Expansion of Smartwatch Use in Daily Life Activities: Reliability in Heart Rate Variability Measurement

Kelly C B da Silva[1], Petrônio C Ferreira[1], Ana L G dos Santos[1], Stella T Maximo[2], Christian G Sassaki[1], Samuel M Camargo[1], Silvia H B de Paula[3], José L Puglisi[4], Daniel G Goroso[1.5]

[1]Universidad Mogi da Cruzes, SP, Brazil
[2]Faculdade de Medicine São Leopoldo Mandic, SP, Brazil
[3]Health Institute of São Paulo, SP, Brazil
[4]College of Medicine, California North State University, EUA
[5]Facultad de Educación Física, Universidad Nacional de Tucumán, Argentina

Abstract

This study investigates the influence of Fitzpatrick skin phototypes on the accuracy of smartwatch-derived blood pressure and heart rate measurements. Forty-eight healthy participants were equally distributed among skin phototypes II, III, and IV (n = 16/group) and were simultaneously evaluated using a Samsung Galaxy Watch 4 (fs = 100 Hz) and an automated sphygmomanometer as the reference method. Systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR) were analyzed using Pearson's correlation and Bland–Altman agreement analyses, both globally and stratified according to skin phototype. Strong correlations were observed across all phototypes ($r = 0.821–0.989$; $p < 0.001$). However, Bland–Altman analysis revealed progressive increases in bias and data dispersion from phototype II to phototype IV. The overall analysis also demonstrated high correlations (SBP: $r = 0.957$; DBP: $r = 0.972$; HR: $r = 0.919$; $p < 0.001$), although agreement decreased due to greater variability, particularly for HR. These findings suggest that skin phototype influences the performance of photoplethysmography-based smartwatch measurements and should be considered during the validation of wearable devices and the development of more robust algorithms for diverse populations.

1. Introduction

Wearable devices equipped with photoplethysmography (PPG) sensors have become increasingly used for continuous physiological monitoring, particularly in cardiovascular applications. These devices enable the non-invasive assessment of physiological parameters, including heart rate, blood pressure estimates, sleep metrics, and anthropometric data, supporting telemedicine, population health screening, and long-term health monitoring [1], [2], [3].

Photoplethysmography estimates physiological parameters by detecting changes in light absorption associated with peripheral blood volume. However, signal quality may be affected by physiological and physical factors, including peripheral perfusion, motion artifacts, and skin pigmentation. Higher melanin concentrations can modify light absorption and scattering, potentially reducing measurement accuracy [3], [4].

Recent studies have reported potential biases in optical sensing technologies associated with skin pigmentation. Although most evidence is related to pulse oximetry, limited data are available regarding the influence of skin phototype on smartwatch-derived blood pressure and heart rate measurements compared with clinical reference methods [5].

As wearable devices become increasingly integrated into clinical and personal health monitoring, validation across different Fitzpatrick skin phototypes is essential to assess algorithm robustness and measurement reliability under varying optical conditions [6].

Therefore, the present study evaluates smartwatch-derived blood pressure and heart rate measurements using a stratified approach based on Fitzpatrick skin phototypes. This methodology allows for the assessment of the potential influence of skin pigmentation on the accuracy and agreement of physiological measurements.

1.1. Hypothesis and objective

This study hypothesizes that skin pigmentation may affect the performance of physiological measurements obtained by a smartwatch when compared with a reference instrument. Accordingly, the aim was to assess the correlation and agreement between smartwatch-derived measurements and those obtained using a sphygmomanometer, both in the overall sample and across Fitzpatrick skin phototypes II, III, and IV, to investigate the influence of skin pigmentation on measurement

performance.

2. Materials and methods

2.1. Materials

The study was conducted at the Polyclinic of the University of Mogi das Cruzes (UMC), São Paulo, Brazil. Physiological measurements were obtained using a Samsung Galaxy Watch 4 smartwatch equipped with an optical photoplethysmography (PPG) sensor and a sampling frequency of 100 Hz. An automated sphygmomanometer (OMRON HEM-7320) was used as the reference method. Measurements obtained from both devices were compared to evaluate the correlation and agreement of blood pressure (BP) and heart rate (HR) measurements.

2.2. Sample

study sample consisted of 48 participants equally distributed across Fitzpatrick skin phototypes II, III, and IV (n = 16 per group). The cohort included 26 women and 22 men, with a mean age of 29.0 ± 10.3 years (range: 22–61 years). Only individuals without comorbidities or physical limitations that could interfere with physiological measurements were included. Participants were recruited voluntarily and provided written informed consent prior to data collection. The study protocol was approved by the Research Ethics Committee under CAAE No. 84992424.4.0000.5497.

2.3. Protocol

Measurements were performed simultaneously using the smartwatch and the reference sphygmomanometer, with participants seated and at rest. The cuff was placed on the right arm and the smartwatch on the contralateral wrist. Systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR) were recorded and subsequently analyzed both globally and stratified according to Fitzpatrick skin phototypes.

2.4. Statistical Analysis

Statistical analyses were conducted using MATLAB® (MathWorks Inc., Natick, MA, USA; R2017a). Pearson's correlation coefficient was used to assess the association between smartwatch and sphygmomanometer measurements. Agreement between methods was evaluated using Bland–Altman analysis, including calculation of the mean bias and 95% limits of agreement (mean difference ± 1.96 standard deviations).

Additionally, proportional bias was assessed by linear regression of the differences against the mean values. Analyses were performed for systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR), both in the overall sample and stratified by Fitzpatrick skin phototypes II, III, and IV. Statistical significance was defined as $p < 0.05$.

3. Results

Analyses were first performed separately for each Fitzpatrick skin phototype and subsequently conducted on the pooled sample.

3.1. Phototype II

Analysis of the phototype II group revealed strong correlations between smartwatch and sphygmomanometer measurements for systolic blood pressure (SBP) ($r = 0.987$, $p < 0.001$) and diastolic blood pressure (DBP) ($r = 0.989$, $p < 0.001$). Heart rate (HR) also showed a significant correlation ($r = 0.821$, $p < 0.001$), although with greater data dispersion.

Bland–Altman analysis demonstrated low positive bias for SBP (*+1.25 mmHg*), DBP (*+0.50 mmHg*), and HR (*+1.26 bpm*), indicating a slight overestimation by the smartwatch. The limits of agreement were narrow, with a homogeneous distribution of residuals and a slight proportional bias observed only for HR.

3.2. Phototype III

Analysis of the phototype III group revealed strong correlations between smartwatch and sphygmomanometer measurements for systolic blood pressure (SBP) ($r = 0.980$, $p < 0.001$), diastolic blood pressure (DBP) ($r = 0.965$, $p < 0.001$), and heart rate (HR) ($r = 0.950$, $p < 0.001$). However, greater data dispersion was observed compared with phototype II, particularly for HR.

Bland–Altman analysis demonstrated increased bias for SBP (*+3.38 mmHg*), DBP (*+1.88 mmHg*), and HR (*+3.75 bpm*), indicating reduced agreement between the methods. The limits of agreement were wider, with proportional bias observed mainly for DBP and HR, as well as mild heteroscedasticity for HR.

3.3. Phototype IV

Analysis of the phototype IV group revealed a lower correlation for systolic blood pressure (SBP) ($r = 0.883$, $p < 0.001$), accompanied by greater data dispersion. In contrast, diastolic blood pressure (DBP) and heart rate (HR) maintained strong correlations ($r = 0.974$ for both, $p < 0.001$), although with greater variability.

Bland–Altman analysis demonstrated greater data dispersion, particularly for SBP, with moderate positive

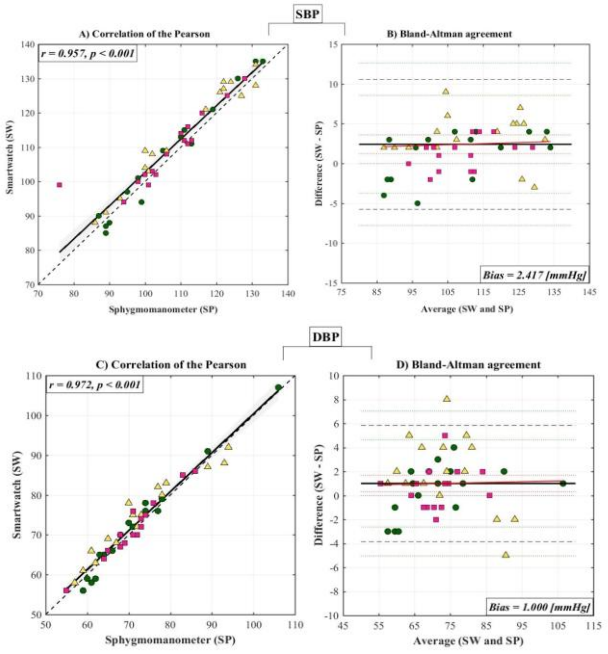
bias for SBP (*+2.63 mmHg*), DBP (*+0.63 mmHg*), and HR (*+0.69 bpm*). Wider residual dispersion and proportional bias were observed for SBP, suggesting reduced agreement and measurement stability between the methods.

3.4. Global Analysis

In the overall analysis, including all participants regardless of skin phototype, strong correlations were observed between the smartwatch and the sphygmomanometer for systolic blood pressure (SBP: $r = 0.957$), diastolic blood pressure (DBP: $r = 0.972$), and heart rate (HR: $r = 0.919$; $p < 0.001$ for all). However, pooling the phototype groups resulted in greater data dispersion, reflecting increased heterogeneity across the sample, particularly for HR (Figure 1A, C, and E).

Bland–Altman analysis demonstrated a consistent positive bias for SBP (*+2.417 mmHg*), DBP (*+1.000 mmHg*), and HR (*+1.896 bpm*), with wider limits of agreement, greater data dispersion, and slight proportional bias, particularly for HR. Although the residuals remained centered around the mean bias, their less homogeneous distribution indicates increased inter-individual variability (Figure 1B, D, and F).

Overall, these findings indicate that, despite the high overall correlation, agreement between the methods varied according to skin phototype. Measurement accuracy decreased with increasing phototype, as reflected by greater variability and bias.



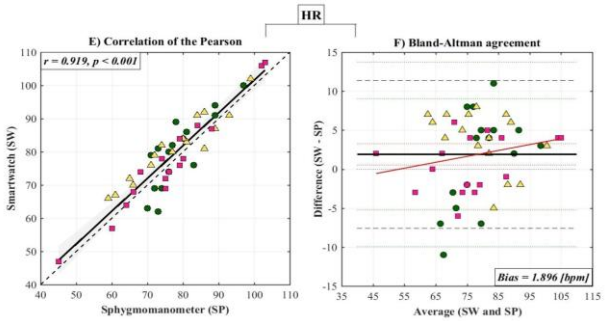


Figure 1: Pearson correlation (A–C–E) and Bland–Altman agreement (B–D–F) analyses of SBP, DBP, and HR measurements obtained using the smartwatch (SW) and sphygmomanometer (SP).

Phototype II Phototype III Phototype IV

4. Discussion

The results demonstrate that, although wearable devices showed strong correlations with the reference method, agreement between measurements was not consistent across skin phototypes, highlighting the importance of complementing correlation analyses with agreement assessment. Overall, diastolic blood pressure (DBP) showed the greatest stability between methods, whereas heart rate (HR) exhibited greater variability. In addition, a progressive increase in data dispersion and bias was observed from phototype II to phototype IV, with systolic blood pressure (SBP) showing the lowest measurement consistency in the darker phototypes. These findings suggest that physiological and optical factors may influence the performance of photoplethysmography (PPG)-based devices [1], [2], [7], [8].

The positive bias observed across all groups indicates a systematic tendency for the smartwatch to overestimate measurements. Although the bias values were relatively low, the wider limits of agreement and the presence of proportional bias, particularly in the darker phototypes, suggest that measurement error varied across the measurement range. Furthermore, the heteroscedasticity observed for heart rate (HR), especially in phototype III, indicates that measurement accuracy may be influenced by physiological signal intensity and the quality of the optical signal [1], [7].

From a physiological perspective, these findings may be explained by the interaction between light and biological tissues. Higher melanin concentrations in darker skin phototypes can alter the absorption and scattering of light emitted by optical sensors, affecting the quality of the photoplethysmography (PPG) signal and, consequently, the accuracy of physiological measurements. This mechanism may explain the greater variability observed in phototypes III and IV, as well as the reduced measurement consistency for specific variables [9].

Another important finding was the distinction between correlation and agreement. Although correlation coefficients remained high across all skin phototypes, Bland–Altman analysis revealed discrepancies between the methods, demonstrating that a strong correlation does not necessarily indicate good agreement. These findings highlight the importance of combining both approaches in validation studies. Furthermore, the variability observed across skin phototypes suggests that smartwatch performance may not be consistent across different populations, emphasizing the need for caution when interpreting their clinical applicability [2], [8].

Finally, these findings emphasize the importance of considering individual characteristics, particularly skin phototype, in the validation of wearable devices, as they may directly affect measurement performance. Future studies should further investigate these effects and support the development of more robust and inclusive algorithms capable of providing reliable measurements across diverse populations [4], [6], [10].

Future work will include a larger sample size, and the inclusion of additional Fitzpatrick skin phototypes (I, V, and VI). Expanding the study will allow a more comprehensive assessment of the performance of systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR) measurements across different skin phototypes.

Acknowledgements

D.G.G. gratefully acknowledges the financial support of the São Paulo Research Foundation (FAPESP), Brazil (Grant No. 2021/14231-0). A.L.G.d.S. acknowledges financial support from the National Council for Scientific and Technological Development (CNPq), Brazil (Grant No. 164066/2024-1). K.C.B.S. acknowledges financial support from the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil (Grant No. 88887.941749/2024-00). The authors also thank Dr. Melquíades M. Portela and Jessica Lelé for their valuable assistance and support throughout the project.

References

[1] B. Nelson e S. Singh, “The Impact of Skin Pigmentation on Pulse Oximetry SpO2 and Wearable Pulse Rate Accuracy: A Meta-Analysis”.

[2] F. Sarhaddi *et al.*, “A comprehensive accuracy assessment of Samsung smartwatch heart rate and heart rate variability”, *PLOS ONE*.

[3] P. H. Charlton *et al.*, “The 2023 wearable photoplethysmography roadmap”, *Physiol. Meas.*, v. 44, n. 11, p. 111001, nov. 2023, doi: 10.1088/1361-6579/acead2.

[4] A. Ometov, “A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges”, *Computer Networks*, 2021.

[5] M. J. Butler, J. A. Crowe, B. R. Hayes-Gill, e P. I. Rodmell, “Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin”, *Physiol. Meas.*, v. 37, n. 5, p. N27–N37, maio 2016, doi: 10.1088/0967-3334/37/5/N27.

[6] K. B. Kim e H. J. Baek, “Photoplethysmography in Wearable Devices: A Comprehensive Review of Technological Advances, Current Challenges, and Future Directions”, 2023.

[7] P. E. Bickler, “Effects of Skin Pigmentation on Pulse Oximeter Accuracy at Low Saturation”, *Anesthesiology*, v. 102, n. 4, 2005, doi: 10.1097/00000542-200504000-00004.

[8] C. Hahnen *et al.*, “Accuracy of Vital Signs Measurements by a Smartwatch and a Portable Health Device: Validation Study”, *JMIR Mhealth Uhealth*, v. 8, n. 2, p. e16811, fev. 2020, doi: 10.2196/16811.

[9] P. H. Charlton, P. A. Kyriacou, J. Mant, V. Marozas, P. Chowienczyk, e J. Alastruey, “Wearable Photoplethysmography for Cardiovascular Monitoring”, *Proc. IEEE*, v. 110, n. 3, p. 355–381, mar. 2022, doi: 10.1109/JPROC.2022.3149785.

[10] R. H. Verzani e A. B. D. S. Serapião, “Contribuições tecnológicas para saúde: olhar sobre a atividade física”, *Ciênc. saúde coletiva*, v. 25, n. 8, p. 3227–3238, ago. 2020, doi: 10.1590/1413-81232020258.19742018.

Address for correspondence:

Daniel Gustavo Goroso, Ph.D.
Biological Systems Modeling and Signal Processing Laboratory
Hospital Policlínica, University of Mogi das Cruzes (UMC); Street Dom Antônio Cândido de Alvarenga, 170, Mogi das Cruzes, SP 08780-911, Brazil
E-mail: danielg@umc.br