Respiratory Rate Accuracy Report of SDK 5.9.1

Executive Summary

Goal

This document evaluates the accuracy of Respiratory Rate in SDK 5.9 [iOS and Android] rPPG with reference devices, using data collected in Israel, India, South Africa and Japan.

Results

The Respiratory rate measured by Binah's SDK was found to be **highly accurate and within the accuracy target (AE≤3 rpm)** in >95.9% of the measurements for iOS and Android and the following confounding factors (see appendix):

- Both female and male
- All skin tones (Fitzpatrick I to VI)
- Ages 18 to 77
- BMI from light to morbid obesity
- Distances close and far from the face
- Luminance from dark to brighter surroundings
- Various face angles, from wide to narrow
- Similar performance in several countries with different ethnicities
- Similar performance on all devices used for recordings.

Conclusions

The respiratory rate measured by Binah's SDK was found to be robust, highly accurate, and within the accuracy target (AE ≤ 3 rpm) in **over 95.9%** of measurements across both iOS and Android operating systems.

Introduction

The main function of the respiratory system is gas exchange. Oxygen is transferred from the environment into the bloodstream, while carbon dioxide is expelled. When inhaling, air passes into the lungs, where gas exchange occurs as oxygen diffuses into the lung capillaries in exchange for carbon dioxide. Exhalation follows this exchange, allowing air containing carbon dioxide to return to the external environment through the nose or mouth.

In addition to gas exchange, the respiratory system performs several secondary functions, including filtering, warming, and humidifying inhaled air.[1],[2] There is a close relationship between respiration and heart activity. Heart rate is regulated by respiration, increasing during inhalation and decreasing during exhalation.[3]

Respiratory Rate (RR), defined as the number of respirations per minute, is a clinical parameter that reflects ventilation — the movement of air in and out of the lungs.[1] While RR varies from person to person, the normal resting rate typically ranges between 12–20 respirations per minute.[4]

RR is a valuable diagnostic and prognostic marker of health, widely used in clinical settings to identify abnormalities.[5] In hospital care, it is recognized as a highly sensitive marker of acute deterioration.[6] For instance, an elevated RR is a known predictor of cardiac arrest[7] and in-hospital mortality,[8] and can indicate respiratory dysfunction.[9]

Despite its clinical importance, RR is still commonly measured by manually counting chest wall movements outside of intensive care settings. This method is often time-consuming, inaccurate,[10],[11] and poorly executed.[12],[13] Consequently, there is a significant need for a non-intrusive, automatic method of measuring RR. The advantage of such a method for monitoring vital signs is unquestionable.

Binah.ai's algorithm leverages photoplethysmography (PPG) technology recorded from facial skin tissue using a remote PPG (rPPG) method. The algorithm extracts face video images, generates an rPPG signal, analyzes the data, and provides the end user with real-time vital sign measurements.

This report presents the results of accuracy studies conducted in Israel, India, South Africa, and Japan, comparing Binah.ai's vital sign measurements with those obtained from approved reference devices.

Methods

Binah.ai's RR - measurements were compared to the Vernier Go Direct® Respiration Belt and/or finger pulse oximeter measurements Masimo finger pulse oximeter measurements. The experiments were conducted in Israel, India, South Africa, and Japan with both healthy participants and participants with a medical background.

Measurement set-up:

At all sites, participants were instructed to sit as still as possible. Recordings were conducted in a controlled testing room with fixed artificial ambient lighting.

The respiration rate reference devices used included the Vernier Go Direct® Respiration Belt and/or the Masimo finger pulse oximeter, which were placed on each participant's finger to measure RR.

For rPPG measurements, a mobile device was placed on a stand or held by the participant's hand in front of them. Participants were positioned so that their face filled most of the camera frame (approximately 20–40 cm distance) and centered within the frame. The camera was set at forehead level and positioned perpendicular to the face. Participants were instructed to look at the screen throughout the recording. Participants were also instructed to remove their glasses, remain still, and avoid movement, including talking. They were required to sit with their feet flat on the floor. Each recording lasted 60 seconds.

Statistical analysis:

Accuracy was calculated using the following parameters:

$$AE\ (Absolute\ Error) = |App_i - Ref_i|$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (App_i - Ref_i)^2}{N}}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |App_i - Ref_i|$$

When,
N is the number of data points.
App is the measurement of the Binah.ai application.
Ref is the measurement of the reference device.
i is the index number of measurements.
Participants with invalid reference device values and with very low signal quality were excluded from the analysis.

For this report, Binah.ai's SDK 5.9 was compared to a reference device.
The measurements were recorded in several locations in Israel, India, South Africa, and Japan using the mobile device models listed below:

- iOS: iPhone 11 Pro, iPhone 13, iPhone 13 Pro Max, iPhone 14 and iPhone 14 Pro Max
- Android: Samsung S21 Ultra, Samsung S23 Ultra, Pixel 6 Pro

Accuracy criteria:

RR: AE ≤ 3 rpm in 85% of measurements.

Results

Measurement disposition

Number of subjects/measurements with reported respiration rate: 880/3232

Number of Unique Subjects and Measurements by Country and Respiration Rate Distribution

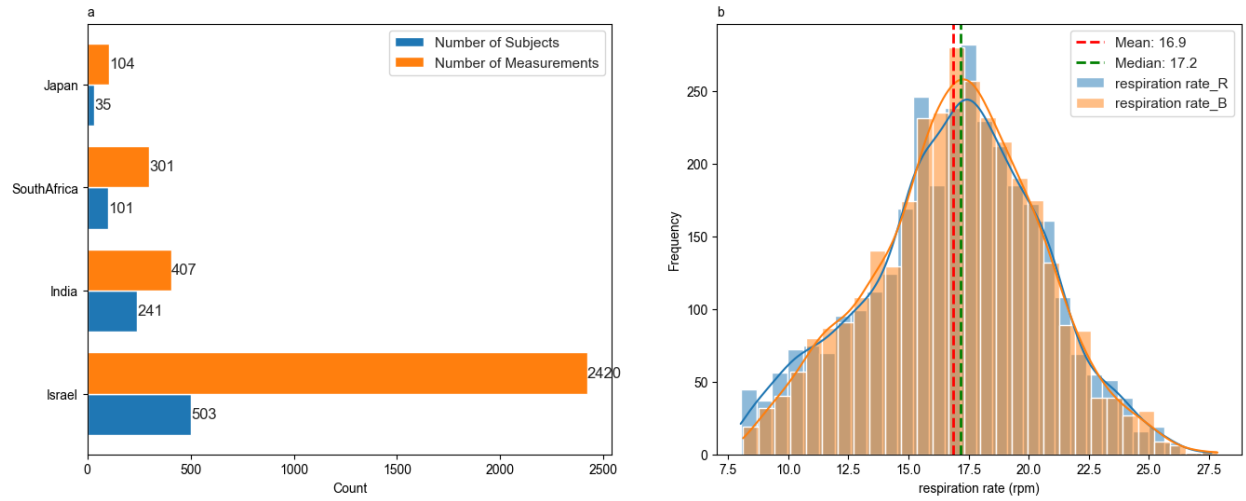

Figure 1:

a. Number of Unique Subjects and Measurements by Country data presented includes all measurements with reference values.

b. Distribution of Respiration rate measured by reference device and Binah.ai's application, both measurements present overlapped normal distribution.

Demographics Data:

Subjects/Measurements	Age (mean ± std)	BMI (mean ± std)	Sex (F/M)	
880/ 3232	35.7 ± 12.5	24.6 ± 4.7	447 / 433	
Fitzpatrick Skin Tone (I/II/III/IV/V/VI)	**Beard (No/Yes)**	**Glasses (No/Yes)**	**Face cream (No/Yes)**	
1 / 261 / 265 / 245 / 69 / 39	285 / 188	351 / 14	327 / 146	
Distance (mean ± std)	**Luminance (mean ± std)**	**Angle yaw (mean ± std)**	**Angle roll (mean ± std)**	**Angle pitch (mean ± std)**
0.29 ± 0.05	80.7 ± 62.8	5.0 ± 3.8	2.0 ± 1.5	11.1 ± 7.5

Table 1: Demographic data for experiments using phones with iOS and Android operating systems.

* Fitzpatrick skin tone classifications are I- Pale white, II- white, III- Darker white, IV- Light brown, V- Brown, VI- Dark brown or black.

Accuracy Data:

OS	Unique Subjects	Measurements	MAE ± STD	Ref Range
iOS	802	1598	0.8 ± 0.9	8.0 - 28.0
Android	806	1634	0.9 ± 0.9	8.0 - 28.0

Table 2: Accuracy data for iOS and Android (MAE ± STD) compared to the reference device in the presented respiration rate range.
MAE - Mean Absolute Error, STD - Standard Deviation

Correlation and Bland-Altman plot by Operating System

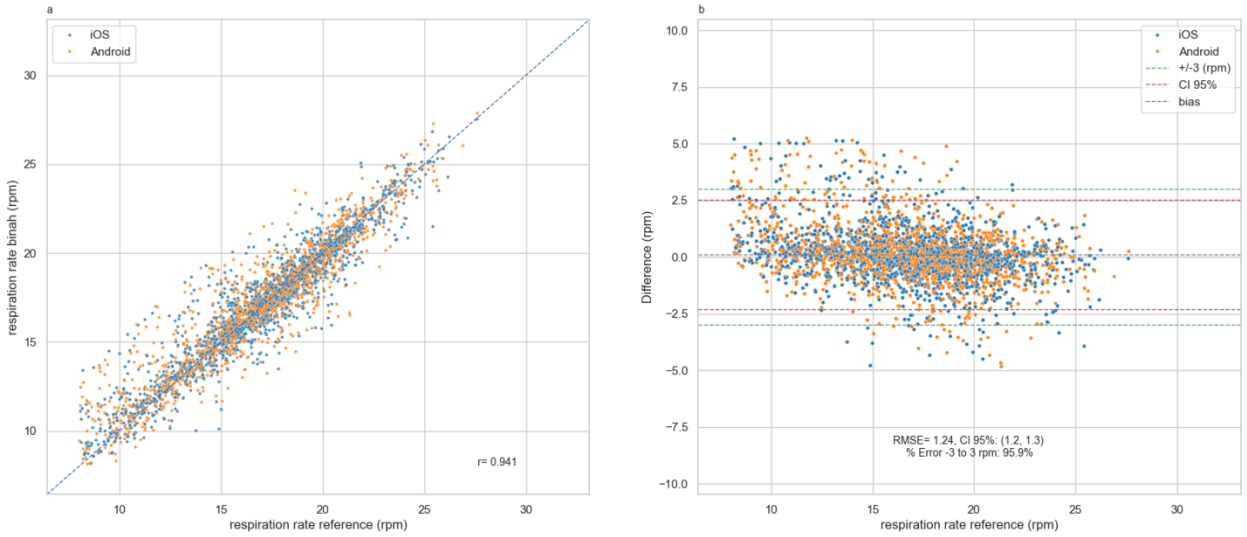

Figure 2:
a. Correlation plot by operating system - Binah.ai's RR estimations versus reference device RR measurements were found to have a correlation of r= 0.941 for both operating systems (Android and iOS).

b. Bland-Altman plot by operating system - Bland-Altman plots for comparison between RR measurements of the two methods (Binah's and the reference device) demonstrated agreement between the two devices (95.9% of the measurements are within target error) for both operating systems (Android and iOS) in the presented respiration rate range.

The "Bias" gray dashed line stands for the mean difference between measurements of Binah.ai and the reference device, the "Error" green dashed lines of ±3 rpm represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Conclusions

This report summarizes the results of an accuracy analysis in which the respiratory rate measured by Binah.ai's SDK was **found to be robust, highly accurate, and within the accuracy target (AE ≤ 3 rpm)** in **95.9%** of measurements across both iOS and Android operating systems.

References

1. Ali, M., Elsayed, A., Mendez, A., Savaria, Y. & Sawan, M. Contact and remote breathing rate monitoring techniques: A review. *IEEE Sens. J.* (2021).
2. Tu, J., Inthavong, K. & Ahmadi, G. *Computational fluid and particle dynamics in the human respiratory system*. (Springer Science & Business Media, 2012).
3. Pallas-Areny, R., Colominas-Balague, J. & Rosell, F. J. The effect of respiration-induced heart movements on the ECG. *IEEE Trans. Biomed. Eng.* **36**, 585–590 (1989).
4. Barrett, K. E., Barman, S. M. & Boitano, S. Ganong's review of medical physiology McGraw-Hill Education. (2019).
5. Charlton, P. H. *et al.* Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review. *IEEE Rev. Biomed. Eng.* **11**, 2–20 (2017).
6. Cretikos, M. *et al.* The objective medical emergency team activation criteria: a case–control study. *Resuscitation* **73**, 62–72 (2007).
7. Schein, R. M. H., Hazday, N., Pena, M., Ruben, B. H. & Sprung, C. L. Clinical antecedents to in-hospital cardiopulmonary arrest. *Chest* **98**, 1388–1392 (1990).
8. Duckitt, R. W. *et al.* Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. *Br. J. Anaesth.* **98**, 769–774 (2007).
9. Gravelyn, T. R. & Weg, J. G. Respiratory rate as an indicator of acute respiratory dysfunction. *Jama* **244**, 1123–1125 (1980).
10. Philip, K. E. J. *et al.* The accuracy of respiratory rate assessment by doctors in a London teaching hospital: a cross-sectional study. *J. Clin. Monit. Comput.* **29**, 455–460 (2015).
11. Lovett, P. B., Buchwald, J. M., Stürmann, K. & Bijur, P. The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Ann. Emerg. Med.* **45**, 68–76 (2005).
12. Pimentel, M. A. F., Charlton, P. H. & Clifton, D. A. Probabilistic estimation of respiratory rate from wearable sensors. in *Wearable electronics sensors* 241–262 (Springer, 2015).
13. Philip, K., Richardson, R. & Cohen, M. Staff perceptions of respiratory rate measurement in a general hospital. *Br. J. Nurs.* **22**, 570–574 (2013).

Appendix

Respiration Rate Error by Age and BMI

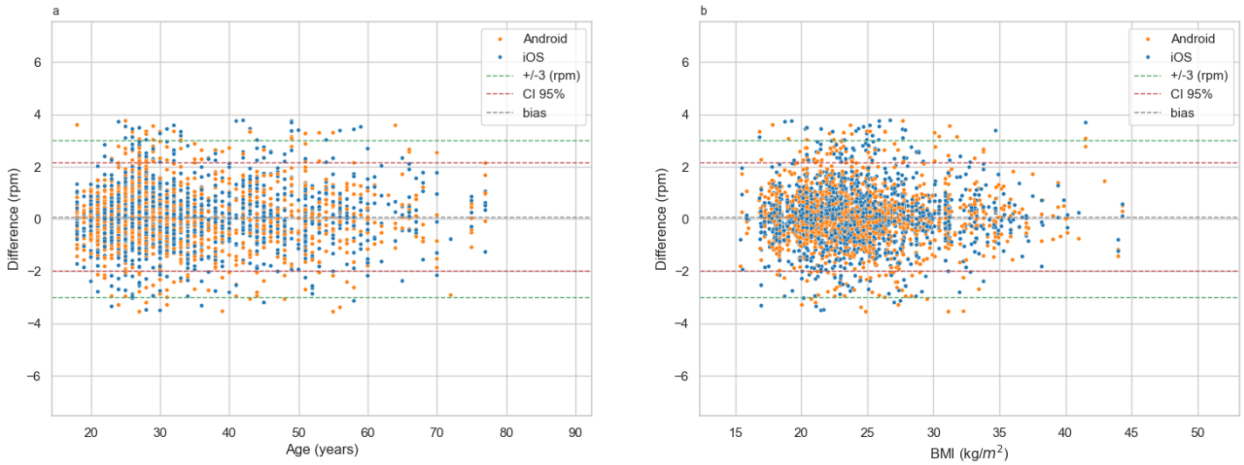

Figure 3:

a. Bland-Altman plot by Age - demonstrated RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented Age range.

b. Bland-Altman plot by BMI - demonstrated RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented BMI range from low to very high.

The "Bias" gray dashed line stands for the mean difference between measurements of Binah.ai and the reference device, the "Error" green dashed lines of ±3 rpm represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Respiration Rate error by skin tone with Gender and Operating system

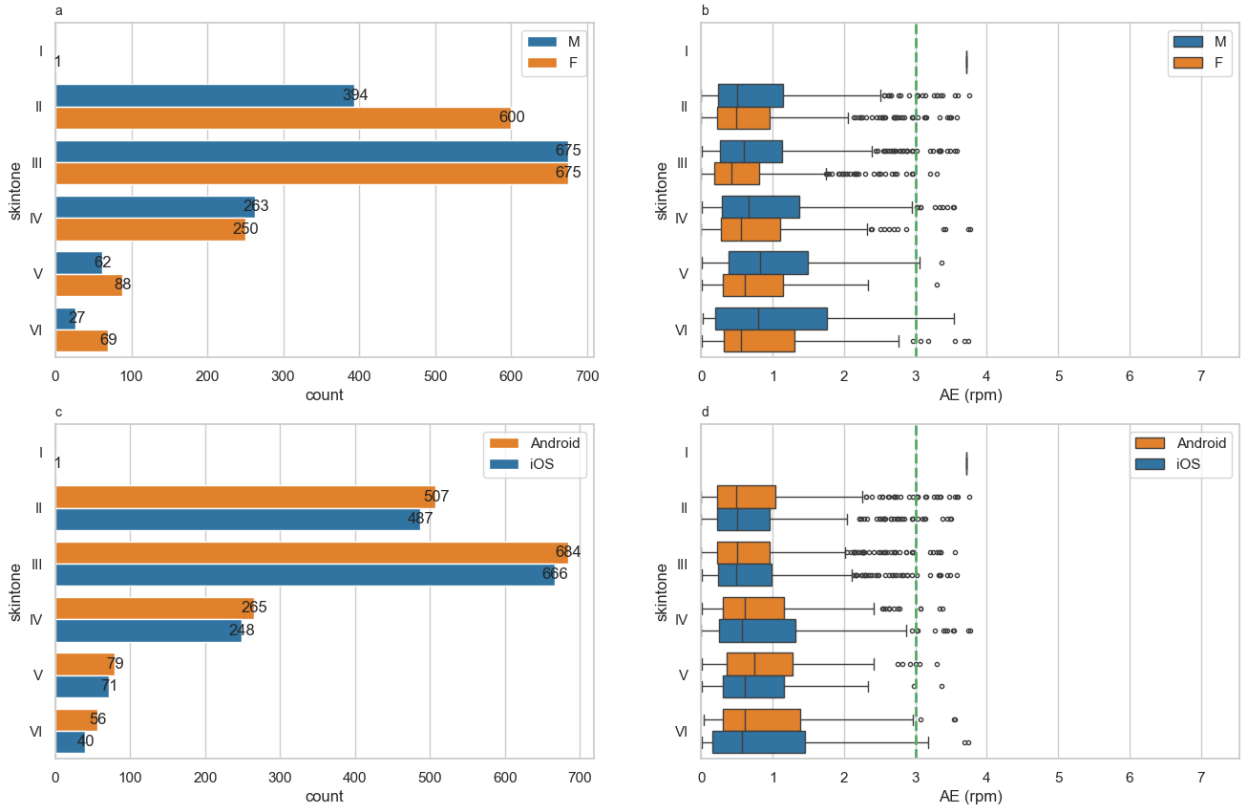

Figure 4:

a. Number of measurements by Fitzpatrick skin tone and sex (female and male).

b. Box plot by Fitzpatrick Skin Tone and Gender – RR measurements obtained by Binah.ai's compared to the reference device are accurate for both genders (female and male) across most of the presented skin tones.
The green dashed "Error" lines set at ≤3 rpm represents the value of the accuracy criterion.

c. Number of measurements by Fitzpatrick skin tone and operating system (Android and iOS).

d. Box plot by Fitzpatrick skin tone and operating system - RR measurements obtained by Binah.ai's versus the reference device are highly accurate for both operating systems (Android and iOS) across all presented skin tones.

The green dashed "Error" lines set at ≤3 rpm represents the value of the accuracy criterion.

Respiration Rate error by Distance and Luminance

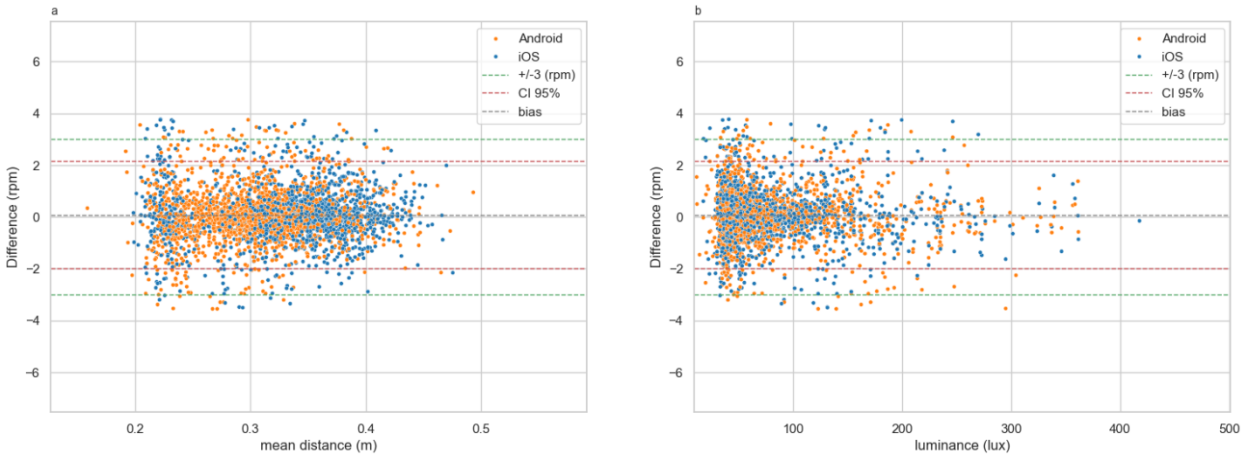

Figure 5:

a. Bland-Altman plot by Distance (m) - demonstrated agreement between RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented distance range between the camera and the subject's face.

b. Bland-Altman plot by Luminance (lux) - demonstrated agreement between RR measurements obtained by Binah.ai's and the reference device for both operating systems (Android and iOS) within the presented luminance range from dark surroundings (<150 lux) to brighter ones.

The "Bias" gray dashed line stands for the mean difference between measurements of Binah.ai and the reference device, the "Error" green dashed lines of ±3 rpm represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Respiration Rate Error by Face Angles

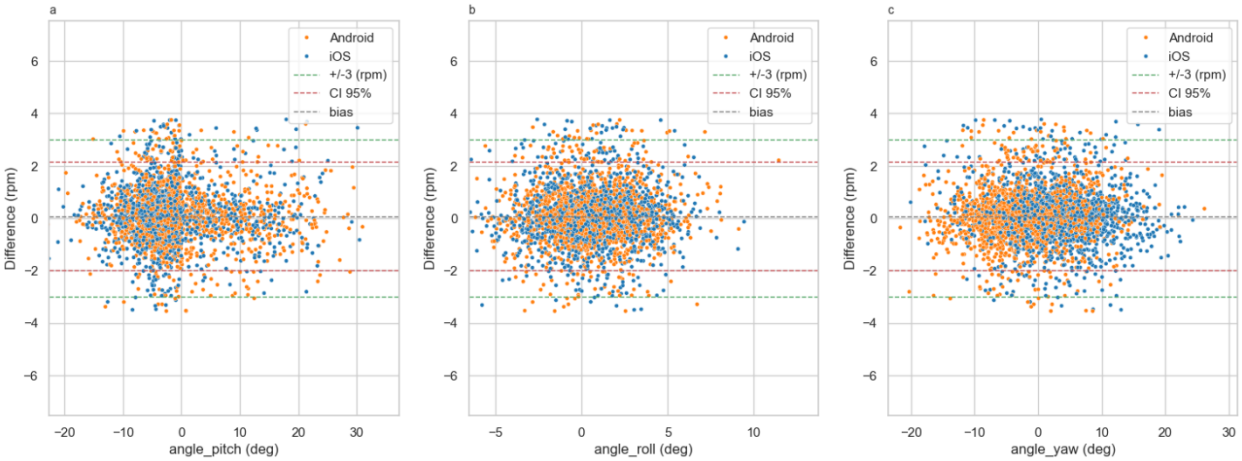

Figure 6:

a. **Bland-Altman plot by pitch angle (deg)** - demonstrated RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented pitch angle range.

b. **Bland-Altman plot by roll angle (deg)** - demonstrated RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented roll angle range.

c. **Bland-Altman plot by yaw angle (deg)** - demonstrated RR measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented yaw angle range.

The gray dashed "Bias" line stands for the mean difference between measurements of Binah.ai and the reference device, the green dashed "Error" lines set at ±3 rpm represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Respiration Rate Error by Devices

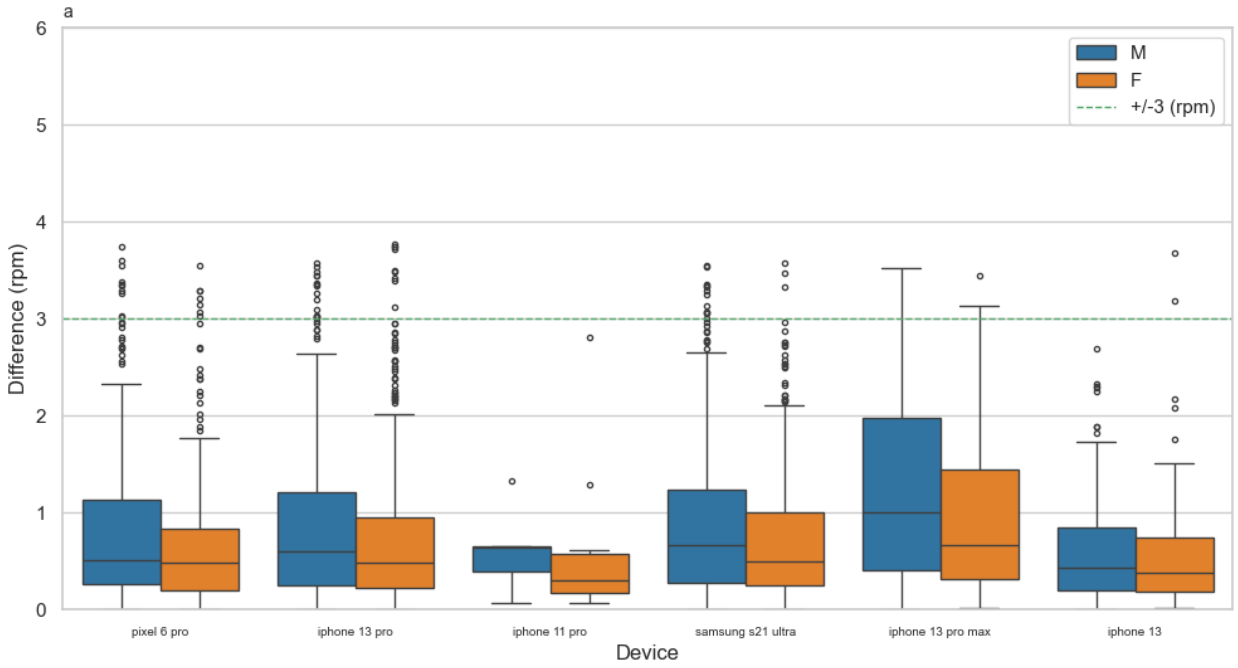

Figure 7:

a. Box plot by device - RR measurements obtained by Binah.ai versus the reference device for both sexes (female and male) on all devices.

The green dashed "Error" lines set at ≤3 rpm represents the value of the accuracy criterion.

Respiration Rate Error by Country by OS

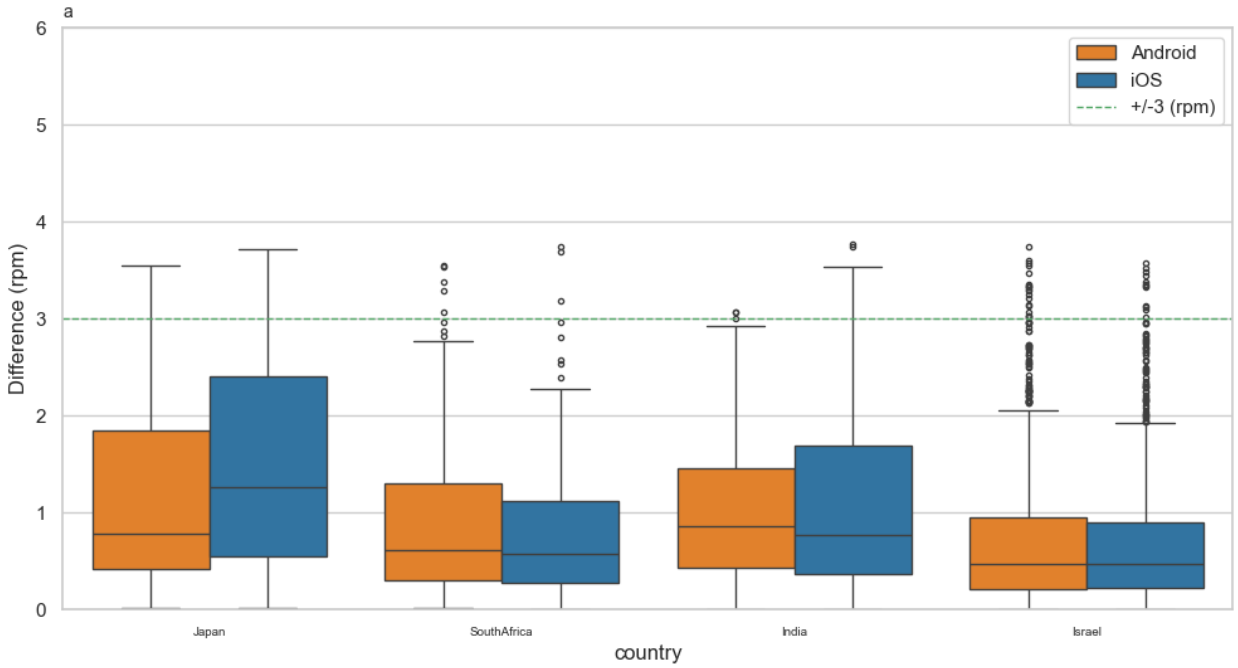

Figure 8:

a. Box plot by country - RR measurements obtained by Binah.ai versus the reference device for both operation systems (Android and iOS) in all countries.

The green dashed "Error" lines set at ≤3 rpm represents the value of the accuracy criterion.