Original Article

The effect of patient ethnicity on the accuracy of peripheral pulse oximetry in patients with COVID-19 pneumonitis: a single-centre, retrospective analysis

M. D. Wiles,^{1,2} (D A. El-Nayal,³ G. Elton,⁴ M. Malaj,³ (D J. Winterbottom,⁴ (D C. Gillies,⁵ I. K. Moppett⁶ (D and K. Bauchmuller¹ (D

1 Consultant, 3 Registrar, 4 Foundation Year-2 Trainee, 5 Charge Nurse, Department of Critical Care, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

2 Honorary Clinical Lecturer, University of Sheffield Medical School, Sheffield, UK

6 Professor, Department of Anaesthesia and Peri-operative Medicine, University of Nottingham, Nottingham, UK

Summary

Pulse oximetry is used widely to titrate oxygen therapy and for triage in patients who are critically ill. However, there are concerns regarding the accuracy of pulse oximetry in patients with COVID-19 pneumonitis and in patients who have a greater degree of skin pigmentation. We aimed to determine the impact of patient ethnicity on the accuracy of peripheral pulse oximetry in patients who were critically ill with COVID-19 pneumonitis by conducting a retrospective observational study comparing paired measurements of arterial oxygen saturation measured by co-oximetry on arterial blood gas analysis (SaO₂) and the corresponding peripheral oxygenation saturation measured by pulse oximetry (S_pO₂). Bias was calculated as the mean difference between SaO₂ and S_pO₂ measurements and limits of agreement were calculated as bias ± 1.96 SD. Data from 194 patients (135 White ethnic origin, 34 Asian ethnic origin, 19 Black ethnic origin and 6 other ethnic origin) were analysed consisting of 6216 paired SaO₂ and S_pO₂ measurements. Bias (limits of agreement) between SaO₂ and S_pO₂ measurements was 0.05% (-2.21–2.30). Patient ethnicity din ot alter this to a clinically significant degree: 0.28% (1.79–2.35), -0.33% (-2.47–2.35) and -0.75% (-3.47–1.97) for patients of White, Asian and Black ethnic origin, respectively. In patients with COVID-19 pneumonitis, S_pO₂ measurements showed a level of agreement with SaO₂ values that was in line with previous work, and this was not affected by patient ethnicity.

Correspondence to: M. D. Wiles Email: matthew.wiles1@nhs.net Accepted: 16 August 2021 Keywords: COVID-19 pneumonitis; critical care; ethnicity; pulse oximetry; skin pigmentation Twitters: @STHJournalClub; @coralpinkness; @JakeDMW96; @lainMoppett; @bauchmueller

Introduction

Pulse oximetry is used widely to titrate oxygen therapy in patients who are critically ill. As pulse oximetry is inexpensive, non-invasive and can be used continuously, it provides a convenient way to monitor the oxygenation of patients via indirect measurement of the saturation of arterial blood. The US Food and Drug Administration recommends that pulse oximetry devices should be validated by directly comparing values with oxygen saturations measured by co-oximetry of an arterial blood sample [1]. Most pulse oximeter validation studies have been performed using healthy volunteers; those that have involved patients who are critically ill have shown variable agreement with arterial blood saturations measured by co-oximetry [2–4].

Pulse oximetry is also used as a triage tool in COVID-19 pneumonitis. However, there are concerns regarding the accuracy in this disease [5] and in patients who have a greater degree of skin pigmentation [6–8]. This is important as patients from non-White ethnic backgrounds have been disproportionally affected by COVID-19 infection, with a greater risk of critical illness as a result [9, 10]. We aimed to determine the impact of patient ethnicity on the accuracy of peripheral pulse oximetry in patients who were critically ill with COVID-19 pneumonitis.

Methods

We conducted a retrospective, observational study of arterial blood oxygen saturation measurement in patients with confirmed COVID-19 pneumonitis who had been admitted to our tertiary general critical care unit with hypoxaemic respiratory failure. The purpose of the study was to quality-assure local practice with respect to saturation targets, due to concerns that peripheral oxygen saturations may be under-estimating the degree of hypoxaemia. The study was registered prospectively with the local clinical effectiveness unit as a service review and was deemed not to require formal ethical approval. Patients were identified by cross-referencing the critical care clinical information system (MetaVision ICU; iMDsoft[®], Tel Aviv, Israel) and the local Intensive Care National Audit and Research Centre (ICNARC) case-mix programme database. COVID-19 pneumonitis was diagnosed by the combination of positive SARS-CoV-19 polymerase chain reaction test, hypoxaemia requiring artificial oxygen administration and characteristic features on chest radiograph or computed tomographic (CT) imaging.

We included all consecutive patients (aged \geq 16 y), who received non-invasive respiratory support (continuous positive airway pressure (CPAP), low-flow/high-flow oxygen therapy) within the first 7 days of critical care admission. To avoid the confounding effects of sedative drugs/vasopressors and make our study relevant to patients with COVID-19 who receive ward-level care, we did not study patients requiring mechanical ventilation. vasopressor support or renal replacement therapy. We compared all paired measurements of arterial oxygen saturations measured by co-oximetry on arterial blood gas analysis (SaO₂) and the corresponding peripheral oxygenation saturations measured via pulse oximetry (S_pO_2) . In order to minimise the inclusion of artefactual S_pO_2 measurements, for example due to patient movement or poor signal quality, we used the mean value of readings recorded over the 4-min period that immediately preceded arterial blood gas analysis. Arterial blood gas analysis was performed using a RAPIDpoint 500 analyser (Siemens Healthcare GmbH, Erlangen, Germany) and S_pO_2 measurement done as part of the B1x5 M/P monitoring system (GE Healthcare, Chicago, IL, USA) using NellcorTM reusable (Medtronic, Watford, UK) or disposable (Mindray, Huntingdon, UK) probes. All devices were maintained and calibrated in accordance with the manufacturers' instructions.

Data retrieved from the database included: patient characteristics (age, sex, BMI, ethnicity); Rockwood clinical frailty core; APACHE-2; fraction of inspired oxygen concentration (F_1O_2); S_pO_2 ; and arterial blood gas measurements (F_1O_2 , SaO₂, PaO₂, PaCO₂, pH, bicarbonate, lactate and haemoglobin). We also recorded the duration of critical care stay.

The linear relationships between paired SaO_2 and S_pO_2 measurements were analysed using linear regression and goodness of fit. The bias and limits of agreement between SaO₂ and S_pO₂ measurements were assessed using Bland-Altman plots, with the difference between SaO_2 and S_pO_2 measurements plotted against the mean saturation measurement. Bias was calculated as the mean difference between SaO_2 and S_pO_2 measurements, limits of agreement were calculated as bias ± 1.96 SD and correction was made for within-subject variation [11]. To investigate if bias was affected at lower, more clinically important levels of oxygenation, analyses were repeated only including paired measurements with SaO₂ ≤94%. Contingency tables were used to calculate the ability of S_pO_2 measurement to detect hypoxaemia, defined as SaO_2 <90%; this value was chosen as being clinically relevant as</p> the target range for oxygen saturations in patients with COVID-19 pneumonitis was typically set at 88–92% in our critical care unit. The effect of patient ethnicity on the measures was then assessed using the same methods of analysis. Data were analysed using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA), R (version 4.1.0) and RStudio (v 1.4.1717).

Results

During the analysis period from 25 February 2020 to 16 December 2020, 489 patients were admitted to critical care. Data from 241 patients meeting the inclusion criteria were identified; 47 patients were not studied (41 had no arterial blood gases recorded in the database and six had an insufficient number of S_pO_2 readings within the 4-min time period before arterial blood gas analysis), leaving the data from 194 patients for analysis. These provided 6216 paired SaO_2/S_pO_2 measurements with mean (SD) S_pO_2 94.3% (3.0) and mean SaO_2 94.4% (2.8). The following ethnicities were recorded in our database: Asian British-Indian; Asian British-Pakistani; Asian British-other; Black British-African; Black British-Caribbean; Black British-other; White-British; White-Irish; White-other; and other ethnic group. For the purposes of analysis, ethnic groups were pooled as follows: Asian (Asian British-Indian; Asian British-Pakistani; Asian British-other); Black (Black British-African; Black British-Caribbean; Black British-other); White (White-British; White-Irish; White-other); and other. Patient characteristics are shown in Table 1 and groups were well matched except for a higher incidence of diabetes mellitus in patients of non-White ethnic origin.

The bias and limits of agreement for paired SaO₂ and S_pO₂ measurements are shown in Table 2. The Bland-Altman plots for paired SaO₂ and S_pO₂ measurements for the whole population and for those measurements taken in the presence of SaO₂ \leq 94% are shown in Figures 1a and 1b, respectively. The effect of patient ethnicity on the bias and

limits of agreement for paired SaO_2 and S_pO_2 measurements is shown in Table 2 and Figures 2–4.

Linear regression analysis showed that correlation between SaO₂ and S_pO₂ was poor overall ($R^2 = 0.4631$) and was worse in patients of Black ($R^2 = 0.2311$) compared with those of White and Asian ethnic origin ($R^2 = 0.4631$ and $R^2 = 0.4576$, respectively).

With respect to the ability of S_pO_2 to detect hypoxaemia, we analysed the number of S_pO_2 measurements that over-read compared with SaO_2 values. In total, there were 353 measurements with SaO_2 <90%; paired S_pO_2 readings were \geq 90%, \geq 92% and \geq 94% in 259 (73.4%), 191 (54.1%) and 132 (37.4%) of measurements, respectively. For patients of White ethnic origin, there were 238 readings with SaO_2 <90%; the paired S_pO_2 values were \geq 90%, \geq 92% and \geq 94% in 164 (68.9%), 111 (46.6%) and 71 (29.8%) of measurements, respectively. For patients of Asian ethnic origin, there were 61 readings with SaO_2 <90%; the paired

Table 1	Characteristics of patients with	n COVID-19 pneumonitis w	ho were admitted to	critical care for non-	invasive respiratory
support. \	/alues are mean (SD), number	(proportion) or median (IQ	R[range]).		

	Total n = 194	White ethnic origin n = 135	Asian ethnic origin n = 34	Black ethnic origin n = 19	Other ethnic origin n = 6
Paired readings	6216	4197	1241	599	179
Age; y	62(12.4)	63 (12.9)	59(11.9)	62 (9.4)	61 (12.9)
Sex; male	140 (72%)	96(71%)	26(76%)	13 (68%)	5 (83%)
BMI; kg.m ⁻²	31 (6.7)	31 (7.1)	30 (5.6)	31 (5.4)	27 (3.9)
Diabetes	67 (35%)	35 (26%)	18 (53%)	9 (47%)	5 (83%)
Admission APACHE-2	15(12–19)[6–45])	14(12–19[6–45])	15 (12–17.5 [6–34])	19(12–21[8–31])	15.5 (13.75–18 [13–24])
Clinical frailty scale	3 (3–4 [1–7])	3 (3–4 [1–7])	3 (3–3 [1–6])	3 (3–3 [2–4])	3 (2.75–4 [2–4])
Duration of critical care stay; days	11 (6–20 [1–78])	11 (6–20 [1–78])	10(7–26.25[2–60])	7 (5–17 [1–33])	11.5 (6–22.5 [6–33])

Table 2 Effect of ethnicity on bias in measurements of oxygen saturation by arterial blood gas analysis (SaO₂) and peripheral oxygen saturation (S_pO_2) in patients with COVID-19 pneumonitis admitted to critical care for non-invasive respiratory support. Values are shown for all readings and those when the patient was hypoxaemic (defined as SaO₂ \leq 94%).

	All measurements			Measurements when SaO ₂ ≤94%			
	Bias	Lower limit of agreement (95%CI)	Upper limit of agreement (95%CI)	Bias	Lower limit of agreement (95%CI)	Upper limit of agreement (95%Cl)	
All patients	0.05	-2.21 (-2.16 to -2.26)	2.30 (2.25 to 2.35)	0.25	-2.26 (-2.17 to -2.34)	2.77 (2.68 to 2.86)	
Patients of White ethnic origin	0.28	-1.79 (-1.74 to -1.85)	2.35 (2.29 to 2.40)	0.44	-1.88 (-1.79 to -1.98)	2.76 (2.67 to 2.85)	
Patients of Asian ethnic origin	-0.33	-2.47 (-2.36 to -2.56)	1.80 (1.69 to 1.90)	0.16	-2.80 (-2.56 to -3.03)	2.48 (2.24 to 2.70)	
Patients of Black ethnic origin	-0.75	-3.47 (-3.26 to -3.64)	1.97 (1.76 to 2.14)	-0.66	-3.85 (-3.46 to -4.34)	2.54 (2.12 to 3.02)	



Figure 1 Bland-Altman plot paired measurements of oxygen saturation by arterial blood gas analysis and peripheral oxygen saturation in 194 patients with COVID-19 pneumonitis who were admitted to critical care for non-invasive respiratory support. Patient ethnic origin is shown by the colour of each datum (White = yellow; Asian = purple; Black = light blue; and Other = green). The solid line represents the bias between the two measurements, the dashed line represents the limits of agreement (bias \pm 1.96 SD) and the dotted line represents the 95%Cl for the limits of agreement. (a) shows all paired measurements and (b) shows only those measurements when the patient was hypoxaemic (defined as SaO₂ ≤94%).



Figure 2 Bland-Altman plot paired measurements of oxygen saturation by arterial blood gas analysis and peripheral oxygen saturation in 135 patients of White ethnic origin with COVID-19 pneumonitis who were admitted to critical care for non-invasive respiratory support. The solid line represents the bias between the two measurements, the dashed line represents the limits of agreement (bias \pm 1.96 SD) and the dotted line represents the 95%Cl for the limits of agreement. (a) shows all paired measurements and (b) shows only those measurements when the patient was hypoxaemic (defined as SaO₂ \leq 94%).



Figure 3 Bland-Altman plot paired measurements of oxygen saturation by arterial blood gas analysis and peripheral oxygen saturation in 34 patients of Asian ethnic origin with COVID-19 pneumonitis who were admitted to critical care for non-invasive respiratory support. The solid line represents the bias between the two measurements, the dashed line represents the limits of agreement (bias \pm 1.96 SD) and the dotted line represents the 95%Cl for the limits of agreement. (a) shows all paired measurements and (b) shows only those measurements when the patient was hypoxaemic (defined as SaO₂ ≤94%).



Figure 4 Bland-Altman plot paired measurements of oxygen saturation by arterial blood gas analysis and peripheral oxygen saturation in 19 patients of Black ethnic origin with COVID-19 pneumonitis who were admitted to critical care for non-invasive respiratory support. The solid line represents the bias between the two measurements, the dashed line represents the limits of agreement (bias \pm 1.96 SD) and the dotted line represents the 95%Cl for the limits of agreement. (a) shows all paired measurements and (b) shows only those measurements when the patient was hypoxaemic (defined as SaO₂ \leq 94%).

S_pO₂ values were ≥90%, ≥92% and ≥94% in 50 (82.0%), 37 (60.7%) and 24 (39.3%) of measurements, respectively. For patients of Black ethnic origin, there were 38 readings with SaO₂ <90%; the paired S_pO₂ values were ≥90%, ≥92% and ≥94% in 32 (84.2%), 30 (78.9%) and 27 (71.1%) of measurements, respectively.

The overall incidence of S_pO_2 measurements being \geq 90% with the paired SaO₂ value being <90% was 5.7% (95%CI 5.1-6.3). However, this is somewhat of an underestimate as the denominator includes measurements with high/normal SaO₂ values; in the patient population of clinical interest (those with $SaO_2 \leq 94\%$), the incidence increased to 14.5% (95%CI 13.1-15.9). The incidence of SO₂ measurements being >90% with the paired SaO₂ value being \leq 90% was broadly similar in each of the ethnic groups when all measurements were analysed: patients of White ethnic origin 5.7% (95%CI 5.0-6.4); patients of Asian ethnic origin 4.9% (95%CI 3.9-6.3); and patients of Black ethnic origin 6.3% (95%CI 4.7–8.6). However, when the population was limited to lower SaO_2 readings (SaO_2 $\leq\!\!94\%$), the incidence was increased by approximately three-fold: patients of White ethnic origin 14.1% (95%Cl 12.5-15.8); patients of Asian ethnic origin 13.9% (95%Cl 11.0-17.5); and patients of Black ethnic origin 15.9% (95%CI 11.8-21.1).

Discussion

In patients who are critically ill with COVID-19 pneumonitis, measurement of oxygen saturation by S_pO_2 equates closely with SaO₂, although within limits of agreement that cover a 4–5% range of saturations. The significance of this will depend upon the clinical situation, with a greater risk of occult hypoxaemia in patients who have a low SaO₂, either due to disease severity or the deliberate targeting of lower levels of oxygenation by clinicians (e.g. as part of lung-protective ventilation strategies, as a triage decision for ward-based care in patients who do not meet other criteria for critical care admission or as a strategy to conserve oxygen supplies). Of note, measurement of S_pO_2 was not affected by patient ethnicity to a clinically significant degree.

The US Food and Drug Administration sets out several standards for studies undertaken to test the accuracy of pulse oximeters including: ≥ 10 healthy subjects that vary in age and sex; ≥ 200 paired S_pO₂/SaO₂ measurements; and a range of skin pigmentations, including at least two darkly pigmented subjects or 15% of the study population (whichever is larger) [1]. Given the huge variation in age, ethnic origin, medical comorbidities and physiological derangements in patients admitted to critical care, this relatively low standard for calibration is surprising. The

recommendations for acceptable accuracy are a root mean square difference between S_pO₂ and SaO₂ measurements of <3% for finger and <3.5% for ear probes. This is calculated using the following formula: $\sqrt{bias^2+precison^2}$). In our study, all the patient groups were within this target accuracy specification: all patients 1.09%; patients of White, Asian and Black ethnic origin 1.08%, 1.13% and 1.56%, respectively.

The COVID-19 pandemic has highlighted significant differences between ethnic groups in terms of illness severity and mortality, with patients of non-White ethnic origin having an increased risk of death with SARS-CoV-2 infection, especially during the first wave of infections in the UK [Wan et al., preprint, https://doi.org/10.1101/2021.07. 05.21260026]. Measurements of S_pO_2 were used commonly in the UK as a triage tool to help guide the requirement for admission to hospital and/or critical care [12], and as a method for the titration of supplemental oxygen therapy and/or respiratory support. A retrospective database analysis of patients in the USA who were critically ill showed that S_pO_2 measurement failed to detect hypoxaemia (defined as SaO₂ <88%) in 11.7% of measurements in patients of Black ethnic origin. This led to concerns that the development of occult hypoxaemia in patients with pigmented skin was one contributory factor to the increased risk of mortality in certain ethnic groups. In our study, we defined hypoxaemia as SaO₂ <90%, and among 6216 paired readings, there were 353 occasions when S_pO_2 measurement failed to detect this (instead giving a reading \geq 90%), with the incidence similar in patients of White, Asian and Black ethnic origin. Our findings suggest that S_pO_2 measurement remains a useful triage tool in monitoring oxygen therapy in patients of Black and Asian ethnic origin with COVID-19 pneumonitis and mild to moderate hypoxaemia. However, the relatively low numbers of patients with profound hypoxaemia in our study should lead to cautious interpretation of this observation, and the potential for patient ethnicity to introduce bias at lower oxygen saturation levels warrants further investigation.

The bias and limits of agreement between SaO₂ and S_pO₂ measurement are lower than that seen in other studies (summarised in Table 3). There are several possible reasons for this. First, it may be that S_pO₂ measurement technology has improved over recent years. Second, all our patients were cared for in an ICU but were only receiving singleorgan support for a respiratory disease; this may have resulted in a greater focus on respiratory parameters by members of the nursing staff. Third, we did not study patients who were sedated and/or were receiving

Table 3 Summary of previous studies assessing bias in the measurements of oxygen saturation by arterial blood gas analysis (SaO_2) and peripheral oxygen saturation (S_pO_2) . Where different ethnic groups or skin pigmentation were analysed these data are included. If not stated in the original study, then limits of agreement were calculated as 1.96 SD. Values are number or mean (SD).

n	Patient population	Mean S _p O ₂ ;%	Ethnic group/skin pigmentation	Bias; %	Limits of agreement; %
54	Medical ICU	Not reported	All	2.7	-1.8-7.2
			Black	3.3	-2.0-8.6
			White	2.2	-1.3-5.7
284	Emergency department	94(5)	Light	2.5	-6.5-11.5
			Intermediate	2.8	-7.4-13.0
			Dark	2.2	-5.1-9.5
404	ICU	95.6(3.0)	All	0.15	-4.2-4.5
41	ICU	94.6(2.7)	All	1.34	-2.3-5.0
30	ICU	96.44 (2.20)	All	0.4	-4.3-5.2
90	Emergency department	93.9(4.8)	All	2.75	-3.4-8.9
102	ICU	90.4 (75–100) ^a	All	-0.02	-4.2-4.2
	n 54 284 404 41 30 90 102	nPatient population54Medical ICU284Emergency department404ICU404ICU30ICU90Emergency department102ICU	nPatient populationMean SpO2;%54Medical ICUNot reported284Emergency department94(5)404ICU95.6(3.0)41ICU94.6(2.7)30ICU96.44(2.20)90Emergency department93.9(4.8)102ICU90.4(75–100) ^a	nPatient populationMean SpO2;%Ethnic group/skin pigmentation54Medical ICUNot reported BlackAll54Medical ICUNot reported UniteBlack284Emergency department Unite94(5)Light Intermediate Dark404ICU95.6(3.0)All41ICU94.6(2.7)All30ICU96.44(2.20)All90Emergency department93.9(4.8)All102ICU90.4(75-100) ^a All	nPatient populationMean SpO2;%Ethnic group/skin pigmentationBias; %54Medical ICUNot reported BlackAll2.754Medical ICUNot reported HarkBlack3.354Emergency department94(5)Light2.2284Emergency department94(5)Light2.87404ICU95.6(3.0)All0.15410ICU94.6(2.7)All1.3430ICU96.44(2.20)All0.490Emergency department93.9(4.8)All2.75102ICU90.4(75-100) ^a All-0.02

^amedian (range) reported.

vasopressor infusions, which is not typical for many patients in an ICU.

Of interest is the fact that bias and limits of agreement between SaO_2 and S_pO_2 measurement were not altered by the exclusion of patients who were normoxic (SaO₂ \geq 94%). This is important as the degree of bias when the oxygen saturation is high-normal is of little clinical significance. For example, a 4% difference in measured S_pO_2 and SaO_2 is unlikely to alter management when these are both \geq 95%; however, the same 4% difference at saturations in the region of 88–92% risks either an unnecessary intervention (such as an arterial blood gas sample or increased respiratory support) or failure to recognise hypoxaemia. It is usual for S_pO_2 to be monitored continuously and readings will be dynamic in nature. Thus, the degree of bias may vary throughout the monitoring period and may not be consistent or predictable. However, despite our ICU policy of targeting an SaO₂ 88–92%, 5172 out of 6216 (82%) SaO₂ measurements were >92%, suggesting that most patients were receiving an excessive F_1O_2 . This limits our ability to determine the bias between S_pO_2 and SaO_2 at lower levels of oxygen saturation. Previous work has suggested that bias is increased at lower oxygen saturation levels (\leq 90%), especially in patients who have greater degrees of skin pigmentation [6, 7].

Clinicians may wish to consider our findings when producing guidelines relating to S_pO_2 targets in patients with COVID-19 pneumonitis. A target S_pO_2 of 88% may risk occult hypoxaemia, given the accuracy and limits of agreement for this measurement, and a S_pO_2 target of \geq 90% may be of value in ward-based environments that do not

have the ability to measure SaO_2 rapidly and repeatedly by sampling from arterial catheters.

There are some limitations to our study. First, this was a single-centre, retrospective study of a single disease state and all measurements were done using one type of equipment. As such, our findings should be seen as being hypothesis-generating and should not be extrapolated uncritically to other institutions, diseases or monitoring/ measurement devices. Second, in order to avoid the potential impact of variations in peripheral vasomotor tone on S_pO_2 readings (either as a result of sedation or vasopressor administration), we did not study patients whose tracheas had been intubated and were receiving mechanical ventilation. Third, ethnic origin was determined by subjective, patient (or next of kin) self-identification. In addition, 70% of our study population was of White ethnic origin; the smaller number of patients of Asian and Black ethnic origin could have been a reason for the lack of precision around limits of agreement. Fourth, S_pO₂ monitoring is dynamic and prone to artefact caused by poor probe positioning or patient movement. We attempted to minimise the effect of such errors by taking a mean of the S_pO_2 readings in the 4-min period before SaO_2 measurement, but we cannot exclude the possibility that some artefactual readings may have been included in the analysis. However, this is a limitation of all studies that analyse real-time physiological measurement data retrospectively and reflects real-world data measurement conditions. We also did not correct for the presence of anaemia, acidaemia or skin temperature, all of which may have a small impact on bias measurement in oxygenation saturation in patients who are critically ill [2, 3]. Finally, we did not perform a formal power calculation and instead chose a sample size of convenience. However, a post-hoc calculation suggested that 4405 paired measurements were necessary in order to detect a 2.4% difference between S_pO_2 and SaO_2 measurements (assuming α = 0.05 and β = 0.10).

In summary, in patients with COVID-19 pneumonitis admitted to critical care for non-invasive respiratory support, S_pO_2 measurements showed a level of agreement with SaO₂ values in line with previous studies. This was not affected by the patients' self-reported ethnicity. However, the limits of agreement between SaO₂ and S_pO_2 values may become clinically significant at lower levels of oxygen saturation and should be taken into account when using S_pO_2 measurements as a triage and therapeutic monitoring tool. In addition, guidelines should recommend a low threshold for the direct measurement of SaO₂ and emphasise the variable accuracy of S_pO_2 measurement.

Acknowledgements

MW and IM are editors of *Anaesthesia*. No external funding or other competing interests declared.

References

- US Food and Drug Administration. Pulse Oximeters Premarket Notification Submissions [510(k)s]: Guidance for Industry and Food and Drug Administration Staff. 2013. https://www.fda. gov/regulatory-information/search-fda-guidance-documents/ pulse-oximeters-premarket-notification-submissions-510ksguidance-industry-and-food-and-drug (accessed 01/07/2021).
- Ebmeier SJ, Barker M, Bacon M, et al. A two centre observational study of simultaneous pulse oximetry and arterial oxygen saturation recordings in intensive care unit patients. *Anaesthesia and Intensive Care* 2018; **46**: 297–303.

- 3. Perkins GD, McAuley DF, Giles S, Routledge H, Gao F. Do changes in pulse oximeter oxygen saturation predict equivalent changes in arterial oxygen saturation? *Critical Care* 2003; **7**: R67–71.
- 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.
- Philip KEJ, Bennett B, Fuller S, et al. Working accuracy of pulse oximetry in COVID-19 patients stepping down from intensive care: a clinical evaluation. *British Medical Journal Open Respiratory Research* 2020; 7: e000778.
- Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology* 2005; **102**: 715–9.
- 7. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesthesia and Analgesia* 2007; **105**: S18–23.
- Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. *New England Journal of Medicine* 2020; 383: 2477–8.
- 9. Doidge JC, Gould DW, Ferrando-Vivas P, et al. Trends in intensive care for patients with COVID-19 in England, Wales, and Northern Ireland. *American Journal of Respiratory and Critical Care Medicine* 2020; **203**: 565–74.
- Lopez L, Hart LH, Katz MH. Racial and ethnic health disparities related to COVID-19. *Journal of American Medical Association* 2021; **325**: 719–20.
- Shieh G. The appropriateness of Bland-Altman's approximate confidence intervals for limits of agreement. *BMC Medical Research Methodology* 2018; **18**: 1–11.
- 12. National Institute for Health and Care Excellence. COVID-19 rapid guideline: Managing COVID-19. [NG191]. 2021. https://www.nice.org.uk/guidance/ng191 (accessed 17/08/2021).
- Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. *Chest* 1990; 97: 1420–5.
- Adler JN, Hughes LA, Vtvilecchia R, Camargo CA Jr. Effect of skin pigmentation on pulse oximetry accuracy in the emergency department. *Academic Emergency Medicine* 1998; 5: 965–70.
- Louw A, Cracco C, Cerf C, et al. Accuracy of pulse oximetry in the intensive care unit. *Intensive Care Medicine* 2001; 27: 1606–13.