% vs ≥95%). Levene's test was used to compare the precision between the subgroups. A p-value less than 0.05 was considered significant. Data processing and analysis were done using Microsoft Excel.

Ethical considerations

The study was conducted in adherence to the Declaration of Helsinki and approved by the Institutional Review Board (IRB)/Ethics Committee of Chong Hua Hospital (IRB reference code 3123-04). Verbal consent was obtained prior to the procedure, as pulse oximeter monitoring is recognized as an acceptable standard of care.

RESULTS

Table 1 shows the demographic and clinical characteristics of the study participants (n = 107). The mean age was 63 years, ranging from 32 to 92 years. The majority of participants were male (78%) and most identified as Asian (95%). The mean hemoglobin level was 11.41 g/dL, with values ranging from 9.1 to 14.9 g/dL.

Table 2 shows the accuracy of OTC pulse oximeter finger devices compared to SaO_2, reported as the mean bias (SD) and ARMS.

Among the seven brands, Yongrow had the lowest mean bias (0.21) and ChoiceMMed MD300C29 had the lowest standard deviation (±1.84). The latter also had the lowest ARMS score (1.89%). In contrast, IMDK, INDOPLAS, and Yongrow had the highest ARMS values.

Table 1. Participant characteristics (n = 107)

Characteristics	
Age, years	
Mean (Min to Max)	63 (32 to 92)
Sex, n (%)	
Male	83 (78)
Female	24 (22)
Race, n (%)	
Asian	102 (95)
White	5 (5)
Hemoglobin, g/dL Mean (Min to Max)	11.41 (9.1 to 14.9)

Table 2. Performance of pulse oximeters compared to arterial blood gas

Pulse oximeter	Mean bias (SD)	ARMS, %
ChoiceMMed MD300C13 (Beijing, China)	0.53 (1.91)	1.98
ChoiceMMed MD300C29 (Beijing, China)	0.49 (1.84)	1.89
ChoiceMMed MD300C5 (Beijing, China)	0.60 (1.95)	2.03
INDOPLAS (Quezon, Philippines)	-0.70 (2.17)	2.27
Contec (Hebei, China)	-0.33 (2.12)	2.14
Yongrow (Xuzhou, China)	0.21 (2.18)	2.18
IMDK (Guangdong, China)	-0.90 (2.47)	2.61

SD, standard deviation; ARMS, average root mean square error

Figures 1.1 to 1.7 present the relationship between SaO_2 levels and bias (SpO_2 – SaO_2, in %) for the different pulse oximeter devices. Each figure includes a linear regression plot illustrating how bias changes with varying SaO_2 levels, alongside Bland-Altman plots displaying mean bias and its limits of agreement. All devices showed a statistically significant relationship between SaO_2 and bias ($p < 0.05$; values not shown), with bias generally decreasing as SaO_2 increased, as indicated by the downward slopes and negative regression coefficients.

ChoiceMMed MD300C13 and ChoiceMMed MD300C29 exhibited fair variability in bias (25.2% and 28.3%, respectively) with strong correlations (multiple R = 0.506 and 0.532, respectively). ChoiceMMed MD300C5 demonstrated moderate variability (37.7%) with strong correlations (multiple R = 0.614). INDOPLAS displayed low variability (17.9%) with moderate correlation (multiple R = 0.423), while CONTEC showed fair variability (21.7%) with moderate correlation (multiple R = 0.466). Yongrow exhibited fair variability (22%) and moderate correlation (multiple R = 0.469). IMDK had poor variability (7.5%) and weak correlation (multiple R = 0.274).

The strength of the correlation varied across devices, with ChoiceMMed MD300C5 showing the strongest correlation (R = 0.614) followed by ChoiceMMed MD300C29 (R = 0.532) and ChoiceMMed MD300C13 (R = 0.506), while IMDK showed the weakest correlation (R = 0.274). The percentage of variability in bias also differed, indicating differences in the consistency of bias across different devices.

Table 3 provides a detailed breakdown of mean bias, precision, and ARMS values across the different devices, for hypoxemic and non-hypoxemic patients.

The majority of patients (81%) had SaO_2 of 95% or greater. Mean bias was significantly greater among patients with hypoxemia compared to patients without hypoxemia. This was noted across all pulse oximeters except IMDK. Precision values were also significantly higher among patients with hypoxemia for all devices, with precision values close to zero being more precise. As to the ARMS of hypoxemic patients, their scores were noted to be higher.

DISCUSSION

In evaluating the performance of each pulse oximeter device, the mean bias was used to represent the average difference between the measured SpO_2 values and the standardized SaO_2. Notably, Yongrow demonstrated the lowest mean bias, suggesting better precision. However, it had a wide standard deviation compared to other devices. Conversely, ChoiceMMed MD300C29 exhibited the lowest standard deviation, reinforcing its precision. Standard deviation provides a measure of the variability or spread of the bias scores, with lower SD indicating less variability and higher SD suggesting greater variability in bias. Both the mean bias and standard deviation should be considered in assessing precision.

Conversely, IMDK exhibited the highest mean bias and a larger standard deviation, indicating lower precision among the evaluated devices. Positive mean biases were observed with the three ChoiceMMed devices and Yongrow, implying a tendency to slightly overestimate SaO_2. In contrast, INDOPLAS, Contec, and IMDK exhibited negative mean biases, suggesting a tendency to slightly underestimate SaO_2.

To assess the overall accuracy of the pulse oximeter devices, ARMS values were computed, with lower values signifying better accuracy. All ChoiceMMed devices demonstrated relatively lower ARMS values, indicating better overall accuracy compared to other devices. Conversely, IMDK recorded the highest ARMS value, indicating a higher level of error in SaO_2 estimation. Nevertheless, ARMS values for all devices were less than 3%, or the acceptable performance range set by the FDA.

The correlation between SpO_2 values from OTC pulse oximeters and standardized SaO_2 displayed variability among different devices. The ChoiceMMed devices exhibited strong correlations while Indoplas, Contec, and Yongrow demonstrated moderate correlations, and IMD showed weak correlation.

Based on the linear regression, ChoiceMMed MD300C5 showed a stronger correlation between SaO_2 and bias, and a more

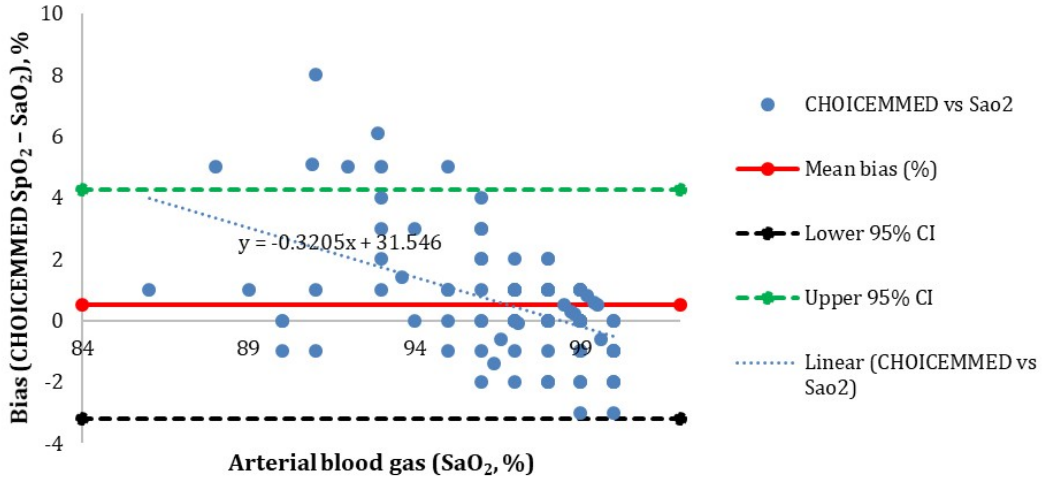

Figure 1.1. Linear regression and Bland-Altman plot of SaO2 vs bias for ChoiceMMed MD300C13

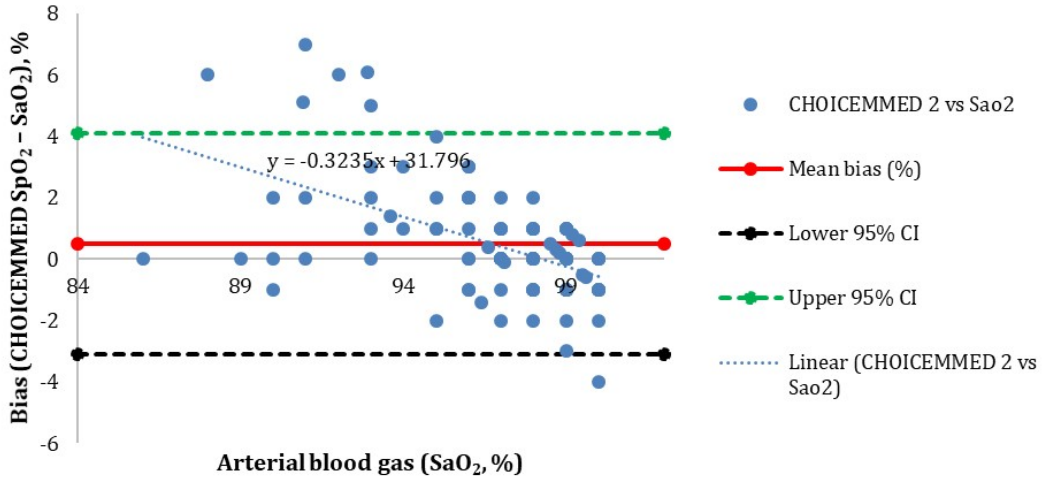

Figure 1.2. Linear regression and Bland-Altman plot of SaO2 vs bias for ChoiceMMed MD300C29

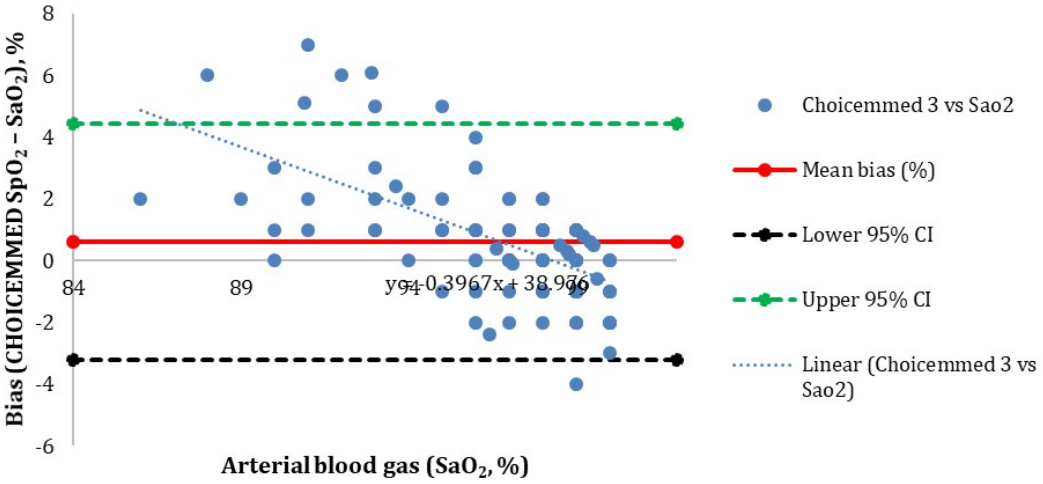

Figure 1.3. Linear regression and Bland-Altman plot of SaO2 vs bias for ChoiceMMed MD300C5

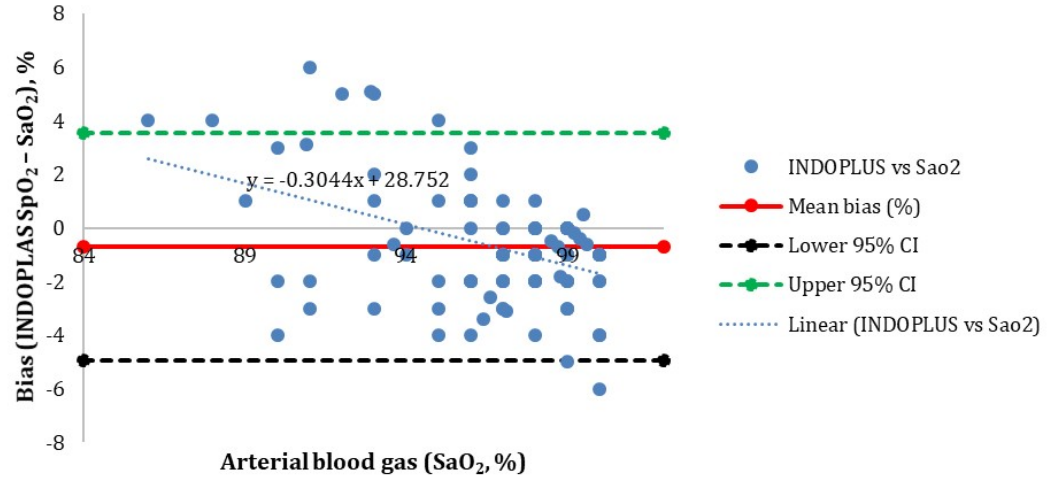

Figure 1.4. Linear regression and Bland-Altman plot of SaO2 vs bias for INDOPLAS

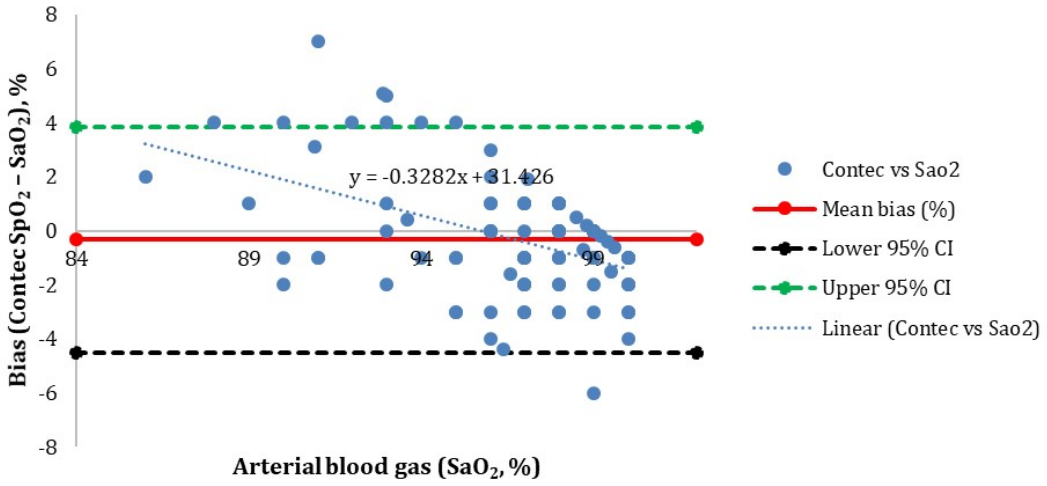

Figure 1.5. Linear regression and Bland-Altman plot of SaO2 vs bias for Contec

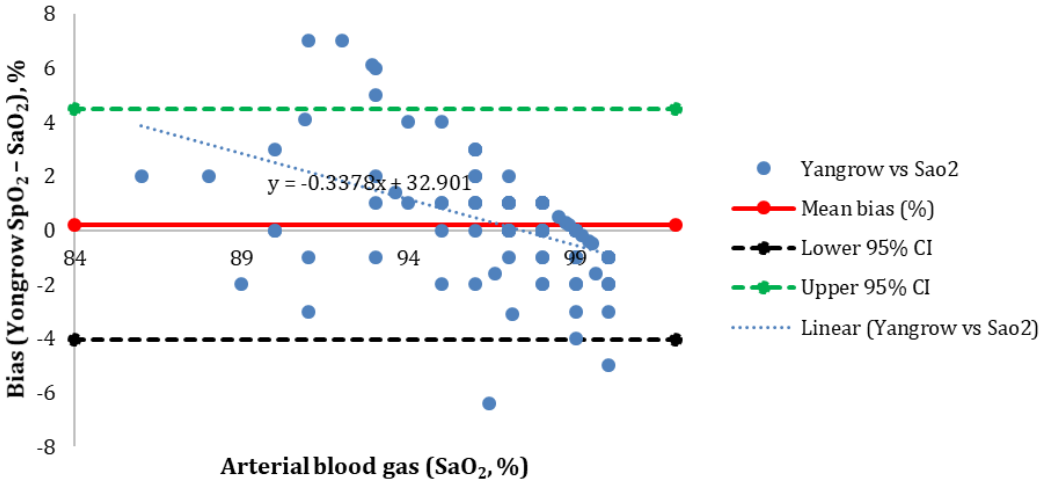

Figure 1.6. Linear regression and Bland-Altman plot of SaO2 vs bias for Yongrow

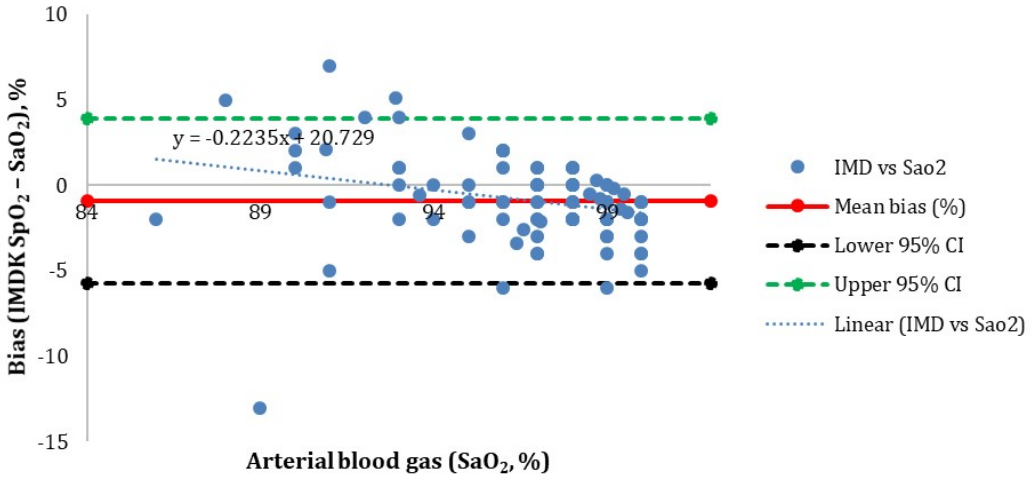

Figure 1.7. Linear regression and Bland-Altman plot of SaO2 vs bias for IMDK

Table 3. Comparison of mean bias, precision, and ARMS in patients with hypoxemia (SaO_2 <95%) and without hypoxemia (SaO_2 ≥95%)

		With hypoxemia (n = 20)	Without hypoxemia (n = 87)	Overall	p-value
ChoiceMMed MD300C13	Mean bias	2.48	0.08	0.53	0.001
	Precision	2.57	1.40	1.91	<0.001
	ARMS (%)	3.52	1.40	1.98	…
ChoiceMMed MD300C29	Mean bias	2.48	0.04	0.49	<0.001
	Precision	2.53	1.27	1.84	<0.001
	ARMS (%)	3.49	1.26	1.89	…
ChoiceMMed MD300C5	Mean bias	2.88	0.07	0.60	<0.001
	Precision	2.19	1.46	1.95	0.004
	ARMS (%)	3.58	1.45	2.03	…
INDOPLAS	Mean bias	1.13	-1.12	-0.70	0.005
	Precision	3.14	1.63	2.17	<0.001
	ARMS (%)	3.26	1.98	2.27	…
Contec	Mean bias	1.83	-0.83	-0.33	<0.001
	Precision	2.72	1.61	2.12	<0.001
	ARMS (%)	3.22	1.80	2.14	…
Yongrow	Mean bias	2.23	-0.25	0.21	0.002
	Precision	2.99	1.64	2.18	<0.001
	ARMS (%)	3.67	1.65	2.18	…
IMDK	Mean bias	0.48	-1.22	-0.90	0.100
	Precision	4.32	1.68	2.47	<0.001
	ARMS (%)	4.24	2.07	2.61	…

significant decrease in bias with increasing SaO_2, suggesting that it may perform best only when the SaO2 is sufficiently high.

Absolute accuracy measures (mean bias and ARMS) provide a more comprehensive overview of the device's performance across all conditions. These metrics are more relevant for determining the overall accuracy of the device's readings. Thus, while ChoiceMMed MD300C5 showed a better linear relationship between SaO_2 and bias, ChoiceMMed MD300C29 demonstrated superior overall accuracy and consistency when considering absolute measures of performance. This explains why ChoiceMMed MD300C29 might be considered better in terms of mean bias and ARMS, even if ChoiceMMed MD300C5 appeared more sensitive to SaO_2 variations. However, it should be noted that the difference between the two might not be significant, which is a limitation in this paper since pairwise comparison of the bias for each pulse oximeter was not done.

The subgroup analysis for hypoxemic and non-hypoxemic patients revealed several important findings. Firstly, the majority of patients had SaO_2 levels of 95% or greater. Secondly, there was a significant increase in bias among patients with SaO_2 less than 95% across all pulse oximeters except IMDK. Thirdly, precision was notably large, which indicates less precision in patients with SaO_2 less than 95% for all pulse oximeters. ARMS values in the hypoxemic subgroup also exceeded the 3% threshold. This indicates that the pulse oximeters, while meeting general accuracy standards, may show decreased precision and accuracy in the context of hypoxemia.

Limitations

The study was a single-center pilot study that used convenience sampling and involved a sole data collector. This may have resulted in a single-observer bias and an increased risk of human error in data collection. Additionally, the restricted variability in SaO_2, with the study population mostly having SaO_2 exceeding 95%, hindered further analysis of patients with lower SaO_2. Potential moment-to-moment differences in pulse oximeter readings during sequential device placement were acknowledged, prompting the randomization of device measurement sequences to minimize this effect. The population size was not able to meet the desired data points set by FDA, which may have resulted in an underpowered analysis. Lastly, the study lacked blinding, as an investigator applied the devices to the patient's finger, while serving as the sole data collector.

Recommendations

It is strongly recommended that future investigations be conducted through a larger, multi-center study to enhance the generalizability and robustness of the results. The inclusion of an FDA-approved medical-grade pulse oximeter as a comparator is advised to establish a more comprehensive benchmark for performance evaluation.

To refine accuracy assessments, it is suggested that SaO_2 be subcategorized into specific ranges (less than 70%, 70% to 80%, 80% to 90%, and 90% to 100%) with equal populations. This subcategorization could offer a nuanced understanding of each device's accuracy across different SaO_2 ranges. However, if not conducted in a controlled experimental environment, a larger population size may be necessary to ensure statistical validity.

To streamline the study and minimize potential confounding factors, it is advised to limit the number of oximeters to less than five. This limitation aims to reduce the time differences in measurements for each device, thereby enhancing reliability of the comparative analyses.

CONCLUSION

This study provides a comprehensive evaluation of OTC pulse oximeter devices by comparing their performance—through mean bias, precision, and accuracy—to standardized SaO_2 measurements. Devices with lower mean bias and smaller standard deviations are more precise, and lower ARMS values are more accurate. ChoiceMMed MD300C29 and Yongrow demonstrate superior precision, with the former emerging as the most accurate. IMDK lags with the lowest precision and accuracy.

The correlation between SpO_2 values from OTC pulse oximeters and standardized SaO_2 displayed variability among different models which underscore the diverse performance characteristics and associations among the evaluated pulse oximeter models.

In states of hypoxemia, precision and accuracy of devices were lower. These findings underscore the importance of considering oxygen saturation levels when assessing the performance of pulse oximeter.

Clinicians and users should consider device-specific characteristics and limitations when interpreting SpO_2 values. Further research is warranted to validate these findings in a larger population group, with variable SaO_2 measurements, and in specific clinical conditions.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Authors' Disclosure

The authors declared no conflict of interest.

Funding Source

The authors did not receive any funding related to the research, authorship, and/or publication of the article.

References

1. U.S. Food and Drug Administration. Performance evaluation of pulse oximeters taking into consideration skin pigmentation, race and ethnicity. Prepared for the February 2, 2024, meeting of the Anesthesiology and Respiratory Therapy Devices Panel of the Medical Devices Advisory Committee. Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration; 2024:1,3,30. Accessed April 28, 2025. https://www.fda.gov/news-events/press-announcements/fda-proposes-updated-recommendations-help-improve-performance-pulse-oximeters-across-skin-tones
2. International Organization for Standardization. ISO 9919:2005(E) Medical electrical equipment—Particular requirements for the basic safety and essential performance of pulse oximeter equipment for medical use. Geneva: ISO; 2005: viii.
3. Pilcher J, Ploen L, McKinstry S, Bardsley G, Chien J, Howard L, et al. A multicentre prospective observational study comparing arterial blood gas values to those obtained by pulse oximeters used in adult patients attending Australian and New Zealand hospitals. *BMC Pulm Med*. 2020;20(1):2. https://doi.org/10.1186/s12890-019-1007-3
4. Harskamp RE, van der Laan AM, van Goor H, et al. Performance of popular pulse oximeters compared with simultaneous arterial oxygen saturation or clinical-grade pulse oximetry: A cross-sectional validation study in intensive care patients. *BMJ Open Respir Res*. 2021;8(1):e000939. https://doi.org/10.1136/bmjresp-2021-000939
5. Schrading WA, McCafferty B, Grove J, Page DB. Portable, consumer-grade pulse oximeters are accurate for home and medical use: implications for use in the COVID-19 pandemic and other resource-limited environments. *JACEP Open*. 2020;1(6):1450-1458. https://doi.org/10.1002/emp2.12292
6. U.S. Food and Drug Administration. Pulse oximeter accuracy and limitations: FDA safety communication. 2022. Accessed April 28, 2025. https://www.fda.gov/media/162709/download
7. Shi C, Goodall M, Dumville J, Hill J, Norman G, Hamer O, et al. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis. *BMC Med*. 2022;20(1):267. https://doi.org/10.1186/s12916-022-02452-8
8. Herbst A, Goel S, Beane A, Brotherton BJ, Dula D, Ely EW, et al. Oxygen saturation targets for adults with acute hypoxemia in low and lower-middle income countries: a scoping review with analysis of