Heart Rate Variability Accuracy Report of SDK 5.9.1

Executive Summary

Goal

This document evaluates the accuracy of Heart Rate Variability (HRV) in SDK 5.9 [iOS and Android] rPPG with reference devices, using data collected in Israel, India, Nepal, South Africa and Italy.

Results

The HRV (mean RRi) measured by Binah's SDK demonstrated high accuracy, consistently meeting the accuracy target (AE ≤ 25 ms) in 100% of measurements across both iOS and Android platforms, regardless of the following confounding factors (see appendix):

- Both female and male
- All skin tones (Fitzpatrick II to VI)
- Ages 18 to 83
- 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 on all devices used for recordings
- Similar performance in several countries with different ethnicities

Conclusions

The mean RRi HRV measured by Binah's SDK was **found to be robust and highly accurate, consistently meeting the accuracy target (AE ≤ 25 ms) in 100%** of the measurements across both iOS and Android operating systems.

Introduction

Heart rate (HR) is the number of heart beats per minute. **Heart Rate Variability (HRV)** is the fluctuation in the time intervals between adjacent heartbeats (RR intervals – RRi).[1] In other words, the variability of RRi is known as the Heart Rate Variability (HRV). HRV is generated by heart-brain interactions and autonomic nervous system processes. HRV helps the body adapt to environmental and psychological challenges. It reflects the regulation of autonomic balance, blood pressure, blood vessel diameter, gas exchange, gut, and heart.[2] A healthy heart is not a metronome, and heartbeats do not occur at constant intervals, but rather with a small variance between them.[3] This variability in heartbeats provides the flexibility to rapidly cope with an uncertain and changing environment. [4]

Physical or emotional stress results in faster, monotonic heartbeats, causing HRV to decrease. Relaxation results in slower, less regular heartbeats, and higher HRV.[5] Normal HRV is associated with a lower risk of developing depression and post-traumatic stress disorder.[6,7] Moreover, decreased HRV has been identified as an independent predictor of cardiovascular and overall mortality.[8–10] Thus, HRV is a noninvasive method that can be used to evaluate autonomic nervous system activity and physical and emotional status in a variety of clinical situations. [11,12].Therefore, the advantage of a non-intrusive, automatic, and accessible method for monitoring this vital sign is unquestionable.

Binah.ai's algorithm utilizes the photoplethysmography (PPG) signal recorded from facial skin tissue (remote PPG - rPPG) to extract and analyze vital signs in real-time. It processes face video images to generate an rPPG signal, identifies heartbeat peaks corresponding to the contraction of heart ventricles (R peaks of the QRS complex in the ECG wave), and delivers accurate physiological measurements to the end user.

Binah.ai's HRV measurements are based on various parameters calculated from RRi values, such as, SDNN [in msec] which represents the standard deviation of RRi, or RMSSD [in msec] which represents the Root Mean Square of Successive Differences of RRi.

This report describes the results of accuracy studies conducted in Israel, India, South Africa, Nepal, and Italy that compare Binah.ai's HRV measurements with the measurements of approved reference devices.

Methods

Binah.ai's HRV measurements were validated in comparison to the Polar H10 Heart Rate SensorTM / Polar Verity Sense measurements. The experiments were conducted in Israel, India, South Africa, Nepal, and Italy with both healthy participants and participants with a medical background.

Measurement set-up:

In all sites, each participant was instructed to sit as stable as possible. Recordings were conducted in a testing room, with controlled and fixed artificial ambient light.
The Pulse rate reference devices that were used included: the Polar H10 Heart Rate Sensor™ or Polar Verity Sense which were placed on each participant's finger to measure HRV.

For rPPG measurements, a mobile device was placed on a stand in front of the participant. The participant's face filled most of the frame's area (distance of about 20-40 cm) and was positioned in the center of the frame. The camera was set at the level of the forehead and positioned perpendicular to the face. Participants were instructed to look at the screen throughout the recording.
Participants were instructed to take off their glasses and to avoid any movement, including talking, and were required to sit still 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 participants with very low signal quality were excluded from the analysis.

Definitions:

- RR intervals (RRi) are defined as the difference between successive R peaks.
- RR intervals are calculated as:

RR(n) = R(n)−R(n−1), where n is the beat index number.

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, Nepal and, Italy using the mobile device models listed below:

- iOS: iPhone 11 Pro, iPhone 13, iPhone 13 Pro and iPhone 13 Pro Max, iPhone 14 pro, iPhone 14 pro max.
- Android: Samsung S21 Ultra, Samsung S22 Ultra, Samsung S23 Ultra, Pixel 6 Pro, Samsung S21 FE.

Accuracy criteria:

HRV: AE ≤ 25 ms in 90% of measurements.

Results

Mean RRi

Measurement disposition

Number of subjects/measurements with reported HRV: 1377/2208

Number of Unique Subjects and Measurements by Country and mean RRi distribution

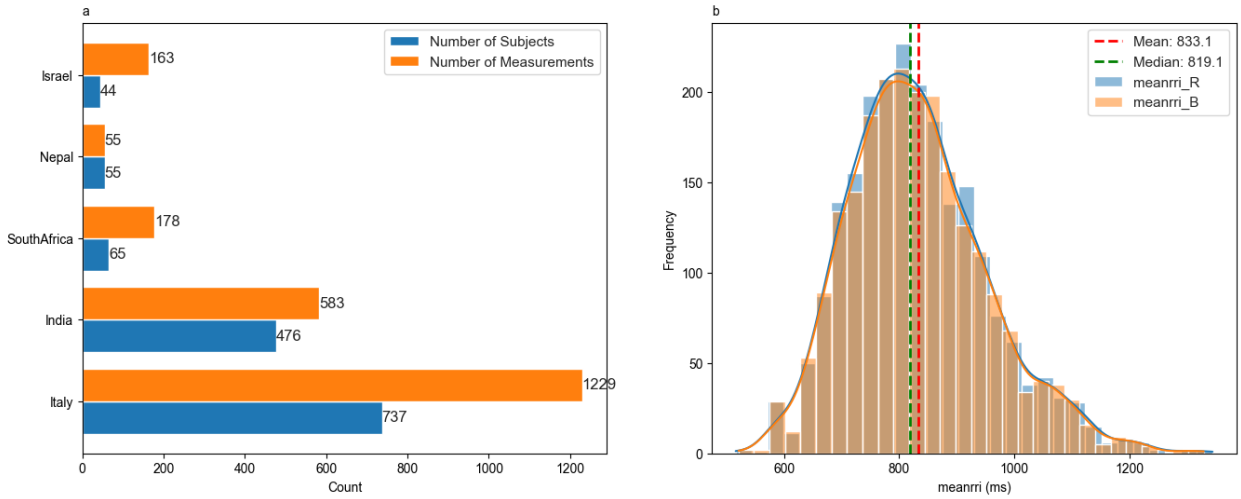

Figure 1:
a. Number of Unique Subjects and Measurements by Country data presented includes all measurements with reference values.
b. Distribution of mean RRi measured by reference device and Binah.ai's application, both measurements present overlapped normal distribution.

Demographics Data:

Subjects/Measurements	Age (mean ± std)	BMI (mean ± std)	Sex (F/M)	
1377/ 2208	51.1 ± 15.1	27.3 ± 5.2	627 / 750	
Fitzpatrick Skin Tone (I/II/III/IV/V/VI)	**Beard (No/Yes)**	**Glasses (No/Yes)**	**Face cream (No/Yes)**	
43 / 501 / 217 / 506 / 85 / 25	876 / 501	1039 / 273	1132 / 245	
Distance (mean ± std)	**Luminance (mean ± std)**	**Angle yaw (mean ± std)**	**Angle roll (mean ± std)**	**Angle pitch (mean ± std)**
0.27 ± 0.05	104.8 ± 70.9	5.3 ± 4.1	2.3 ± 1.9	8.8 ± 6.7

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. ** Skin tone, beard, glasses, and face cream information do not exist for all subjects.

Accuracy Data:

Operation System	Unique Subjects	Measurements	MAE ± STD	Ref Range
iOS	1162	1228	2.7 ± 2.9	516.0 - 1299.0
Android	902	980	3.0 ± 3.2	577.0 - 1344.0

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

Correlation and Bland-Altman plot by operating system

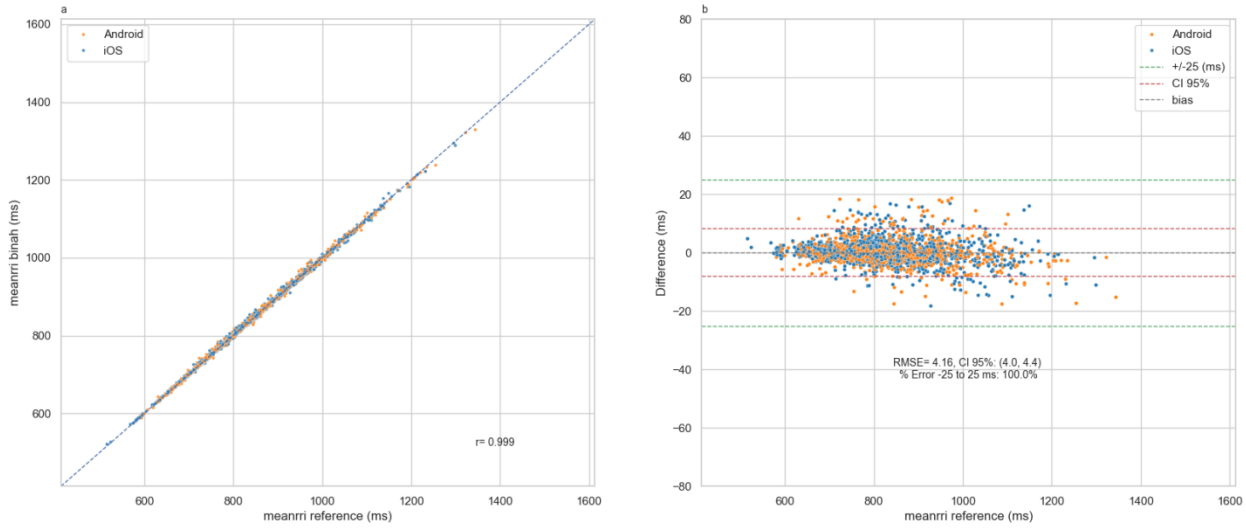

Figure 2:
a. Correlation plot by operating system - Binah.ai's mean RRi estimations versus reference device mean RRi measurements were found to be very high (r=0.999) for both operating systems (Android and iOS).

b. Bland-Altman plot by operating system - Bland-Altman plots for comparison between mean RRi measurements of the two methods (Binah's and the reference device) demonstrated high agreement between the two devices (100% of the measurements are within target error) for both operating systems (Android and iOS) in the presented mean RRi 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 ± 25 ms represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

SDNN

Accuracy Data:

Operation System	Unique Subjects	Measurements	MAE ± STD	Ref Range
iOS	1165	1231	6.5 ± 6.4	2.0 - 123.0
Android	916	994	7.2 ± 6.7	2.0 - 98.0

Table 3: Accuracy data for iOS and Android (MAE±STD) when compared to the reference device in the presented SDNN range.

RMSSD

Accuracy Data:

Operation System	Unique Subjects	Measurements	MAE ± STD	Ref Range
iOS	1162	1228	11.7 ± 13.3	3.0 - 160.0
Android	920	998	11.5 ± 12.8	3.0 - 138.0

Table 4: Accuracy data for iOS and Android (MAE±STD) when compared to the reference device in the presented RMSSD range.
MAE -Mean Absolute Error, STD - Standard Deviation

Conclusions

This report summarizes the results of an accuracy analysis demonstrating that the HRV measured by Binah.ai's SDK was **robust, highly accurate, and consistently met the mean RRi accuracy target (AE ≤ 25 ms) in 100%** of the measurements across both iOS and Android operating systems.

References

1. Selye, H. The stress of life. (1956).
2. Campkin, M. Stress management in primary care. Fam. Pract. 17, 98–99 (2000).
3. Thayer, J. F., Åhs, F., Fredrikson, M., Sollers III, J. J. & Wager, T. D. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci. Biobehav. Rev. 36, 747–756 (2012).
4. Malik, M. & Camm, A. J. Heart rate variability. Clin. Cardiol. 13, 570–576 (1990).
5. Salahuddin, L., Cho, J., Jeong, M. G. & Kim, D. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. in 2007 29th annual international conference of the ieee engineering in medicine and biology society 4656–4659 (IEEE, 2007).
6. Huang, M. et al. Association of depressive symptoms and heart rate variability in Vietnam war–era twins: a longitudinal twin difference study. Jama Psychiatry 75, 705–712 (2018).
7. Minassian, A. et al. Association of predeployment heart rate variability with risk of postdeployment posttraumatic stress disorder in active-duty marines. JAMA psychiatry 72, 979–986 (2015).
8. Carney, R. M. et al. Heart rate turbulence, depression, and survival after acute myocardial infarction. Psychosom. Med. 69, 4–9 (2007).
9. Stein, P. K., Domitrovich, P. P., Huikuri, H. V, Kleiger, R. E. & Investigators, C. Traditional and nonlinear heart rate variability are each independently associated with mortality after myocardial infarction. J. Cardiovasc. Electrophysiol. 16, 13–20 (2005).
10. Shah, A. S. et al. Association of Psychosocial Factors With Short-Term Resting Heart Rate Variability: The Atherosclerosis Risk in Communities Study. J. Am. Heart Assoc. 10, e017172 (2021).
11. Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C. M. & Suri, J. S. Heart rate variability: a review. Med. Biol. Eng. Comput. 44, 1031–1051 (2006).
12. Sztajzel, J. Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. Swiss Med. Wkly. 134, 514–522 (2004).

Appendix

Mean RRi error by Age and BMI

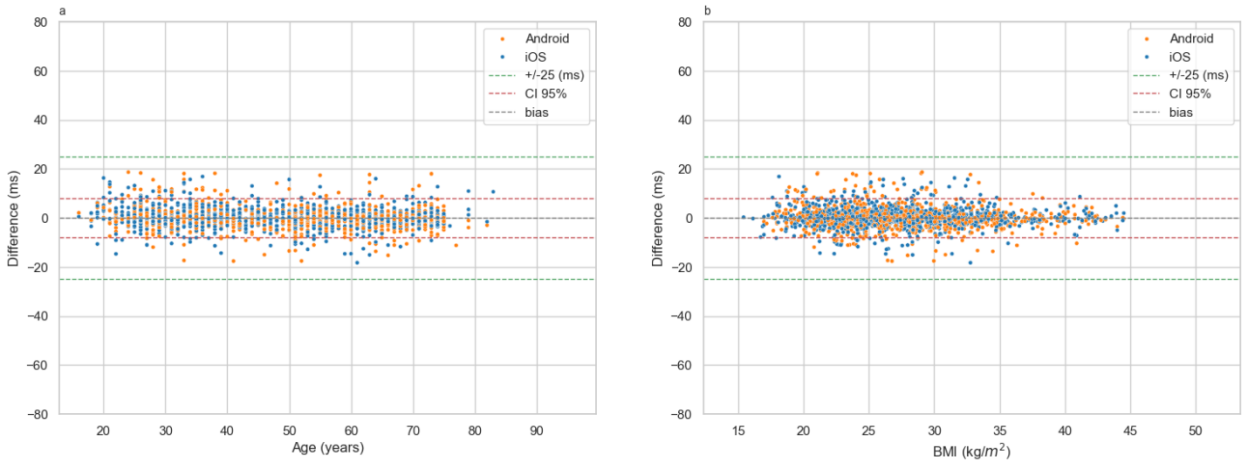

Figure 5:

a. Bland-Altman plot by age - demonstrated high agreement between mean RRi 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 high agreement between mean RRi 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 ±25 ms represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Mean RRi error by skin tone with Gender and Operating system

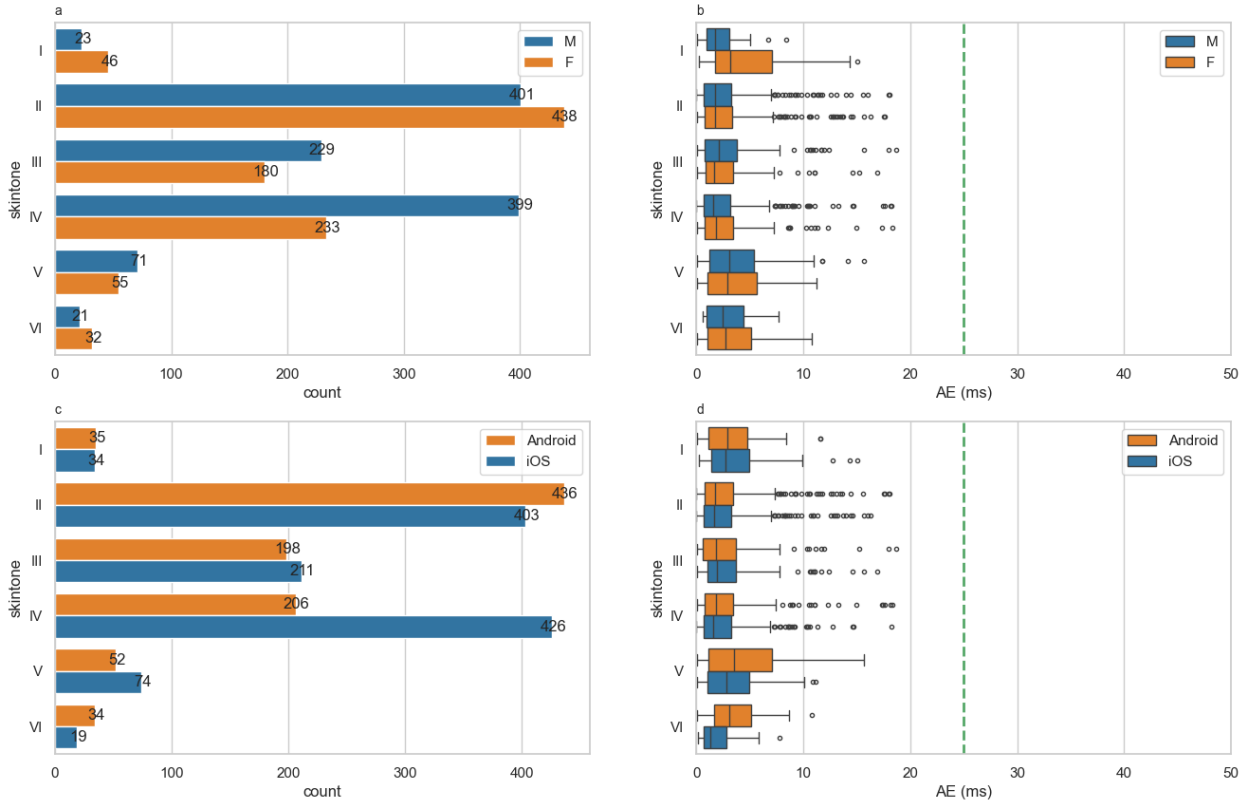

Figure 6:

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

b. **Box plot by Fitzpatrick skin tone and Sex** – mean RRi measurements obtained by Binah.ai's in comparison to the reference device are highly accurate for both sexes (female and male) across all presented skin tones.
The green dashed "Error" lines set at ≤25 ms represent 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** - mean RRi 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 ≤25 ms represent the value of the accuracy criterion.

Mean RRi error by distance and luminance

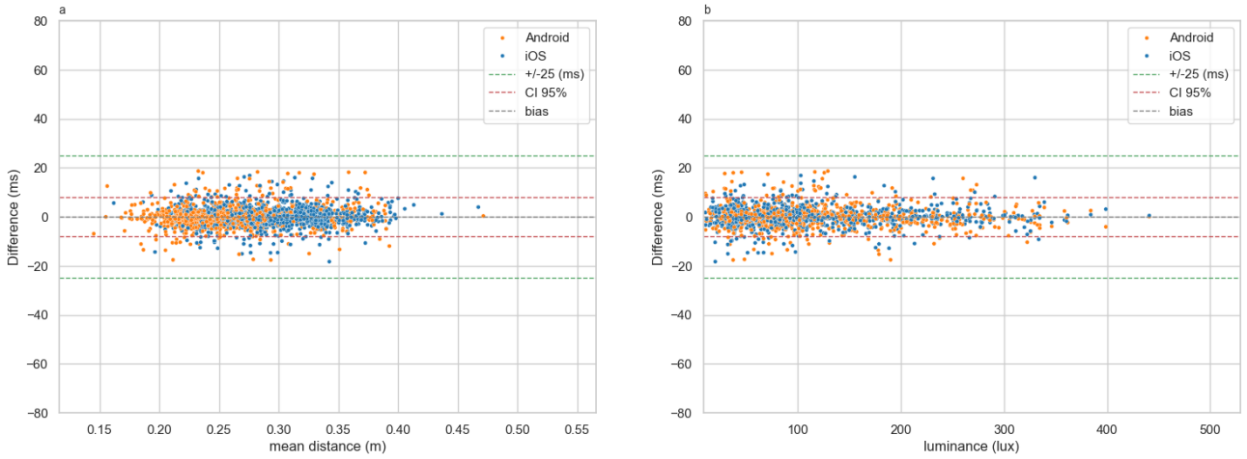

Figure 7:

a. **Bland-Altman plot by Distance (m)** - demonstrated high agreement between mean RRi measurements obtained by Binah.ai's 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 high agreement between mean RRi 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 gray dashed "Bias" line stands for the mean difference between measurements of Binah.ai and the reference device, the Th
The green dashed "Error" lines set at ±25 ms represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Mean RRi error by face Angles

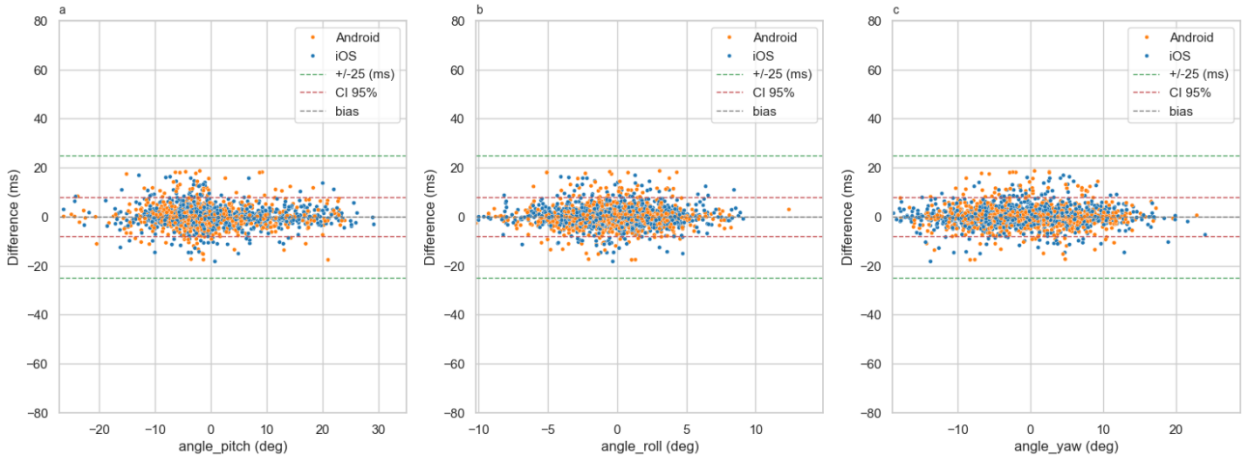

Figure 8:

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

b. **Bland-Altman plot by roll angle** - Comparison between mean RRi measurements obtained by Binah.ai and the

reference device for both operating systems (Android and iOS) within the presented roll range.
c. **Bland-Altman plot by yaw angle** - Comparison between mean RRi measurements obtained by Binah.ai and the reference device for both operating systems (Android and iOS) within the presented yaw range.

The gray dashed "Bias" line stands for the mean difference between measurements of Binah.ai and the reference device, the Th
The green dashed "Error" lines set at ±25 ms represent the value of the accuracy criterion, the "Limits of agreement" lines mark the limit of 95% of the samples.

Mean RRi error by Mobile Device Models

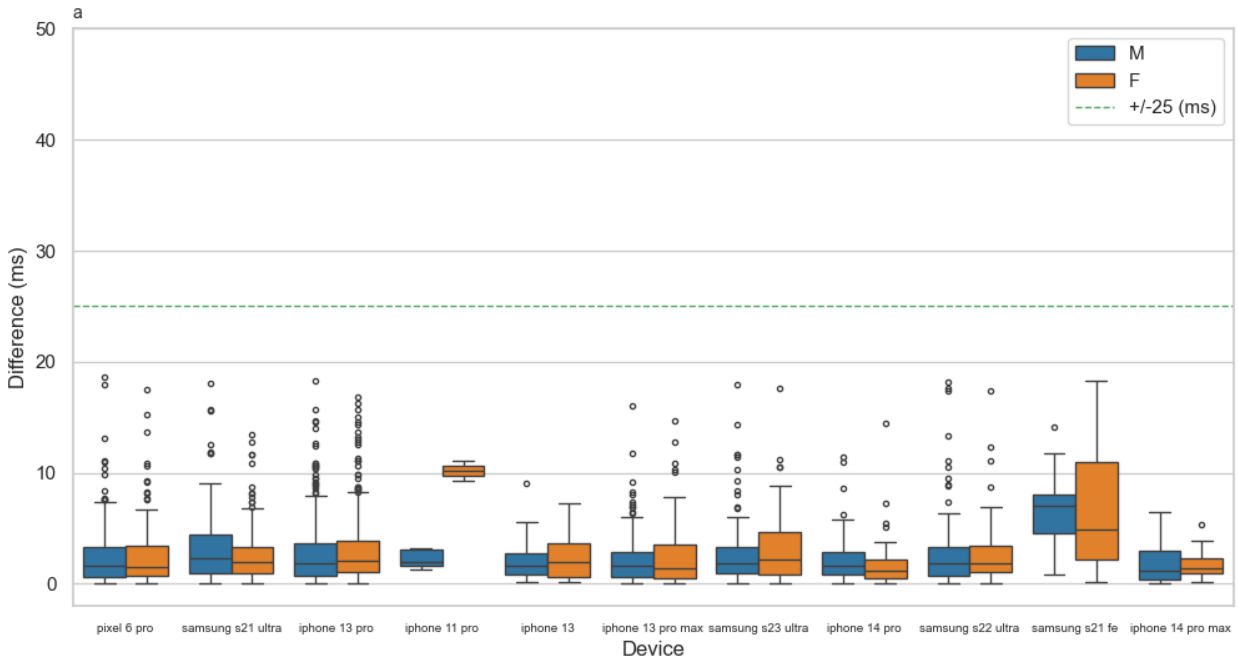

Figure 9:
a. **Box plot by device** – mean RRi measurements obtained by Binah.ai's versus the reference device are highly accurate for both sexes (female and male) on all devices.
The green dashed "Error" lines set at ≤ 25 ms represent the value of the accuracy criterion

Mean RRi error by Country

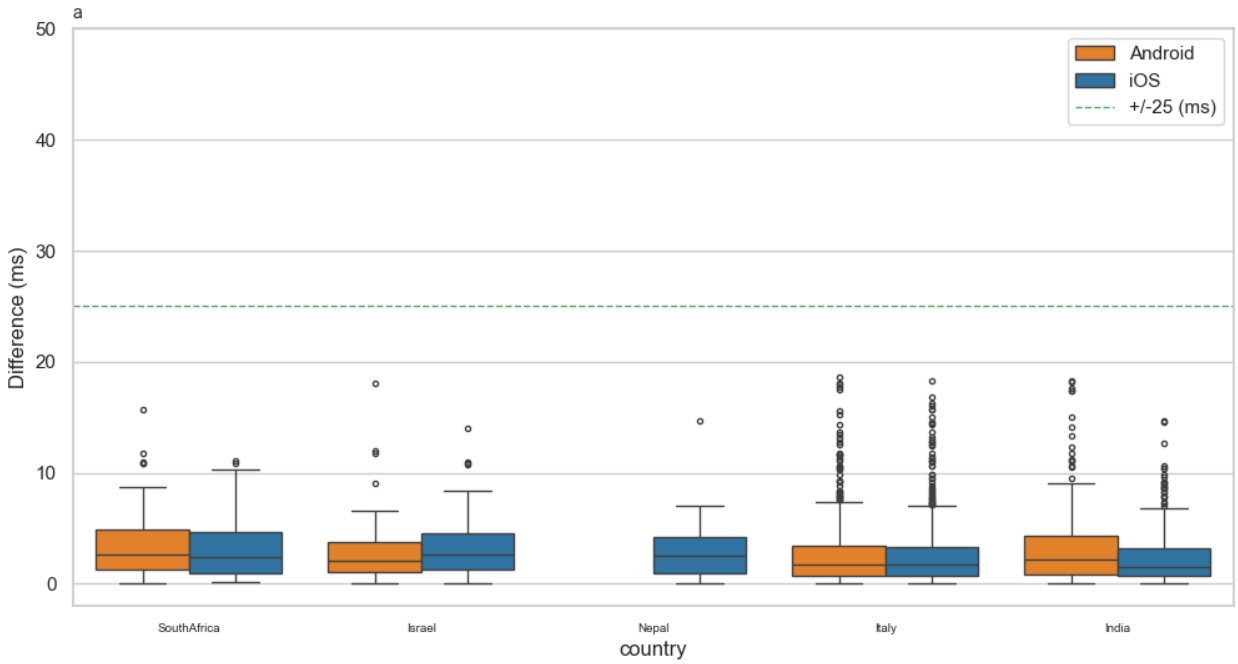

Figure 10:

a. **Box plot by country** - mean RRi measurements obtained by Binah.ai's versus the reference device is highly accurate for both operation systems (iOS and android) on all countries.

The green dashed "Error" lines set at ≤ 25 ms represent the value of the accuracy criterion.