2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 979-8-3503-2447-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMBC40787.2023.10341069

Reducing racial bias in SpO_2 estimation: The effects of skin pigmentation

Matteo Bermond, Harry J. Davies, Edoardo Occhipinti, Amir Nassibi and Danilo P. Mandic*

Abstract— Accurate pulse-oximeter readings are critical for clinical decisions, especially when arterial blood-gas tests — the gold standard for determining oxygen saturation levels — are not available, such as when determining COVID-19 severity. Several studies demonstrate that pulse oxygen saturation estimated from photoplethysmography (PPG) introduces a racial bias due to the more profound scattering of light in subjects with darker skin due to the increased presence of melanin. This leads to an overestimation of blood oxygen saturation in those with darker skin that is increased for low blood oxygen levels and can result in a patient not receiving potentially life-saving supplemental oxygen. This racial bias has been comprehensively studied in conventional finger pulse oximetry but in other less commonly used measurement sites, such as in-ear pulse oximetry, it remains unexplored. Different measurement sites can have thinner epidermis compared with the finger and lower exposure to sunlight (such as is the case with the ear canal), and we hypothesise that this could reduce the bias introduced by skin tone on pulse oximetry. To this end, we compute SpO_2 in different body locations, during rest and breath-holds, and compare with the index finger. The study involves a participant pool covering 6-pigmentation categories from Fitzpatrick's Skin Pigmentation scale. These preliminary results indicate that locations characterized by cartilaginous highly vascularized tissues may be less prone to the influence of melanin and pigmentation in the estimation of SpO_2, paving the way for the development of non-discriminatory pulse oximetry devices.

I. INTRODUCTION

Photoplethysmography (PPG) is a non-invasive optical technique to measure changes in blood volume. The typical shape of the PPG signal is given in Figure 1B and has two characteristic patterns: Systolic, representing the inflow of blood and the expansion of the arterial system, and Diastolic, corresponding to the outflow of blood and the contraction of the arterial system. The PPG signal is composed of two main parts: the pulsatile (AC) component generated by arterial blood flow and the non-pulsatile (DC) component obtained from the venous and capillary blood and the absorption due to tissue, skin, and bone. When implemented at two different wavelengths, namely red and infrared light, PPG allows for the estimation of blood oxygen saturation (SpO_2). Red and infrared light are used since oxygenated haemoglobin absorbs infrared light more efficiently than red light, whereas the opposite is true for deoxygenated haemoglobin. This combination of different wavelengths of PPG for the purposes of estimating blood oxygen saturation is referred to as Pulse Oximetry.

This simple, low-cost, non-invasive technique has become increasingly important for clinical decisions during the COVID-19 pandemic given the early warning sign of low blood oxygen levels without breathlessness (silent hypoxia). However, there is strong evidence that pulse oximeters are less accurate in dark-skinned individuals, especially in critical cases of lower saturation (<80%) where it results in overestimation [1]. Long-standing oximetry theory does not fully account for the way in which photons are scattered by an increased presence of melanin in the epidermis, making corrections for the effect of pigmentation difficult to implement. Calibration studies compound this problem because they typically oversample people with lighter skin. This is an issue of major concern since SpO_2 values are used to inform treatment and an overestimation of these values can result in patients not being given access to supplemental oxygen when it is needed.

The USA Food and Drug Administration (FDA) requires each clinical study to include at least two participants with dark pigmentation or 15% of the participant pool, whichever is larger. It is worth mentioning that the FDA published a safety communication in February 2021 highlighting the limitations of current pulse oximeters [2], which overestimate the oxygen concentration in some individuals with darker skin, leading to inaccurate estimation of dangerously low oxygen saturation levels (Figure 1C). Driven by clinical experiences, early in the pandemic of COVID-19, Sjoding et al. [3] published a retrospective report showing that pulse oximeters overestimate the true oxygen saturation of individuals with darker skin. This inaccuracy means that the diagnosis of hypoxaemia, the condition of having low levels of oxygen saturation in the blood, is approximately three times more likely to be missed in darker skin patients than in white patients. Misdiagnosed patients are said to have occult hypoxaemia when arterial blood-gas tests indicate oxygen saturation levels of less than 88% (signalling hypoxaemia), despite pulse oximeters measuring healthy oxygenation of more than 92%. Feiner et al demonstrated that the overestimation in blood oxygen saturation increases as the saturation drops introducing an error of up to 7% in the SpO_2 for subjects with darker skin types [4].

Existing investigations of racial bias have been carried out in the conventional finger PPG sensor, but several emerging locations remain uninvestigated. Importantly, locations across the head can have reduced skin thickness when compared with the finger, better perfusion, and even reduced exposure to sunlight which is the case with the ear canal. We hypothesize that different measuring sites for pulse oximetry, such as the ear canal, auricle, nose, and forehead,

* M.B., H.J.D., A.N. and D.P.M. are with the Electrical and Electronic Engineering Department, Imperial College London, United Kingdom.

E.O. is with the UKRI Centre for Doctoral Training in AI for Healthcare, Imperial College London, United Kingdom. E.O. is supported by the UKRI Centre for Doctoral Training in AI for Healthcare grant number EP/S023283/1.

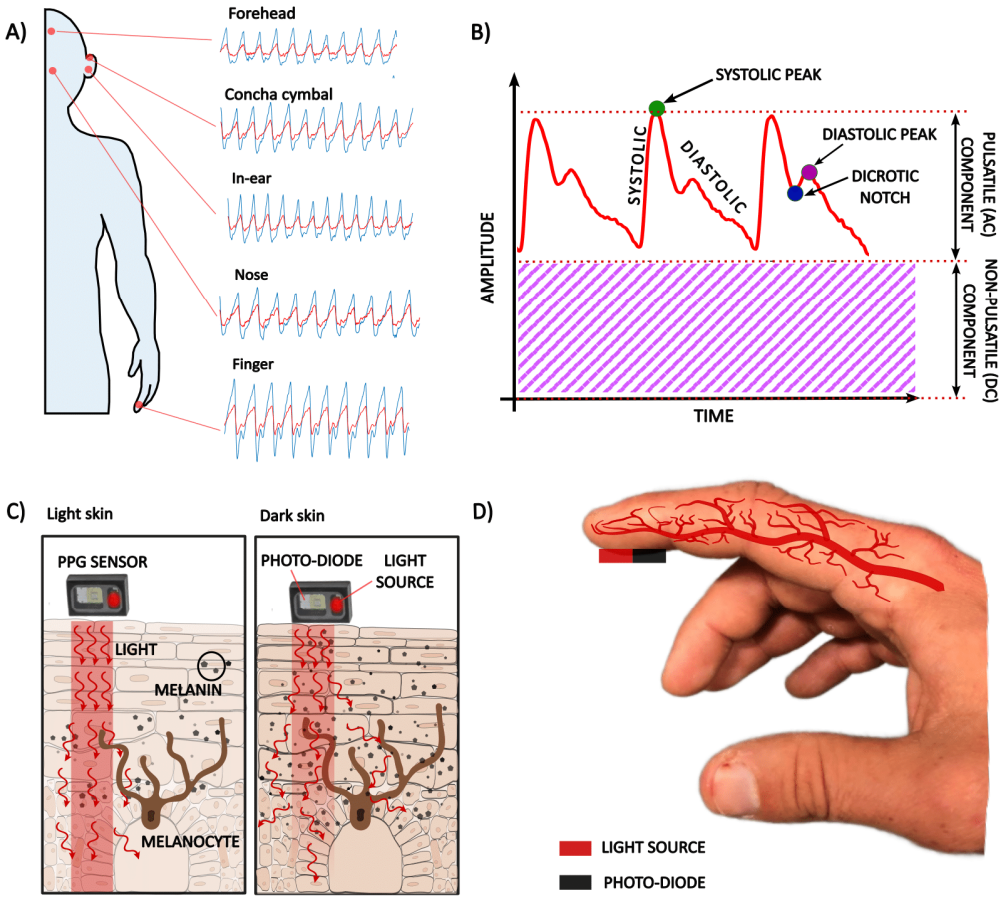

Figure 1: Pulse-oximetry accuracy varies with skin tone. A) Evaluation of PPG signal quality in different body locations. B) A typical waveform of the PPG and its characteristic parameters. C) Light signals are affected by melanin, which is distributed through the skin in structures, known as melanosomes, that are produced by cells called melanocytes. Melanosomes in dark skin are both larger and more numerous than those in light skin. D) Pulse oximeter is used in this study to estimate the oxygen concentration in a person's blood by shining the red and infrared light through their skin. Oxygenated haemoglobin absorbs infrared light more efficiently than red light, whereas the opposite is true for deoxygenated haemoglobin.

may help mitigate the impact of skin tone on the overestimation of SpO_2. To this end, this pilot study estimates the SpO_2 from five locations: index finger, concha cymbal, ear canal, forehead, and nose. Every location but the concha cymbal has been used in previous work [5][6], including investigations of their feasibility and limitation in SpO_2 estimation. The concha cymbal thus presents a novel location for the estimation of blood oxygen saturation using pulse plethysmography. Each location has been compared with the index finger as it is the common position where the oximeter probe is placed in medical practices.

II. Methods

A. Subjects pool

In order to have a sample of all types of pigmentation, and their effect on oxygen blood saturation, the analysis involved twelve healthy, non-smoking subjects ranging in age between 25 and 40 years. None of the subjects had lung disease, obesity, or cardiovascular problems. Since different gender can introduce different errors in the SpO_2 estimation [4], we recorded one female and one male subject for each class of the six-skin pigmentation class according to Fitzpatrick's skin type classification [7]. The recordings were performed under the Imperial College London ethics committee approval JRCO 20IC6414, and the subjects were given full informed consent.

B. Protocol

The SpO_2 was estimated by placing a MAX30101 digital PPG sensor at different body positions in order to assess the relation between the recording site and the effect of skin pigmentation. The sensor comprises green (527 nm), red (660 nm) and infrared (880 nm) light-emitting diodes, as well as a photodiode to measure the reflected light (Figure 1D). The three light wavelengths gave three measurements of the pulse, however, at this stage of the study, we were mainly interested in the red and infrared wavelengths as their ratio allows for the estimation of SpO_2. The considered locations were the index finger, nose, forehead, ear canal, and concha cymbal (Figure 1A). Data were collected from subjects for 6 minutes while in a supine and relaxed position, breathing normally. At every 2-minute interval, they were required to breathe out and hold their breath. All the subjects were able to hold their breath for more than 12 seconds. This protocol ensures two visible drops of the SpO_2 at each recording location, which thus allowed us to evaluate the accuracy of each location with respect to the index finger. considered as a gold standard. Figure 2B shows an example of the collected information from the different locations.

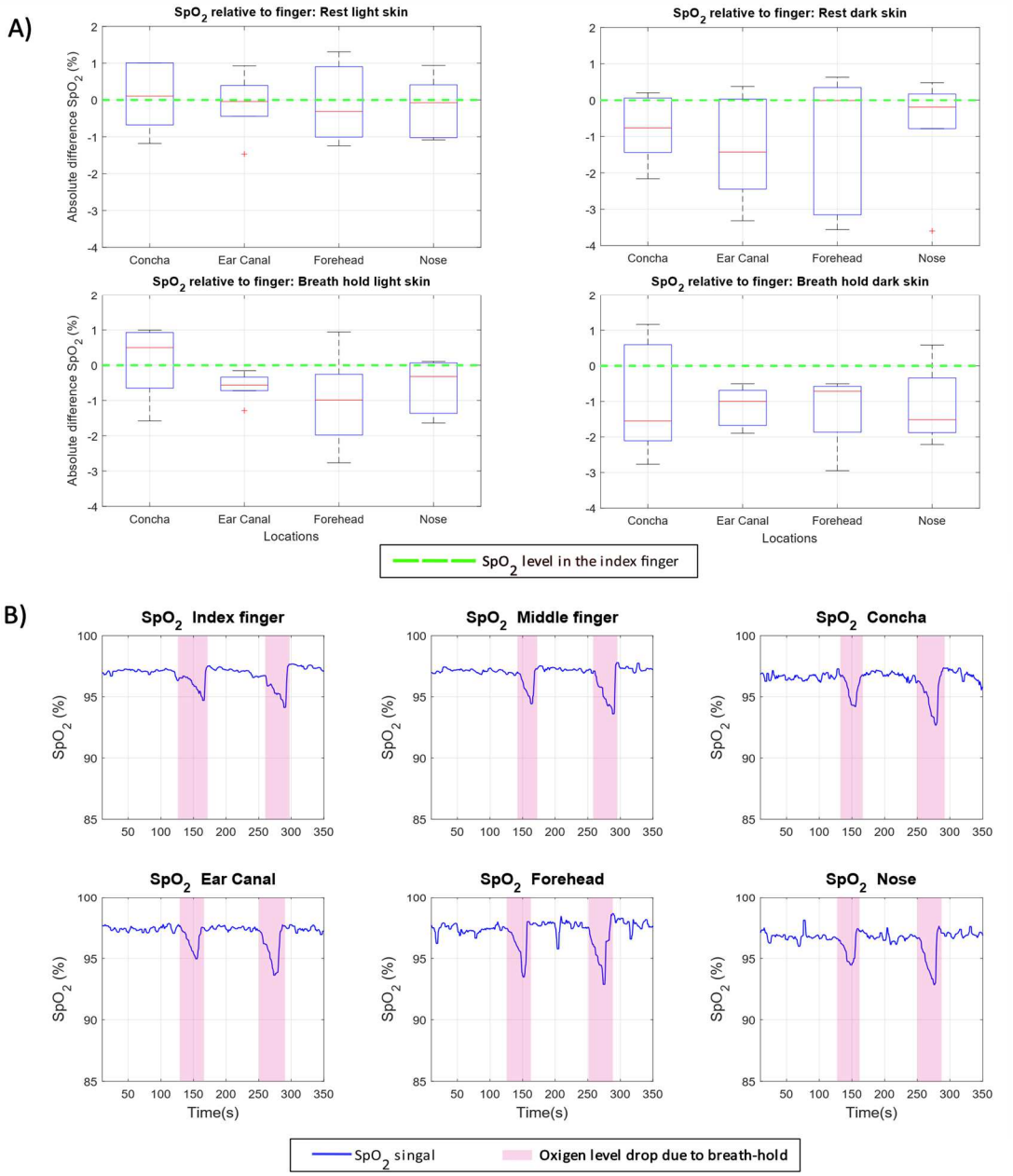

Figure 2: SpO_2 estimation from different body locations. A) Absolute difference between the SpO_2 estimates at the considered location and the SpO_2 from the index finger. The difference in the SpO_2 was computed for each location dividing the subjects pool into two main categories: Light skin (type 1 and type 2) and Dark skin (type 3 to type 6). The green dashed line shows the SpO_2 value detected in the index finger during the saturation drop. B) SpO_2 computed in every location considered in this study, for a representative subject of the pool. Observe how holding the breath after complete exhalation results in significant drops in blood oxygenation.

C. Signal processing

As explained in [5], the ratio of absorbance of the infrared and red light within the PPG sensor changes depending on the variation of oxygenated haemoglobin in the blood. To this end, for each recording location, the ratio of the pulsatile (AC) wave and non-pulsatile (DC) PPG wave was computed for both the red and infrared light.

To acquire the AC components within the PPG measurements, the raw signals were firstly band-pass filtered between 1 Hz and 30 Hz and peak detection was implemented using findpeaks in Matlab to find the peaks and troughs. In addition, peaks corresponding to artefacts were manually excluded, to ensure that results were not influenced by motion artefacts and noise [11]. The peak and trough values were separated and interpolated before their absolute values were added together to give a constant estimate of the AC amplitude. The DC components were obtained by low pass filtering the raw signal at 0.01 Hz. The ratio between the red and the infrared component was then quantified and the SpO2 value as a proxy for blood oxygen saturation was calculating using an empirically derived linear approximation given by (1).

$$SpO_2 = 104 - 17\frac{{AC_{red}}/{DC_{red}}}{{AC_{infrared}}/{DC_{infrared}}}. \qquad (1)$$

Once the saturation was computed at every location, we split the signal into four parts, corresponding to the recording segments in which the subject was breathing normally and those where the drops in the SpO_2 were identified. These drops were identified using the findpeak function on the amplitude-reversed signal. For every section, we computed the mean and median of the SpO_2 values. The Root Mean Square Error (RMSE) and the absolute difference with respect to the index finger were calculated, based on the location of each sensor, as a measure of error magnitude.

III. RESULTS AND DISCUSSION

As the error in the SpO_2 increases for values below 88%, the RMSE analysis was used to classify the quality of the different measurement locations with respect to the index finger. The recording location was marked as a more valuable position than the index finger for the estimation of the SpO_2 if it presented a similar SpO_2 value in the resting sections and a lower value in the drops in the SpO_2 caused by breath-holding.

Figure 2A shows the standard deviation of the SpO_2 at rest as an estimate of signal quality, and the SpO_2 computation was computed as a median over a 5-second window to reduce the effects of artefacts. It was observed that locations such as the concha cymbal exhibit negligible variation over the subject with lighter skin in both rest condition and during the breath hold (absolute difference $SpO_2 < 1\%$); on the contrary, for the darker skin subjects, while the rest conditions showed similar variability in the pulse-ox estimation, the concha cymbal showed a higher difference during the drops compared to the index finger. This difference (lower estimates of the SpO_2) leads to the conclusion that the concha cymbal might be a more accurate location for the pulse-ox estimation since appear less affected by melanin. The forehead and the ear canal exhibited high standard deviation across all the subjects likely due to artefacts and the difficulty of achieving a stable sensor position (Table 1). Given that a poor signal-to-noise ratio can bias SpO_2 estimation downwards, it was difficult to make any conclusions on these positions at this stage. The nose and the concha cymbal, on the other hand, exhibited lower standard deviation over the subjects (Table 1), and a consistent coherence with the index finger (absolute difference $SpO_2 < 1\%$) in the rest section. The nose did not show a significant difference in SpO_2 estimation between light and dark skin, but the concha cymbal exhibited a larger drop compared to the finger in subjects with dark skin. The result indicates that in a darker skin population, SpO_2 recorded from the concha may produce a more accurate estimate of blood oxygen saturation in hypoxia scenarios, compared to the same sensor placed on the finger. A possible reason for this is that the concha cymbal is cartilaginous with a thin epidermis, whilst being strongly vascularised. These structural differences may help to mitigate the impact of melanin on blood oxygen estimation, resulting in a SpO_2 measurement that is not racially biased.

Previous works such as [8] [9] [10] analyze in depth the light absorption and scattering mechanism from in vivo skin showing how those are affected both by wavelength and concentration of melanin. It was demonstrated that the remittance of light in human skin can be modelled as a straight line in the visual spectrum (400 – 720 nm). Furthermore, it was shown that the slope of this straight line is a reliable indicator of the pigment level in the skin. As a matter of fact, a higher concentration of melanin, hence a darker skin pigment, exhibits a steeper and higher absorption spectrum.

To obtain reliable data to overcome the lack of features of commercial oximeters regarding dark skin pigmentation, future studies involving several participants covering all the spectrum defined by Fitzpatrick's Skin Pigmentation scale should be performed. This will help to redesign the SpO_2 calculation algorithm related to light absorption with a calibration based on different wavelengths. This would also yield a new equation for the computation of the SpO_2, which would modify the parameters of equation (1) depending on the skin type. This will also allow to maintain the current design of pulse oximeters, guaranteeing real-time calibration based on skin pigmentation levels. In addition, the non-invasive nature of PPG sensors allows for their placement at different body locations where the effects of melanin and pigmentation are reduced.

Table 1: Standard deviation computed over the subjects at rest and normalized to the index finger.

Subjects	Standard deviation at rest in different body locations as a proxy for signal quality			
	Concha [%]	In-ear [%]	Forehead [%]	Nose [%]
All	1.167	2.520	2.486	1.510
Light skin	1.122	3.768	3.641	1.980
Dark skin	1.190	1.882	1.837	1.270

IV. CONCLUSION

We have performed a comprehensive evaluation of blood oxygen saturation from PPG, with the aim to reduce the racial bias in current pulse oximeters due to melanin and skin pigmentation. Out of the four locations investigated (concha cymbal, ear canal, forehead, and nose), the concha cymbal has exhibited a similar behaviour to the index finger when estimating SpO_2 in resting periods, while, as desired, resulting in a consistently lower value when blood oxygen saturation drops below 90%. Despite the small sample size, this pilot study lays the ground for novel pulse oximetry locations that could mitigate the impact of darker skin on the bias in blood oxygen estimation. A pulse oximeter without racial bias would dramatically reduce the chances of misdiagnosed pathologies and promote fairness in health diagnosis. Following this proof of concept, a comprehensive study over a much larger cohort of subjects is underway.

REFERENCES

[1] A. M. Cabanas, M. Fuentes-Guajardo, K. Latorre, D. León, and P. Martín-Escudero, "Skin pigmentation influence on pulse oximetry accuracy. A systematic review and bibliometric analysis", Sensors, vol. 22, 2022, pp. 3402.

[2] "Pulse oximeter accuracy and limitations FDA safety communication" 2022. [Online]. Available: https://www.fda.gov/medical-devices/safety-communications/pulse-oximeter-accuracy-and-limitations-fda-safety-communication.

[3] M. W. Sjoding, R. P. Dickson, T. J. Iwashyna, S. E. Gay, and T. S. Valley, "Racial bias in pulse oximetry measurement", New England Journal of Medicine, vol.383, 2020, pp. 2477–2478.

[4] J. R. Feiner, J. W. Severinghaus , P. E. Bickler, "Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender", Anesthesia & Analgesia, vol 105, 6 Suppl, 2007, pp. 18-23.

[5] H. J. Davies, I. Williams, N. S. Peters, and D. P. Mandic, "In-Ear SpO_2: A tool for wearable, unobtrusive monitoring of core blood oxygen saturation", Sensors, vol. 20, no. 17, 2020, pp. 4879.

[6] M. Huotari, J. Röning and K. Määttä, "P170 a forehead and nasal bridge pulse oximeter comparison measurements on healthy subjects", Artery Research, vol. 24, 2018, pp. 130.

[7] S. Del Bino, C. Duval, and F. Bernerd, "Clinical and biological characterization of skin pigmentation diversity and its consequences on UV impact", International Journal of Molecular Sciences, vol. 19, 2018, pp. 2668.

[8] N. Kollias, and A. H. Baqer, "Absorption mechanisms of human melanin in the visible, 400-720 nm", Journal of Investigative Dermatology, vol.89, 1987, pp. 384–388.

[9] N. Kollias, "The spectroscopy of human melanin pigmentation, melanin: its role in human photoprotection", Valden, 1995, pp. 31-38.

[10] V. V. Barun, A. P. Ivanov, A. V. Volotovskaya, V. S. Ulashchik, "Absorption spectra and light penetration depth of normal and pathologically altered human skin", Journal of Applied Spectroscopy, vol.74, 2007, pp. 430–439.

[11] E. Occhipinti, H. J. Davies, G. M. Hammour, and D. P. Mandic, "Hearables: artefact removal in ear EEG for continuous 24/7 monitoring", Proceedings of the IEEE International Joint Conference of Neural Networks (IJCNN), 2022.