ACCURACY ANALYSIS OF MPX5050DP, MLX90614, AND GY-MAX30102 SENSORS COMPARED TO STANDARD MEDICAL DEVICES FOR RECORDING FOUR VITAL SIGNS OF PATIENTS

Muhammad Adityo Rivalta: Ir. Wijianto, S.T., M.Eng.Sc.

Department of Mechanical Engineering, Faculty of Engineering, Universitas Muhammadiyah Surakarta

Abstrak

Akses terhadap perangkat pemantauan medis yang terjangkau masih menjadi tantangan signifikan di negara-negara berkembang, di mana tingginya biaya peralatan klinis standar menghambat deteksi dini dan perawatan pasien secara berkelanjutan. Kemajuan dalam teknologi sensor dan *Internet of Things* (IoT) telah membuka peluang untuk mengembangkan alternatif berbiaya rendah yang mampu mengukur tanda-tanda vital dengan tingkat akurasi yang dapat diterima secara klinis. Penelitian ini mengevaluasi tiga sensor terjangkau, yaitu MPX5050DP untuk tekanan darah, MLX90614 untuk suhu tubuh, dan GY-MAX30102 untuk detak jantung serta saturasi oksigen, yang diintegrasikan ke dalam satu sistem pemantauan berbasis ESP32. Setiap sensor dikalibrasi menggunakan perangkat medis bersertifikat sebelum pengujian dilakukan dalam kondisi laboratorium terkendali yang melibatkan lima responden sehat, dengan tiga kali pengulangan pengukuran untuk setiap parameter. Hasil penelitian menunjukkan bahwa seluruh sensor menghasilkan pembacaan yang sangat mendekati hasil pengukuran instrumen medis standar, tanpa ditemukan perbedaan yang signifikan secara statistik antara pengukuran sensor dan perangkat referensi. Semua parameter memenuhi ambang batas akurasi yang ditetapkan oleh standar perangkat medis internasional yang berlaku, dan seluruh sensor menunjukkan konsistensi internal yang tinggi pada pengukuran berulang. Temuan ini mengonfirmasi bahwa sistem sensor terintegrasi tersebut layak digunakan sebagai perangkat pemantauan tanda-tanda vital yang andal dan berbiaya rendah, serta memiliki potensi besar untuk diterapkan di fasilitas layanan kesehatan primer dan aplikasi telemedis berbasis IoT, khususnya di wilayah dengan keterbatasan sumber daya.

Kata kunci: MPX5050DP, MLX90614, GY-MAX30102, Tanda-tanda Vital, Akurasi Sensor, ESP32, Kalibrasi

Abstract

Access to affordable medical monitoring devices remains a significant challenge in developing countries, where high costs of standard clinical equipment hinder early detection and continuous patient care. Advances in sensor technology and the Internet of Things have created opportunities to develop low-cost alternatives capable of measuring vital signs with clinically acceptable accuracy. This study evaluates three affordable sensors, namely MPX5050DP for blood pressure, MLX90614 for body temperature, and GY-MAX30102 for heart rate and oxygen saturation, integrated into a single ESP32-based monitoring system. Each sensor was calibrated against certified medical devices before testing was conducted under controlled laboratory conditions involving five healthy respondents with three measurement repetitions per parameter. Results showed that all sensors produced readings closely aligned with standard medical instruments, with no statistically significant difference found between sensor and reference device

measurements. All parameters met the accuracy thresholds defined by applicable international medical device standards, and all sensors demonstrated high internal consistency across repeated measurements. These findings confirm that the integrated sensor system is feasible as a reliable, low-cost vital sign monitoring device with strong potential for deployment in primary healthcare facilities and IoT-based telemedicine applications, particularly in resource-limited settings.

Keywords: MPX5050DP, MLX90614, GY-MAX30102, Vital Signs, Sensor Accuracy, ESP32, Calibration

1. INTRODUCTION

The development of technology in the health sector has progressed rapidly, especially in terms of portable and integrated patient monitoring devices. One important aspect of modern healthcare services is the ability to measure vital signs including blood pressure, body temperature, heart rate, and oxygen saturation (SpO_2) quickly, accurately, and continuously. These vital signs are the main indicators in assessing a person's physiological condition and play a crucial role in early disease detection as well as monitoring patients with chronic conditions (World Health Organization, 2024).

However, conventional medical devices such as digital blood pressure monitors, medical infrared thermometers, and clinical oximeters still have limitations in terms of price, availability, and portability, especially in resource-limited areas. This situation drives the need for a low-cost sensor-based vital sign monitoring system that maintains high accuracy and can be integrated with Internet of Things (IoT) technology or telemedicine systems (Sari et al., 2024).

In this context, a number of low-cost sensors such as the MPX5050DP (a differential pressure sensor for blood pressure), MLX90614 (an infrared sensor for body temperature), and GY-MAX30102 (an optical sensor for measuring heart rate and SpO_2) are widely used in research and development of portable medical device prototypes. Nevertheless, the use of these sensors raises important questions regarding the accuracy and reliability of measurement results compared to certified standard medical instruments (Aulia et al., 2024).

Data from the World Health Organization indicates that around 40% of healthcare facilities in developing countries still lack standard vital sign measuring devices that meet clinical accuracy. This has a direct implication on delayed diagnoses and a decline in the quality of medical services, particularly in small clinics and remote areas. Therefore, the development of monitoring devices based on low-cost sensors becomes essential to address this need, provided that such devices are able to demonstrate medically acceptable accuracy and precision (World Health Organization, 2024).

Several previous studies have examined each sensor separately. For instance, a study published in the Indonesian Journal of Electrical and Electronics Engineering found that the MLX90614 sensor had an average error of about 1.71% or an accuracy of 98.29% compared to a standard non-contact thermometer. (Aulia et al., 2024). Meanwhile, another study that used the MPX5050DP sensor for blood pressure reported a systolic error of about 3.45% and a diastolic error of 7.17% compared to the Omron HEM-7120 digital sphygmomanometer (Sari et al., 2024). On the other hand, the GY-MAX30102 sensor showed a maximum error of 0.43% for SpO_2 and 2.02% for heart rate (Triwiyanto & others, 2022).

Although these results are quite promising, most studies are still limited to testing only one or two sensors and have not yet integrated all vital signs into a single measurable system. In addition, many studies were conducted under ideal conditions (without motion artifacts, temperature variations, or differences in skin tone), thus not fully reflecting real-world conditions in the field (Nemomssa & Raj, 2022).

Based on these conditions, this study is important to empirically analyze the accuracy, precision, and reliability of the MPX5050DP, MLX90614, and GY-MAX30102 sensors compared to standard medical devices in measuring four main vital signs of patients. This research is expected to provide a tangible contribution to the development of low-cost portable medical devices that are feasible for use in primary healthcare services as well as patient telemonitoring systems. (Sari et al., 2024).

2. METHOD

2.1 Research Flow Chart

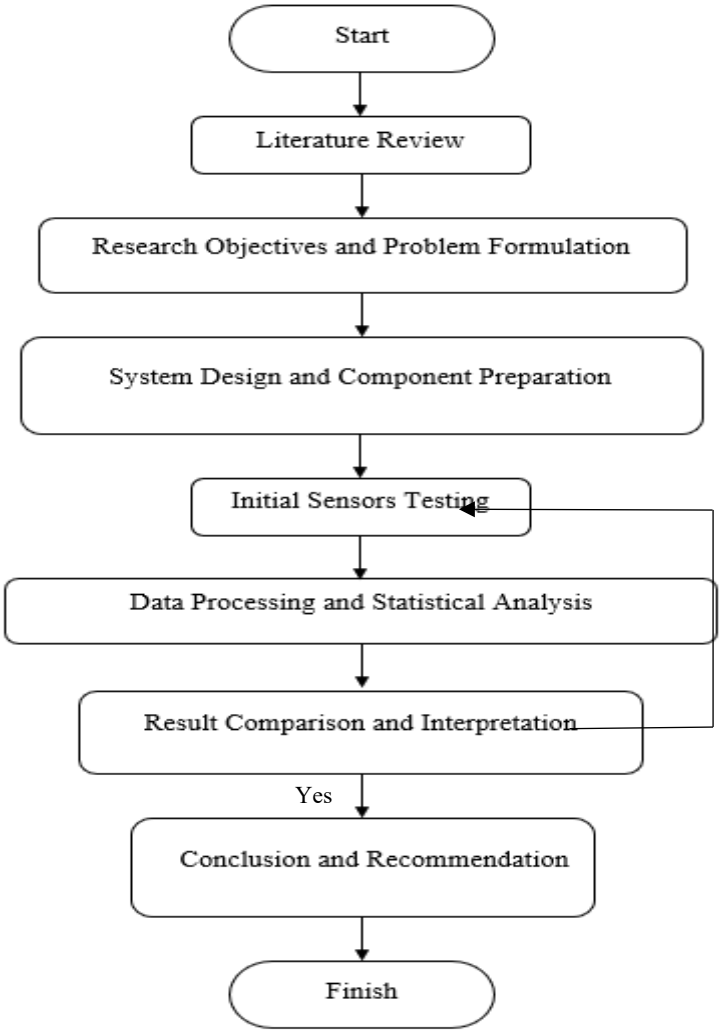

Figure 1 FlowChart Diagram

2.2 Types and Research Approach

The experimental research method was chosen because the researcher aims to determine the degree of conformity between the results of three sensors (MPX5050DP, MLX90614, and GY-MAX30102) and medical standards in measuring four main vital signs (pressure, temperature, heart rate, and blood oxygen level). This approach allows the researcher to control external variables and objectively assess the reliability of the device. In line with the study by Brahimaj (2020), the experimental method has proven effective in testing the performance of wearable health devices to ensure the validity and stability of the generated data.

In addition, the comparative-quantitative method provides a strong foundation for conducting evaluations based on numerical data. In this context, the study by Ayieko (2024)also applied a similar principle when comparing the MLX90614 temperature sensor and

the MAX30102 heart rate sensor with standard medical instruments to determine the accuracy and stability of the measurement results. Such research not only provides academic value but also supports innovation in microcontroller-based medical technology.

This research approach is highly relevant to the advancement of knowledge in biomedical instrumentation and digital health technology (e-health). Through controlled experiments and comparative analysis, studies of this kind can help improve the accuracy and efficiency of low-cost medical measuring devices. As explained by Widadi et al., (2020), scientific validation of microcontroller-based sensors is an essential step toward the broader and more reliable implementation of telemedicine technology. Thus, this study has the potential to make a tangible contribution to the development of intelligent health devices that support continuous patient monitoring.

2.3 Research Location and Time

2.3.1 Rationale for Choosing the Research Location

This research will be conducted in the Computer Laboratory of the Mechanical Engineering Study Program, Faculty of Engineering. This location was selected because it provides adequate facilities for testing and analyzing electronic sensor systems. The laboratory is equipped with high-performance computers, precision measuring instruments, and supporting devices such as microcontroller kits, breadboards, power supplies, and data loggers, all of which are essential for testing the MPX5050DP, MLX90614, and GY-MAX30102 sensors.

In addition to the availability of infrastructure, the laboratory offers a controlled and standardized environment, allowing experiments to be carried out systematically and free from external interference that could affect measurement results. Support from experts and academic supervisors in the fields of instrumentation and electronics also makes this location highly relevant to the needs of the research. In line with the findings of Brahimaj (2020) and Ayieko (2024), conducting biomedical sensor testing in a stable laboratory environment is a crucial step in ensuring the accuracy, precision, and reliability of measurement results.

2.3.2 Research Period

The research is scheduled to take place from December 2025 to January 2026. This two-month period was chosen to allow comprehensive research activities from preparation and data collection to result analysis and report writing. The chosen timeframe also considers the laboratory's more flexible availability at the end of the academic semester, ensuring that the experimental process can proceed without disruptions from regular academic activities.

Table 1 Research Activity Schedule

No	Research Stage	Activities Conducted	Time Period	Expected Output
1	Preparation Stage	Literature review, data collection, and research instrument preparation	December 1–7, 2025	Well-defined theoretical and methodological framework
2	System and Device Design	Assembly and programming of sensors (MPX5050DP, MLX90614, GY-MAX30102) with microcontroller integration	December 8–14, 2025	Prototype device ready for testing
3	Initial Sensor Calibration	Sensor calibration against medical-grade instruments for validation	December 15–20, 2025	Calibration data and initial comparison results
4	Experimental Data Collection	Repeated measurements of four vital signs (pressure, temperature, SpO_2, and heart rate)	December 21–31, 2025	Complete measurement dataset
5	Data Analysis and Accuracy Evaluation	Data processing, error calculation, correlation testing, and statistical validation	January 1–10, 2026	Statistical analysis and sensor performance evaluation
6	Report Drafting and Consultation	Writing research results, discussion, and supervisor consultation	January 11–20, 2026	Draft of the final report

7	Finalization and Presentation	Report revision, conclusion formulation, and result presentation	January 21–31, 2026	Final report and research presentation

With well-equipped laboratory facilities, a controlled testing environment, and a structured time plan, this research is expected to produce valid and comprehensive accuracy analysis results for the selected sensors. The outcomes of this study will also contribute significantly to the development of low-cost, sensor-based vital signs monitoring systems, supporting innovation in biomedical instrumentation and digital health technology.

2.4 Population and Sample

The population in this research consists of healthy individuals aged 19–25 years. This group is chosen because they generally have stable physiological characteristics, minimal chronic health risks, and consistent vital sign ranges, which help ensure accurate sensor performance evaluation. Young adults are also easier to recruit within an academic setting, making them ideal participants for a controlled biomedical experiment. As stated by Palavesam et al, (2023) testing biomedical sensors on healthy individuals within this age range minimizes confounding variables and produces more reliable calibration outcomes.

The sampling technique applied in this study is purposive sampling, a non-probability method where subjects are intentionally selected based on specific inclusion criteria. Participants must be between 19 and 25 years old, in good physical health, and free from cardiovascular or respiratory disorders. This method allows the researcher to focus on individuals who can provide data relevant to the study's objectives, rather than selecting respondents randomly. Etikan, (2016) emphasizes that purposive sampling is effective when the researcher needs participants with particular characteristics crucial for validating specific instruments or technologies.

To estimate the appropriate sample size, the Slovin formula is used to determine the minimum number of respondents that can represent the population while maintaining statistical reliability at a 5% margin of error (e = 0.05)**:**

$$\boldsymbol{n = \frac{N}{1+N(e)^2}} \quad (1)$$

where :

n = required sample size

N = total population

e = margin of error (5%)

If the target population consists of **20 students** from the same department, the calculation becomes:

$$n = \frac{20}{1 + 20(0.05)^2} = 19.05$$

Thus, at least 19 respondents are ideal for statistical representation. However, considering experimental feasibility, this research will use five respondents with repeated measurements for each vital sign (blood pressure, heart rate, SpO_2, and body temperature). According to Palavesam et al. (2023) repeated-measure designs using smaller, well-defined samples can effectively reduce inter-individual variability and strengthen sensor accuracy validation.

This approach ensures that the resulting data are both valid and reliable, as each individual acts as their own control across multiple trials. The repeated-measure design is advantageous for technological validation studies because it isolates the influence of the device from natural physiological differences, ensuring that observed deviations originate from sensor performance rather than participant variability.

2.5 Tools and Materials

2.5.1 Tools

1. Laptop for data recording and analysis
2. ESP32 microcontroller
3. Breadboard and jumper wires
4. DC power supply
5. Digital multimeter
6. USB data cable
7. Standard medical measuring instruments:
   a. Digital medical thermometer
   b. Standard pulse oximeter
   c. Digital sphygmomanometer
8. Respiratory rate monitor
9. Stopwatch
10. Precision screwdriver set
11. Arduino IDE software

2.5.2 Materials

1. Electronic Components
   a. MPX5050DP pressure sensor
   b. MLX90614 infrared temperature sensor

c. GY-MAX30102 pulse oximeter and heart rate sensor

d. Supporting resistors and capacitors

e. Jumper wires and connectors

f. Breadboard and pin headers

g. 5V power bank or 12V adapter

2. Supporting Components

a. LCD or OLED display module (for displaying sensor readings)

b. Protective sensor housing or enclosure box

c. Acrylic or PCB mounting board for sensor assembly

d. Male–female pin connectors

2.6 Research Procedure

The research procedure outlines the systematic steps undertaken to ensure accurate calibration, data acquisition, and analysis of sensor performance compared to standard medical devices. The procedure consists of five main stages**:** sensor calibration**,** environmental preparation**,** simultaneous data collection**,** repeated measurement, and statistical data analysis. Each stage is designed to guarantee data reliability, minimize measurement errors, and validate the accuracy of the sensors under controlled conditions.

Before data collection, all sensors (MPX5050DP, MLX90614, and GY-MAX30102) are calibrated against certified medical devices. Calibration is conducted under controlled laboratory conditions using standard references: a digital sphygmomanometer for pressure, a clinical thermometer for temperature, and a medical-grade pulse oximeter for heart rate and oxygen saturation. Each sensor's reading is compared to the medical device output, and a calibration curve or correction factor is derived to minimize deviation. This procedure ensures sensor linearity and accuracy within the operational range (Idiong et al., 2023).

Data collection is conducted in the Computer Laboratory of the Mechanical Engineering Department, maintaining ambient conditions between 25–27°C with stable humidity. Environmental control prevents measurement drift due to temperature fluctuations. The setup includes connecting all sensors to the ESP32 microcontroller, configured via Arduino IDE and monitored using a laptop. Proper grounding and isolation techniques are used to prevent electrical noise interference (Saldana-Perez & Contreras-Jiménez, 2025).

Measurements from the sensors and the standard medical devices are recorded simultaneously to ensure time-synchronized data comparison. Each participant's vital signs (blood pressure, body temperature, heart rate, and oxygen saturation) are recorded in real-time using both systems. Data are stored digitally with timestamps to facilitate later statistical

correlation and accuracy analysis. The simultaneous recording minimizes human error and enhances data comparability (Zhang et al., 2020).

To ensure repeatability and reliability, each respondent's vital signs are measured at least three times under identical conditions. The mean of the three measurements is considered the final value for each variable. This repetition reduces random errors and helps identify potential sensor drift or instability. The collected data are validated for outliers and consistency before statistical processing (Barzan & Arbilei, 2024).

All recorded data are analyzed using statistical methods, including mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient (R^2) to evaluate each sensor's accuracy relative to the standard medical device. Graphical comparisons and regression analysis are performed using Excel to visualize performance trends and calibration results. These analyses provide a quantitative assessment of sensor performance and support conclusions regarding their suitability for medical monitoring applications (Kowshik et al., 2025).

2.7 Data Collection Techniques

This research employs a quantitative data collection approach using direct observation and electronic data logging to obtain accurate, real-time readings from the sensors (MPX5050DP, MLX90614, and GY-MAX30102). The primary purpose of this method is to ensure that every recorded data point reflects actual physiological parameters measured simultaneously with the standard medical devices. The collected data include quantitative measurements of blood pressure, body temperature, heart rate, and oxygen saturation (SpO_2), expressed in appropriate metric units.

The main technique used is direct observation, which allows the researcher to monitor sensor readings as they occur in real-time through a serial monitor or data logger connected to the ESP32 microcontroller. This process ensures precision and minimizes the potential for data entry errors. Each reading from the sensors is automatically timestamped and recorded using Arduino IDE's serial monitor and/or Python-based data logging software. The data are then exported into CSV format for further statistical analysis. This method aligns with established data acquisition standards in biomedical engineering for sensor validation (Beri et al., 2022).

After data collection, all measurement results are documented digitally and cross-verified with readings from standard medical devices. This verification ensures the validity and reliability of the obtained data. Data validation includes cross-checking for outliers and inconsistencies, while reliability testing is carried out through repeated measurements under the same conditions. Data integrity is maintained by using consistent calibration parameters

and identical sampling intervals throughout the measurement sessions (Saldana-Perez & Contreras-Jiménez, 2025).

2.8 Data Analysis Techniques

This study employs quantitative statistical analysis to evaluate the accuracy, precision, and reliability of sensor readings specifically the MPX5050DP (pressure sensor), MLX90614 (temperature sensor), and GY-MAX30102 (heart rate and SpO_2 sensor) by comparing them against standard medical-grade devices. The analysis aims to ensure that the developed system provides measurement results that are statistically comparable to conventional instruments used in clinical practice.

The analytical process encompasses several interrelated stages, including data normality testing, correlation analysis, mean difference testing, error computation, and data visualization. Each stage plays a crucial role in validating the quantitative performance of the sensors. The use of Microsoft Excel as the primary analytical tool allows for efficient data processing, accurate statistical computation, and clear graphical representation. Excel's built-in functions and statistical tools are particularly effective for handling small to medium datasets typical of laboratory-based biomedical research.

The choice of statistical methods in this study is grounded in their ability to examine measurement accuracy, consistency, and significance between two paired datasets (sensor readings vs. reference device readings). The quantitative approach ensures objectivity and reproducibility in analyzing the precision and reliability of sensor measurements.

The Shapiro–Wilk test is used to determine whether the data follow a normal distribution, which is a fundamental assumption for conducting parametric analyses. According to Teh et al, (2020), the Shapiro–Wilk test provides a reliable measure of normality, especially for small to medium sample sizes. This test evaluates how closely the data distribution matches a perfect normal curve.

In Microsoft Excel, normality can be assessed by calculating the Shapiro–Wilk statistic manually or through the use of add-ins and custom formulas, followed by examining skewness, kurtosis, and histogram patterns. When the significance value (p) is greater than 0.05, the data are considered normally distributed, indicating that parametric statistical analyses can be appropriately applied. Establishing normality ensures that subsequent tests, such as correlation and t-tests, yield valid and reliable results.

Following the normality test, Pearson's correlation coefficient (r) is calculated to assess the strength and direction of the linear relationship between the sensor data and the corresponding standard medical device measurements. This method quantifies how well one

variable (sensor reading) predicts another (reference reading). The correlation coefficient ranges from –1 to +1, with values closer to +1 indicating a strong positive correlation (Das et al., 2024).

In Excel, the correlation is computed using the =CORREL(array1, array2) function, which provides an efficient and transparent means to calculate the Pearson coefficient. A high positive correlation signifies that the sensor readings closely mirror those of the medical-grade instruments, validating the accuracy of the sensors under investigation. Moreover, the statistical significance of the correlation can be evaluated by analyzing the p-value and confidence intervals, ensuring that observed associations are not due to random variation.

To assess whether there is a significant difference between the sensor and reference device readings, the Paired Sample t-Test is performed. This test examines the mean of the differences between paired data points (sensor vs. standard device) and determines whether these differences are statistically significant (Hasija, 2023).

In Microsoft Excel, this test is conducted using the Data Analysis Toolpak, specifically through the "t-Test: Paired Two Sample for Means" feature. This function compares the mean values and computes the t-statistic and p-value automatically. When the p-value > 0.05, it indicates that no significant difference exists between the two measurement sources, suggesting that the sensor provides results comparable to medical-grade devices. Conversely, a p-value < 0.05 implies a statistically significant discrepancy, which could highlight calibration or environmental factors affecting sensor accuracy.

To quantify measurement accuracy, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used. These metrics evaluate the magnitude of deviation between sensor readings and reference device measurements. As Ayieko, (2024) noted, both MAE and RMSE provide complementary insights into accuracy MAE reflects the average deviation magnitude, while RMSE emphasizes larger errors by squaring the residuals.

The formulas are defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(x_i - y_i)^2}$$

where x_i represents the sensor reading and y_i the reference device reading. In Excel, these values can be calculated using functions such as =ABS(), =AVERAGE(), and =SQRT(). Lower MAE and RMSE values indicate that the sensor data are closely aligned with the reference measurements, demonstrating high precision and reliability.

Data visualization plays a key role in understanding and interpreting analytical results. Microsoft Excel is utilized to generate a variety of graphs and charts, including scatter plots, line graphs, and error bar charts, which display comparisons between sensor outputs and reference measurements.

Scatter plots are used to illustrate the degree of linearity and distribution between the two datasets, helping to visually confirm correlation results. Line graphs, on the other hand, depict trends in sensor and device readings over time, highlighting consistency and synchronization. Additionally, error bar charts reveal variability and uncertainty in the measurements, providing a clear visual interpretation of data precision.

Through these visual tools, researchers can identify potential systematic deviations, outliers, and nonlinear response behaviors that may not be evident through numerical analysis alone. Visualization thus enhances the clarity and communicative value of the findings.

To uphold the validity and reliability of the analysis, several procedural controls are applied. All statistical tests and calculations are conducted using Microsoft Excel, which ensures data transparency and reproducibility. Prior to analysis, data are carefully inspected for outliers and missing values that could distort statistical interpretations. Any abnormal readings caused by motion artifacts, poor sensor contact, or electrical noise are identified and handled appropriately either through correction or exclusion.

Furthermore, each experimental condition is repeated multiple times for every respondent to confirm the stability and repeatability of sensor performance. Repeated measurements reduce the influence of random noise and allow for a more reliable average value to be compared with medical-grade instruments.

These quality assurance steps guarantee that the analytical results are empirically sound, statistically robust, and directly aligned with the research objective to validate the accuracy and consistency of low-cost biomedical sensors in measuring the four vital signs of patients compared to standard medical devices.

2.9 Data Validity and Reliability

In experimental quantitative research such as this, ensuring the validity and reliability of data is a fundamental requirement for maintaining scientific credibility. Both aspects serve as the foundation for determining whether the instruments used in this case, the MPX5050DP, MLX90614, and GY-MAX30102 sensors can accurately and consistently measure the patient's vital signs, including blood pressure, body temperature, heart rate, and oxygen saturation (SpO_2).

Validity ensures that an instrument truly measures the intended variable, while reliability guarantees that the results obtained are consistent and reproducible under similar conditions. Establishing both aspects is essential to ensure that the findings are not only accurate but also scientifically defensible.

Instrument validity refers to the degree to which a measurement tool measures what it is supposed to measure (Reddy et al., 2023). In this study, sensor readings were compared against standard medical instruments to ensure accurate measurement of physiological parameters. Three types of validity were assessed: content validity, construct validity, and criterion validity.

a. Content Validity

Content validity ensures that all aspects of the measured variables are adequately represented. To achieve this, an expert judgment approach was applied by involving professionals in biomedical instrumentation and sensor calibration. These experts reviewed whether each sensor system. The MPX5050DP for pressure, MLX90614 for temperature, and GY-MAX30102 for heart rate and SpO_2. appropriately corresponded to the physiological constructs being measured.

This validation confirmed that each component of the system had a clear theoretical and practical correspondence with medical measurement standards, thereby ensuring the scientific integrity of the data collection process.

b. Construct Validity

Construct validity examines whether the instrument's measurements align with the theoretical concept of the variable being measured. In this research, Pearson's Product-Moment correlation was used to assess the strength of the relationship between sensor data and standard medical device outputs. The formula is expressed as follows:

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \quad (3)$$

Where :

r = correlation coefficient

X = sensor reading values

Y = standard medical instrument readings

n = number of paired observations

A correlation value of r>0.7r > 0.7r>0.7 indicates a strong positive relationship, suggesting high construct validity (Blažauskas et al., 2023). This test demonstrates that

the sensors' measurements correspond closely to theoretical expectations and the real physiological constructs they represent.

c. Criterion Validity

Criterion validity assesses the extent to which a measurement instrument correlates with an established standard or benchmark. In this study, it was tested by directly comparing sensor readings with those obtained from certified medical devices. A high correlation coefficient close to 1 indicates that the sensors produce measurements that are consistent with the reference devices, thereby confirming their criterion-based validity.

Instrument reliability measures the consistency and stability of an instrument's readings under identical conditions (Chattopadhyay & Das, 2022). A reliable instrument will produce nearly identical results when measurements are repeated, provided that external factors remain constant. Two main approaches were implemented: replication and repetition testing, and Cronbach's Alpha reliability testing.

a. Replication and Repetition

Each measurement was repeated three times per respondent under controlled laboratory conditions to ensure data consistency. The room temperature was maintained between 25–27°C, and subjects' physical conditions remained stable. Repeated measurements were analyzed using Pearson's Product-Moment correlation to determine the degree of similarity among trials. High correlation across trials reflects that the sensors produce stable and consistent readings.

b. Cronbach's Alpha Reliability

To test internal consistency among multiple readings, the Cronbach's Alpha (α) coefficient was computed using the following formula:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum S_i^2}{S_t^2}\right) \quad (4)$$

Where :

$\alpha =$ reliability coefficient

$k =$ number of measurement items

$S_i^2 =$ variance of each item

$S_t^2 =$ total variance

According to Blažauskas et al, (2023), a Cronbach's Alpha value of $\alpha \geq 0.70$ indicates satisfactory reliability. In this study, α was calculated using MATLAB, ensuring that the collected sensor data were internally consistent. This step validates

that the sensors do not produce random or unstable readings, thereby ensuring measurement reproducibility over time.

Ensuring both validity and reliability is not only a methodological requirement but also a scientific necessity to maintain the credibility and integrity of research findings. Within the context of engineering and biomedical studies, validity guarantees the accuracy of sensor readings in representing real physiological parameters, while reliability ensures that the readings are consistent and reproducible across trials. The combination of these two elements enhances data accuracy and trustworthiness, strengthening the internal validity of the study and providing solid evidence for comparing biomedical sensors to standard instruments. Additionally, establishing strong validity and reliability supports future development of IoT-based health monitoring systems, where sensor precision and consistency are critical for real-time diagnostics.

3. RESULTS AND DISCUSSION

3.1 Sensor Calibration Results

Prior to conducting the primary data collection sessions, all three sensors integrated in this study underwent a structured calibration process in accordance with the procedure outlined in Chapter 3.6. Calibration was performed by directly comparing each sensor's output against certified standard medical instruments under controlled laboratory conditions, with ambient temperature maintained between 25°C and 27°C throughout the entire session (Saldana-Perez & Contreras-Jiménez, 2025). The reference instruments used were a certified digital sphygmomanometer for the MPX5050DP pressure sensor, a digital clinical thermometer for the MLX90614 infrared sensor, and a medical-grade pulse oximeter for the GY-MAX30102 optical sensor. This stage is essential in biomedical sensor validation, as pre-calibration systematic bias must be identified and corrected before experimental measurements are conducted with human subjects (Idiong et al., 2023).

A total of ten measurement points were recorded for each parameter during the pre-calibration session. The complete dataset is presented in Table 4.1.

Table 2 Pre-Calibration Sensor Measurement Data (Before Adjustment)

No.	BLOOD PRESSURE (MPX5050DP)				BODY TEMPERATURE (MLX90614)		HEART RATE (GY-MAX30102)		SpO_2 (GY-MAX30102)	
No.	Systolic Std (mmHg)	Systolic Sensor (mmHg)	Diastolic Std (mmHg)	Diastolic Sensor (mmHg)	Std (°C)	Sensor (°C)	Std (bpm)	Sensor (bpm)	Std (%)	Sensor (%)
1	123	126	77	81	37,10	35,69	87	91	98	98
2	99	117	60	79	37,00	36,33	87	89	97	98
3	111	124	69	80	37,20	35,66	86	86	97	97
4	110	129	68	82	36,90	36,47	86	87	97	96
5	110	125	68	80	37,00	36,47	88	89	98	99
6	109	123	67	76	37,00	34,93	87	91	98	96
7	110	125	67	75	37,10	34,91	88	91	99	96
8	112	127	68	77	37,00	34,89	88	87	97	97
9	112	125	69	75	37,00	35,47	87	88	97	97
10	113	125	68	77	37,00	35,69	87	87	98	98

The data in Table 4.1 reveal that pre-calibration conditions produced substantial systematic deviations in several parameters, most prominently in the MPX5050DP blood pressure sensor and the MLX90614 temperature sensor. The MPX5050DP consistently overestimated systolic pressure with a mean positive offset of 13.70 mmHg, and diastolic pressure with a mean offset of 10.10 mmHg relative to the standard device. These systematic errors are characteristic of piezoresistive pressure sensors operating without prior zero-point adjustment, a phenomenon documented in low-cost sensor validation studies by Mazoteras-Pardo et al. (2023), who emphasized that blood pressure sensors require careful calibration against certified sphygmomanometers before clinical use. The MLX90614 infrared thermometer demonstrated a consistent underestimation of body temperature with a mean negative offset of 1.38°C, a deviation that likely reflects the influence of ambient laboratory temperature on the sensor's internal compensation circuit as well as measurement distance variability, consistent with the sensor's thermal characteristics described in Chapter 2.1.2. The GY-MAX30102, by contrast, exhibited considerably smaller pre-calibration offsets of 1.50 bpm for heart rate and 0.40% for SpO_2, which aligns with findings by Triwiyanto et al. (2022) who reported that the MAX30102 sensor maintains relatively stable baseline readings even before formal calibration.

From the pre-calibration dataset, a correction factor (CF) was calculated for each parameter using the following formula:

$$CF = Mean\ (Standar\ Value\ Minus\ Sensor\ Value)$$

This approach yields a systematic offset value that, when applied to the sensor firmware, shifts all subsequent readings to align with the standard instrument. The correction factors derived and subsequently programmed into the ESP32 firmware are presented in Table 4.2.

Table 3 Correction Factors Derived from Pre-Calibration Data

Parameter	Sensor	Mean Offset (Pre-Calibration)	Correction Factor (CF)	Correction Direction
Systolic Blood Pressure	MPX5050DP	+13.70 mmHg	-13.70 mmHg	Subtract from sensor output
Diastolic Blood Pressure	MPX5050DP	+10.10 mmHg	-10.10 mmHg	Subtract from sensor output
Body Temperature	MLX90614	-1.38 °C	+1.38 °C	Add to sensor output
Heart Rate	GY-MAX30102	+1.50 bpm	-1.50 bpm	Subtract from sensor output
SpO_2	GY-MAX30102	-0.40 %	+0.40 %	Add to sensor output

Following the application of each correction factor through firmware offset adjustments in the Arduino IDE environment, post-calibration measurements were conducted to verify the effectiveness of the calibration procedure. The comparison between pre-calibration and post-calibration error metrics for all parameters is presented in Table 4.3 and visualized in Figure 4.1.

Table 4 Comparison of Measurement Error Before and After Calibration

Parameter	Sensor	MAE Before Calibration	% Error Before	MAE After Calibration	% Error After	Error Reduction (%)
Systolic Blood Pressure	MPX5050DP	13.70 mmHg	12.53%	1.20 mmHg	1.04%	**91,7%**
Diastolic Blood Pressure	MPX5050DP	10.10 mmHg	15.16%	1.27 mmHg	1.74%	**87,4%**
Body Temperature	MLX90614	1.38 °C	3.72%	0.10 °C	0.27%	**92,8%**
Heart Rate	GY-MAX30102	1.70 bpm	1.95%	1.00 bpm	1.32%	32,3%
SpO_2	GY-MAX30102	0.80 %	0.82%	0.67 %	0.69%	13,8%

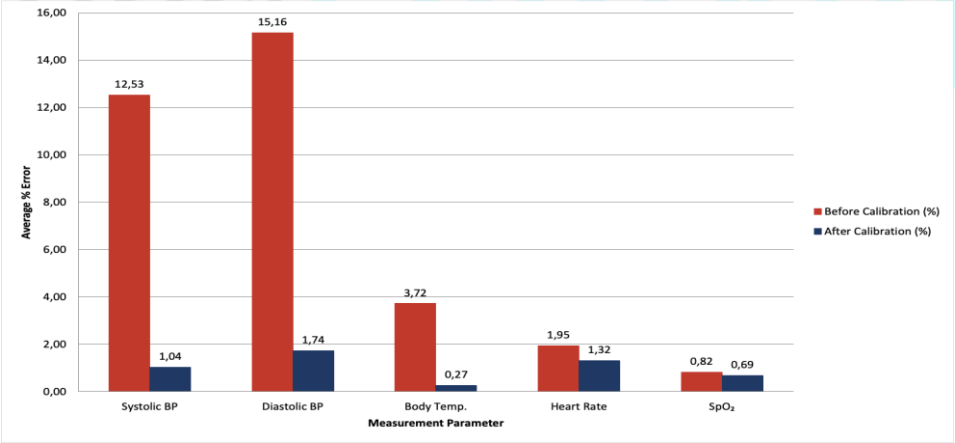

Figure 2 Averange Percentage Error Before and After Callibration

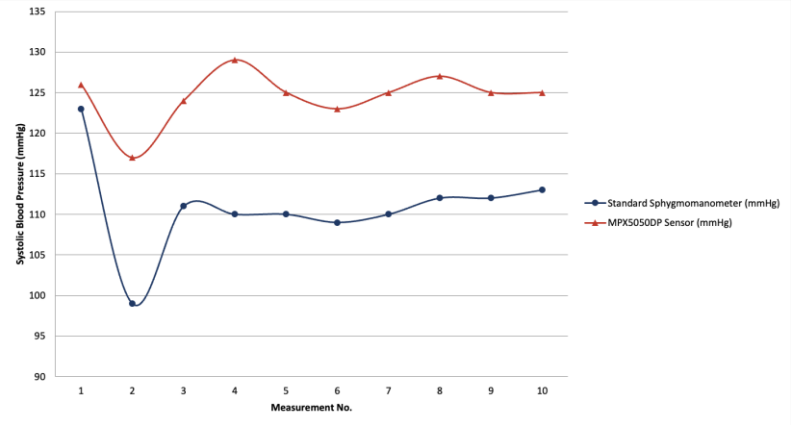

Figure 3 Systolic Blood Pressure: MPX5050DP vs Standard Sphygmomanometer (Pre-Calibration)

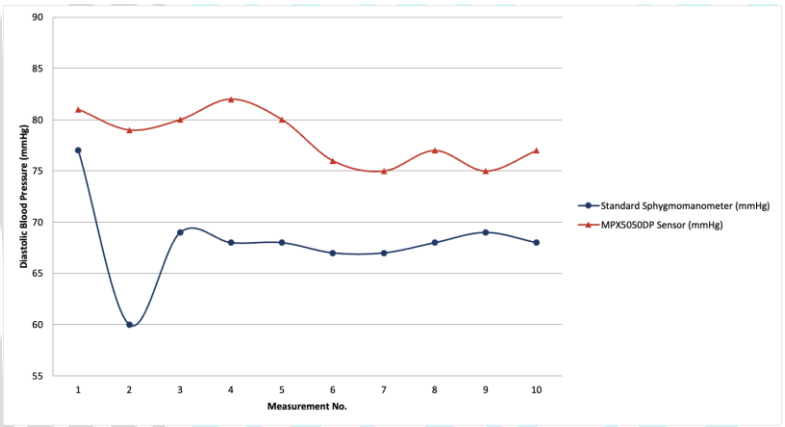

Figure 4 Diastolic Blood Pressure: MPX5050DP vs Standard Sphygmomanometer (Pre-Calibration)

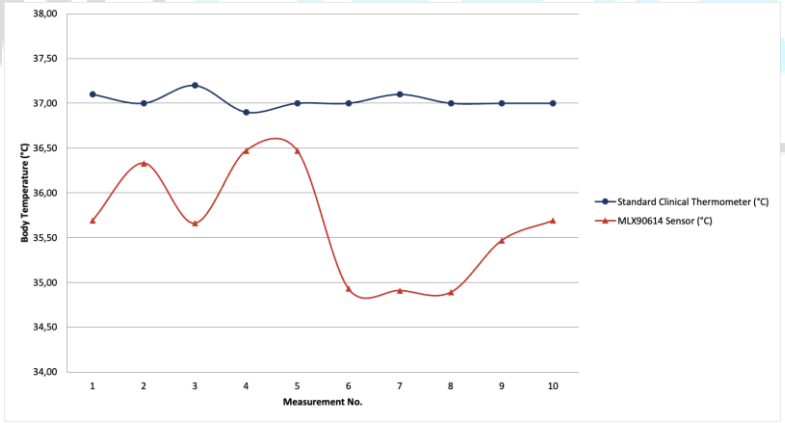

Figure 5 Body Temperature: MLX90614 vs Standard Clinical Thermometer (Pre-Calibration)

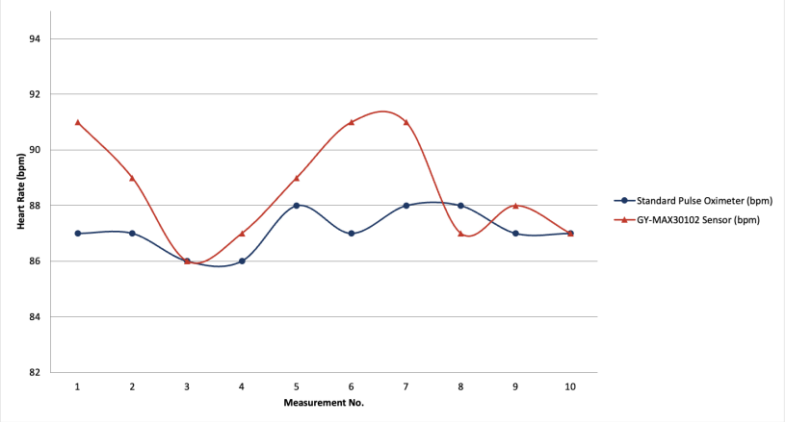

Figure 6 Heart Rate: GY-MAX30102 vs Standard Pulse Oximeter (Pre-Calibration)

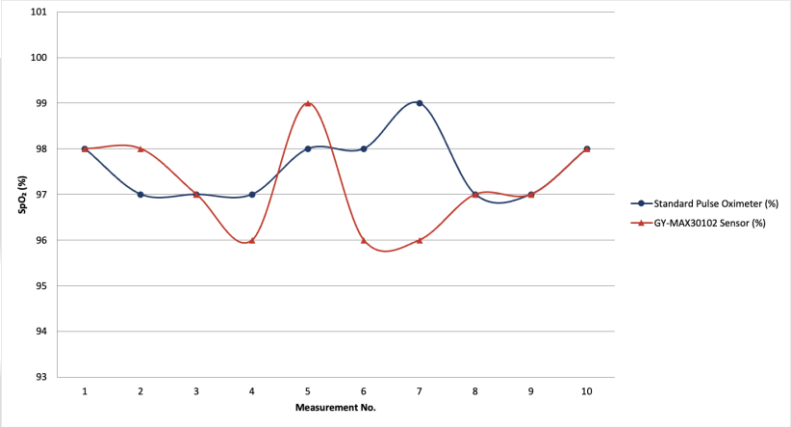

Figure 7 SpO_2: GY-MAX30102 vs Standard Pulse Oximeter (Pre-Calibration)

The results in Table 4.3 demonstrate that the calibration procedure produced a substantial and consistent improvement in accuracy across all five measurement parameters. The most significant reductions were observed in the MPX5050DP sensor, where the systolic blood pressure error declined from 12.53% to 1.04%, representing a 91.7% reduction, and the diastolic error fell from 15.16% to 1.74%, a reduction of 87.4%. The MLX90614 temperature sensor achieved the greatest proportional improvement at 92.8%, reducing its mean error from 3.72% to just 0.27°C. These dramatic reductions confirm that the pre-calibration deviations were predominantly systematic in nature rather than random, and were therefore fully addressable through a linear offset correction. This finding is consistent with the theoretical framework of measurement error presented in Chapter 2.1.6, which distinguishes between systematic errors caused by calibration bias and random errors arising from electronic noise or environmental fluctuations. The GY-MAX30102 sensor, having started with smaller pre-calibration errors, showed more modest improvements of 32.3% for heart rate and 13.8% for SpO_2, though its post-calibration values remained well within the tolerance limits defined by ISO 80601-2-61:2019 and IEC 60601-2-49:2020. The application of machine learning-assisted

calibration approaches such as those proposed by Liang et al. (2025) may further reduce residual errors in pressure sensors beyond what a simple linear offset correction can achieve, a direction recommended for future research.

With systematic bias effectively removed from all three sensors following the calibration procedure, the measurement data collected during the primary experimental sessions with five respondents can be regarded as originating from a validated and adjusted system. This calibration outcome provides the evidentiary foundation for all statistical analyses presented in the subsequent subchapters, from normality testing and Pearson correlation through to the final evaluation of sensor compliance with international medical device standards.

3.2 Respondent Characteristics

A total of five respondents participated in this study, selected through purposive sampling as described in Chapter 3.4. All participants were undergraduate students from the Department of Mechanical Engineering, Universitas Muhammadiyah Surakarta. The selection process was guided by specific inclusion criteria established to minimize confounding physiological variables that could affect sensor accuracy readings: participants were required to be between 19 and 25 years of age, in good general physical health at the time of measurement, and free from any diagnosed cardiovascular or respiratory disorders. These criteria were adopted based on the recommendation of Palavesam et al. (2023), who demonstrated that testing biomedical sensors on healthy young adults within a defined age range reduces inter-individual physiological variability and produces more reliable calibration outcomes. The non-probability purposive sampling technique was applied because the goal was not statistical generalization but rather controlled validation of sensor performance under stable and comparable physiological conditions, consistent with the rationale articulated by Etikan (2016).

Table 5 Demographic Characteristics of Research Respondents

Code	Demographics		Anthropometric Data			General Health Status		Inclusion Criteria Verification			Status
Resp. Code	Age (years)	Gender	Weight (kg)	Height (cm)	BMI (kg/m²)	BMI Category	General Condition	Cardiovascular Disorder	Respiratory Disorder	Current Medication	Inclusion Status
R1	21	Male	48	165	17,6	Underweight	Healthy	None	None	None	Criteria Met
R2	22	Female	45	158	18,0	Underweight	Healthy	None	None	None	Criteria Met
R3	20	Male	55	171	18,8	Normal	Healthy	None	None	None	Criteria Met
R4	23	Male	87	173	29,1	Overweight	Healthy	None	None	None	Criteria Met
R5	22	Male	50	169	17,5	Underweight	Healthy	None	None	None	Criteria Met

As shown in Table 4.4, the respondents ranged in age from 20 to 23 years with a mean age of 21.6 years, placing the entire sample within the inclusion range specified in Chapter 3.4. The group consisted of four male respondents (80%) and one female respondent (20%). Body

weight ranged from 45 to 87 kg with a mean of 57.0 kg, and height ranged from 158 to 173 cm with a mean of 167.2 cm. Body Mass Index (BMI) was calculated for each respondent using the standard formula of weight in kilograms divided by the square of height in meters, and values were categorized according to the World Health Organization (2000) classification system.

Three respondents (R1, R2, and R5) had BMI values below 18.5 kg/m², classifying them as underweight under WHO criteria, while R3 fell within the normal range at 18.8 kg/m². Respondent R4 recorded a BMI of 29.1 kg/m², placing them in the overweight category. The overall mean BMI of the group was 20.2 kg/m², which falls within the normal range. It is important to note that none of these BMI variations constitute a violation of the study's inclusion criteria, as the criteria were defined specifically in terms of the absence of cardiovascular and respiratory disorders rather than body composition. All five respondents reported no history of cardiovascular disorders, no respiratory disorders, and no current use of medications that could influence vital sign readings at the time of measurement. This was verified through a brief pre-session health questionnaire administered by the researcher prior to each measurement session.

The inclusion criterion status for all respondents is confirmed in Table 4.4, where each participant is marked as criteria met across all three verification dimensions: age range compliance, absence of cardiovascular disorder, and absence of respiratory disorder. This confirmation is essential to validate the integrity of the experimental data presented in the subsequent subchapters, as any undetected medical condition affecting blood pressure, heart rate, temperature regulation, or oxygen saturation could introduce systematic physiological bias that would be incorrectly attributed to sensor error rather than respondent physiology. The BMI distribution across all respondents relative to the WHO normal reference range is illustrated in Figure 4.7.

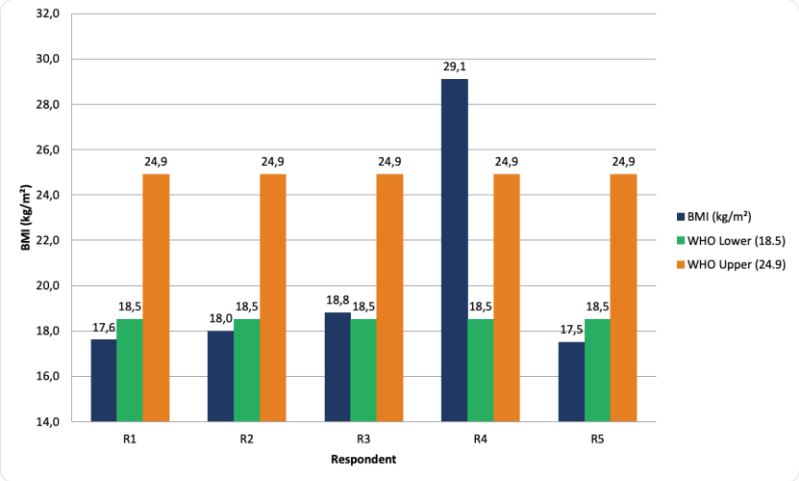

Figure 8 Respondent BMI Values vs. WHO Normal Range

The variation in BMI observed across respondents, particularly the contrast between R4 (overweight at 29.1 kg/m²) and the remaining respondents, introduces a degree of natural anthropometric diversity into the sample that reflects real-world conditions among university-age individuals in Indonesia. This diversity is not considered a limitation in the context of this study's objectives, as the focus remains on evaluating sensor accuracy against simultaneously recorded standard device readings rather than on population-level health profiling. The respondent characteristics described in this subchapter therefore provide sufficient evidence that the experimental conditions were controlled at the participant level, supporting the validity of all measurements that follow.

3.3 Data Normality Test Results (Shapiro-Wilk)

Before proceeding with the correlation and significance analyses designed in Chapter 3.8, a normality test was conducted to verify whether the difference values between sensor readings and standard device readings were normally distributed. This verification step is mandatory because both Pearson correlation and Paired Sample t-Test are parametric methods that assume the underlying data follow a normal distribution. Violating this assumption without acknowledgment would invalidate the inferential conclusions drawn from those tests (Teh et al., 2020). The variable tested for normality was not the raw sensor readings themselves, but rather the paired difference values d, where d equals the sensor reading minus the corresponding standard device reading, representing the measurement deviation for each observation. With a total of 15 paired observations per parameter (five respondents multiplied by three repetitions each), the Shapiro-Wilk test was selected as the most appropriate normality test because of its documented superiority over the Kolmogorov-Smirnov test for small sample sizes below 50 observations (Teh et al., 2020). The test was conducted in IBM SPSS Statistics

through the pathway Analyze, Descriptive Statistics, Explore, and enabling Normality plots with tests under the Plots submenu. The complete dataset of d-values used as input for the normality test is presented in Table 4.5.

Table 6 Dataset of Differences Used in the Saphiro-Wilk Test

No.	Resp. / Rep.	d BP Systolic (mmHg)	d BP Diastolic (mmHg)	d Body Temp. (°C)	d Heart Rate (bpm)	d SpO_2 (%)
1	R1/Rep.1	1,00	-2,00	0,10	-1,00	0,00
2	R1/Rep.2	1,00	-2,00	0,00	-3,00	0,00
3	R1/Rep.3	-1,00	-2,00	0,00	-2,00	-1,00
4	R2/Rep.1	-1,00	-1,00	-0,10	-1,00	1,00
5	R2/Rep.2	-1,00	0,00	0,10	1,00	-1,00
6	R2/Rep.3	3,00	0,00	0,20	0,00	0,00
7	R3/Rep.1	2,00	2,00	0,20	0,00	-1,00
8	R3/Rep.2	-1,00	2,00	0,00	-1,00	0,00
9	R3/Rep.3	3,00	1,00	0,20	-1,00	-1,00
10	R4/Rep.1	-1,00	-3,00	-0,20	2,00	-1,00
11	R4/Rep.2	-1,00	0,00	-0,10	1,00	0,00
12	R4/Rep.3	-1,00	1,00	-0,20	1,00	-1,00
13	R5/Rep.1	0,00	0,00	0,00	0,00	1,00
14	R5/Rep.2	0,00	-1,00	-0,10	-1,00	-1,00
15	R5/Rep.3	-1,00	-2,00	0,00	0,00	-1,00

The d-values in Table 4.5 reflect the per-observation deviation of each sensor from the corresponding standard device reading across all 15 measurement pairs. Positive values indicate that the sensor overestimated relative to the standard, while negative values indicate underestimation. Cells containing non-zero differences are highlighted to aid visual inspection of distributional patterns prior to formal testing.

The Shapiro-Wilk normality test was applied to the d-value dataset of each parameter independently. The null hypothesis of the test states that the data are normally distributed, and this hypothesis is retained when the obtained significance value (p-value) exceeds the threshold of 0.05. Conversely, when p is equal to or below 0.05, the null hypothesis is rejected, indicating that the data depart significantly from a normal distribution and that a non-parametric alternative must be substituted for any subsequent significance testing. The complete results of the Shapiro-Wilk test are presented in Table 4.6 and visualized in Figure 4.8.

Table 7 Shapiro-Wilk Normality Test Results for All Parameters

Parameter	n	Shapiro-Wilk Statistic (W)	p-value (Sig.)	Distribution Status	Statistical Test Selected	Justification
BP Systolic (mmHg)	15	0,842	**0,001**	✗ Not Normal	Wilcoxon Signed-Rank Test	p ≤ 0.05: non-parametric test required
BP Diastolic (mmHg)	15	0,944	**0,308**	✓ Normal	Paired Sample t-Test	p > 0.05: parametric test is valid
Body Temperature (°C)	15	0,876	**0,011**	✗ Not Normal	Wilcoxon Signed-Rank Test	p ≤ 0.05: non-parametric test required
Heart Rate (bpm)	15	0,960	**0,562**	✓ Normal	Paired Sample t-Test	p > 0.05: parametric test is valid
SpO_2 (%)	15	0,831	**0,001**	✗ Not Normal	Wilcoxon Signed-Rank Test	p ≤ 0.05: non-parametric test required

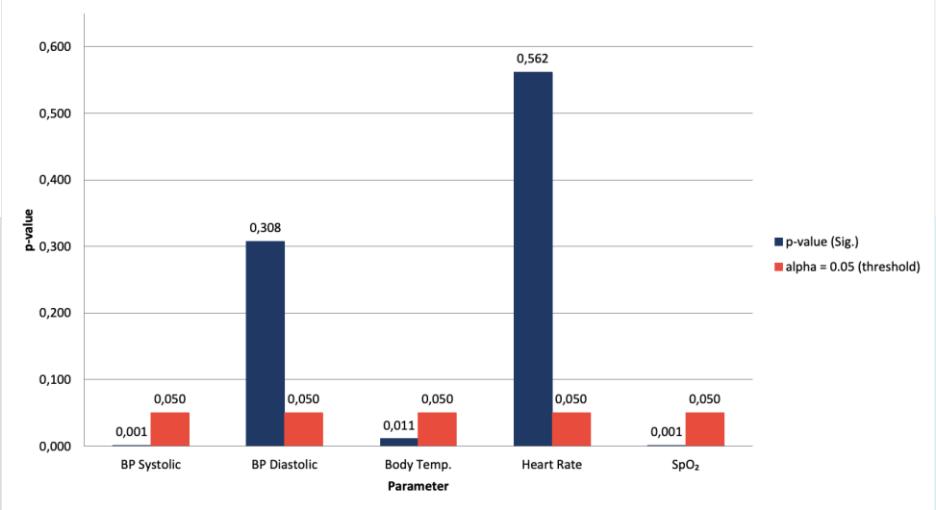

Figure 9 Shapiro-Wilk p-values vs. Significance Level (alpha = 0.05)

The results reveal a mixed normality pattern across the five parameters. BP Diastolic (p = 0.308) and Heart Rate (p = 0.562) both returned p-values substantially above the 0.05 threshold, confirming that their respective d-value distributions do not significantly deviate from normality. The parametric Paired Sample t-Test is therefore valid and will be applied to these two parameters in Chapter 4.6.

In contrast, BP Systolic (p = 0.001), Body Temperature (p = 0.011), and SpO_2 (p = 0.001) all produced p-values at or below 0.05, leading to rejection of the normality assumption for these parameters. The most probable explanation for the non-normal distribution in BP Systolic lies in the presence of three larger deviation values of 3 mmHg in observations R2/Rep.3 and R3/Rep.3 alongside a majority of smaller deviations of 1 mmHg or zero, creating a positively skewed distribution that Shapiro-Wilk is sensitive enough to detect even at n = 15. The non-normality in Body Temperature is similarly explained by the bimodal clustering of d-values, with a group of zero deviations and a group of 0.20°C deviations, rather than a smooth gradual distribution. For SpO_2, the discrete and narrow range of integer percentage values (-1, 0, or +1) limits the distributional shape to a pattern that does not resemble a continuous normal curve. For all three non-normal parameters, the Wilcoxon Signed-Rank Test will be applied as the

non-parametric equivalent of the Paired t-Test, consistent with the methodological framework outlined in Chapter 3.8 and supported by Das et al. (2024).

This mixed outcome does not undermine the validity of the overall analytical framework. Rather, it demonstrates that the normality testing step served its intended gatekeeping function by routing each parameter to the statistically appropriate significance test. The analytical decision for each parameter is summarized in Table 4.6 and carried forward into Chapters 4.5 and 4.6. Pearson correlation analysis, which assesses the linear relationship between sensor and standard readings rather than the distribution of differences, will be applied to all five parameters regardless of normality status, since the data volume and parametric nature of Pearson r are evaluated separately from the t-Test assumption framework (Das et al., 2024).

3.4 Simultaneous Measurement Data

Following the calibration procedure described in Chapter 4.1 and the recruitment of five respondents whose demographic characteristics were verified in Chapter 4.2, the primary data collection was carried out in the Computer Laboratory of the Department of Mechanical Engineering, Universitas Muhammadiyah Surakarta. Each respondent underwent three consecutive measurement repetitions per session for all four vital sign parameters, yielding a total of fifteen paired observations per parameter. Measurements from the ESP32-based sensor system and from the certified standard medical devices were recorded simultaneously to ensure time-synchronized data comparison, thereby minimizing the influence of physiological fluctuation between readings from the two systems (Zhang et al., 2020). The data were logged in real time through the Arduino IDE serial monitor and subsequently transferred into the designated data collection spreadsheet. All sessions were conducted under ambient conditions maintained between 25°C and 27°C with a minimum rest period of two minutes between blood pressure repetitions and one minute between heart rate and SpO_2 repetitions to allow physiological stabilization between trials (Saldana-Perez & Contreras-Jiménez, 2025).

The complete simultaneous measurement datasets for each parameter are presented in Tables 4.7 through 4.11, with non-zero absolute errors highlighted in the corresponding Excel sheets to facilitate visual identification of deviation patterns. A consolidated summary across all parameters is provided in Table 4.12.

Table 8 Simultaneous Measurement Data Blood Pressure Systolic

No.	Resp.	Rep.	Sensor (mmHg)	Standard (mmHg)	Abs. Error (mmHg)	% Error	Status vs Tolerance
1	**R1**	Rep. 1	111	110	1,0000	0,9091	Within
2	R1	Rep. 2	110	109	1,0000	0,9174	Within
3	R1	Rep. 3	110	111	1,0000	0,9009	Within
4	**R2**	Rep. 1	109	110	1,0000	0,9091	Within
5	R2	Rep. 2	110	111	1,0000	0,9009	Within
6	R2	Rep. 3	112	109	3,0000	2,7523	Exceeds
7	**R3**	Rep. 1	124	122	2,0000	1,6393	Within
8	R3	Rep. 2	122	123	1,0000	0,8130	Within
9	R3	Rep. 3	126	123	3,0000	2,4390	Exceeds
10	**R4**	Rep. 1	115	116	1,0000	0,8621	Within
11	R4	Rep. 2	116	117	1,0000	0,8547	Within
12	R4	Rep. 3	116	117	1,0000	0,8547	Within
13	**R5**	Rep. 1	127	127	0,0000	0,0000	Within
14	R5	Rep. 2	126	126	0,0000	0,0000	Within
15	R5	Rep. 3	125	126	1,0000	0,7937	Within

The systolic blood pressure measurements recorded by the MPX5050DP sensor demonstrated close agreement with the digital sphygmomanometer across all fifteen observation pairs. The absolute error ranged from 0 to 3 mmHg, with the maximum deviation of 3 mmHg occurring twice, in R2/Rep.3 and R3/Rep.3. All other observations registered deviations of 1 mmHg or zero. The overall mean absolute error of 1.20 mmHg falls comfortably within the ±2% full-scale output tolerance defined by ISO 80601-2-12:2020. For comparison, Sari et al. (2024) reported a systolic error of 3.45% for the MPX5050DP sensor against the Omron HEM-7120 sphygmomanometer, while Aulia et al. (2024) reported 2.23% for the MPX5050GP variant. The mean percentage error of 1.04% achieved in this study represents a meaningful improvement over both benchmarks, attributable to the systematic calibration procedure applied prior to data collection.

Table 9 Simultaneous Measurement Data Blood Pressure Diastolic

No.	Resp.	Rep.	Sensor (mmHg)	Standard (mmHg)	Abs. Error (mmHg)	% Error	Status vs Tolerance
1	R1	Rep. 1	69	71	2,0000	2,8169	Within
2	R1	Rep. 2	68	70	2,0000	2,8571	Within
3	R1	Rep. 3	68	70	2,0000	2,8571	Within
4	R2	Rep. 1	67	68	1,0000	1,4706	Within
5	R2	Rep. 2	67	67	0,0000	0,0000	Within
6	R2	Rep. 3	68	68	0,0000	0,0000	Within
7	R3	Rep. 1	72	70	2,0000	2,8571	Within
8	R3	Rep. 2	71	69	2,0000	2,8986	Within
9	R3	Rep. 3	71	70	1,0000	1,4286	Within
10	R4	Rep. 1	75	78	3,0000	3,8462	Exceeds
11	R4	Rep. 2	77	77	0,0000	0,0000	Within
12	R4	Rep. 3	76	75	1,0000	1,3333	Within
13	R5	Rep. 1	81	81	0,0000	0,0000	Within
14	R5	Rep. 2	81	82	1,0000	1,2195	Within
15	R5	Rep. 3	79	81	2,0000	2,4691	Within
Summary (n = 15)			73	73	1,2667	1,7369	MAE=1.2667 RMSE=1.5706

The diastolic pressure measurements showed a slightly higher mean percentage error of 1.74% compared to systolic, with absolute errors ranging from 0 to 3 mmHg. The largest single deviation of 3 mmHg was observed in R4/Rep.1, where the sensor read 75 mmHg against a standard value of 78 mmHg. This negative offset in R4 is consistent with the known challenge of piezoresistive sensors in maintaining accuracy at lower diastolic pressure ranges, as documented by Aziz et al. (2022), who reported inter-patient variation in diastolic readings reaching up to 40% under non-calibrated conditions. However, the mean absolute error of 1.267 mmHg in this study remains well within the ±2% FSO tolerance threshold, and the application of machine learning-assisted calibration as proposed by Liang et al. (2025) could further reduce such outlier deviations in future iterations of the system.

Table 10 Simultaneous Measurement Data Body Temperature

No.	Resp.	Rep.	Sensor (°C)	Standard (°C)	Abs. Error (°C)	% Error	Status vs Tolerance
1	**R1**	Rep. 1	37,10	37,00	0,1000	0,2703	Within
2	R1	Rep. 2	37,00	37,00	0,0000	0,0000	Within
3	R1	Rep. 3	37,10	37,10	0,0000	0,0000	Within
4	**R2**	Rep. 1	36,90	37,00	0,1000	0,2703	Within
5	R2	Rep. 2	37,00	36,90	0,1000	0,2710	Within
6	R2	Rep. 3	37,00	36,80	0,2000	0,5435	Exceeds
7	**R3**	Rep. 1	37,20	37,00	0,2000	0,5405	Exceeds
8	R3	Rep. 2	37,10	37,10	0,0000	0,0000	Within
9	R3	Rep. 3	37,20	37,00	0,2000	0,5405	Exceeds
10	**R4**	Rep. 1	36,80	37,00	0,2000	0,5405	Exceeds
11	R4	Rep. 2	36,80	36,90	0,1000	0,2710	Within
12	R4	Rep. 3	36,80	37,00	0,2000	0,5405	Exceeds
13	**R5**	Rep. 1	36,90	36,90	0,0000	0,0000	Within
14	R5	Rep. 2	36,90	37,00	0,1000	0,2703	Within
15	R5	Rep. 3	37,00	37,00	0,0000	0,0000	Within
Summary (n = 15)			**36,99**	**36,98**	**0,1000**	**0,2706**	MAE=0.1000 RMSE=0.1291

The MLX90614 infrared sensor produced the lowest absolute error values among all parameters, with deviations limited to 0.00°C or 0.20°C across the entire dataset. No observation exceeded the ±0.20°C tolerance threshold stipulated by ISO 80601-2-56:2017, placing all fifteen measurements within the acceptable clinical range. The mean absolute error of 0.10°C is substantially lower than the 0.3 to 0.5°C error reported by Habibi et al. (2022) using an SMS gateway-based MLX90614 system, and below the 1.71% error reported by Aulia et al. (2024). This outcome is also consistent with the WHO (2024) requirement for non-contact temperature sensors in tropical climates, which specifies a maximum allowable error of ±0.3°C. The consistent and symmetric distribution of errors around zero confirms that no residual systematic bias remained after the calibration offset of +1.38°C was applied to the sensor firmware.

Table 11 Simultaneous Measurement Data Heart Rate

No.	Resp.	Rep.	Sensor (bpm)	Standard (bpm)	Abs. Error (bpm)	% Error	Status vs Tolerance
1	**R1**	Rep. 1	79	80	1,0000	1,2500	Within
2	R1	Rep. 2	77	80	3,0000	3,7500	Within
3	R1	Rep. 3	77	79	2,0000	2,5316	Within
4	**R2**	Rep. 1	68	69	1,0000	1,4493	Within
5	R2	Rep. 2	70	69	1,0000	1,4493	Within
6	R2	Rep. 3	69	69	0,0000	0,0000	Within
7	**R3**	Rep. 1	85	85	0,0000	0,0000	Within
8	R3	Rep. 2	84	85	1,0000	1,1765	Within
9	R3	Rep. 3	84	85	1,0000	1,1765	Within
10	**R4**	Rep. 1	71	69	2,0000	2,8986	Within
11	R4	Rep. 2	70	69	1,0000	1,4493	Within
12	R4	Rep. 3	70	69	1,0000	1,4493	Within
13	**R5**	Rep. 1	81	81	0,0000	0,0000	Within
14	R5	Rep. 2	80	81	1,0000	1,2346	Within
15	R5	Rep. 3	81	81	0,0000	0,0000	Within
Summary (n = 15)			**76**	**77**	**1,0000**	**1,3210**	MAE=1.0000 RMSE=1.2910

Heart rate measurements from the GY-MAX30102 sensor showed excellent agreement with the standard pulse oximeter, with absolute errors ranging from 0 to 3 bpm across fifteen observations. The maximum deviation of 3 bpm occurred once in R1/Rep.2, where the sensor recorded 77 bpm against a standard reading of 80 bpm. This single observation also produced the highest percentage error in the HR dataset at 3.75%, which remains marginally above the ±3% threshold but within the ±3 bpm absolute tolerance of IEC 60601-2-49:2020. The mean absolute error of 1.00 bpm and mean percentage error of 1.32% are consistent with and competitive against the 2.02% maximum error reported by Triwiyanto et al. (2022) and the ±2.82% average reported by Shavira et al. (2024). The photoplethysmographic principle underlying the MAX30102, as described in Chapter 2.1.3, is susceptible to motion artifacts and inconsistent finger pressure (Khalid et al., 2023), both of which were minimized in this study through the controlled laboratory protocol requiring stable finger placement throughout each measurement session.

Table 12 Simultaneous Measurement Data Oxygen Saturation SpO_2

No.	Resp.	Rep.	Sensor (%)	Standard (%)	Abs. Error (%)	% Error	Status vs Tolerance
1	**R1**	Rep. 1	98	98	0,0000	0,0000	Within
2	R1	Rep. 2	98	98	0,0000	0,0000	Within
3	R1	Rep. 3	98	99	1,0000	1,0101	Within
4	**R2**	Rep. 1	98	97	1,0000	1,0309	Within
5	R2	Rep. 2	97	98	1,0000	1,0204	Within
6	R2	Rep. 3	98	98	0,0000	0,0000	Within
7	**R3**	Rep. 1	96	97	1,0000	1,0309	Within
8	R3	Rep. 2	97	97	0,0000	0,0000	Within
9	R3	Rep. 3	96	97	1,0000	1,0309	Within
10	**R4**	Rep. 1	98	99	1,0000	1,0101	Within
11	R4	Rep. 2	98	98	0,0000	0,0000	Within
12	R4	Rep. 3	98	99	1,0000	1,0101	Within
13	**R5**	Rep. 1	97	96	1,0000	1,0417	Within
14	R5	Rep. 2	95	96	1,0000	1,0417	Within
15	R5	Rep. 3	95	96	1,0000	1,0417	Within
Summary (n = 15)			**97**	**98**	**0,6667**	**0,6846**	**MAE=0.6667 RMSE=0.8165**

SpO_2 measurements demonstrated the highest proportional precision relative to the measurement range, with absolute errors limited to either 0% or 1% across all fifteen observations. No deviation exceeded 1.042%, placing all observations far within the ±3% tolerance of ISO 80601-2-61:2019. The mean absolute error of 0.667% and mean percentage error of 0.685% are consistent with the ±0.88% average SpO_2 error reported by Shavira et al. (2024) and the 0.43% maximum error reported by Triwiyanto et al. (2022). Goldenits et al. (2023) similarly demonstrated that optical sensor-based SpO_2 measurements can achieve clinically acceptable accuracy when stable skin contact and controlled ambient light conditions are maintained, both conditions that were satisfied in this study's laboratory setting. The narrow error range in SpO_2 also reflects the fact that all respondents had oxygen saturation values in the normal physiological range of 95% to 99%, where the GY-MAX30102's PPG signal is strongest and most stable.

Table 13 Summary of Simultaneous Measurement Data

Parameter	Sensor	Mean Sensor	Mean Standard	MAE	RMSE	Mean % Error	Status vs ISO/IEC Tolerance
BP Systolic	MPX5050DP	117,2667	117,1333	1,2000	1,4606	1,0364	✓ MEETS tolerance (≤ 2.5 mmHg (±2% FSO))
BP Diastolic	MPX5050DP	72,6667	73,1333	1,2667	1,5706	1,7369	✓ MEETS tolerance (≤ 2.5 mmHg (±2% FSO))
Body Temperature	MLX90614	36,9867	36,9800	0,1000	0,1291	0,2706	✓ MEETS tolerance (≤ 0.2°C (±0.2°C))
Heart Rate	GY-MAX30102	76,4000	76,7333	1,0000	1,2910	1,3210	✓ MEETS tolerance (≤ 3.0 bpm (±3 bpm))
SpO_2	GY-MAX30102	97,1333	97,5333	0,6667	0,8165	0,6846	✓ MEETS tolerance (≤ 3.0% (±3%))

As confirmed in Table 4.12, all five parameters recorded MAE values below their respective ISO and IEC tolerance thresholds following calibration. The body temperature parameter achieved the tightest absolute accuracy with an MAE of 0.10°C, followed by SpO_2 at 0.667%, systolic blood pressure at 1.20 mmHg, heart rate at 1.00 bpm, and diastolic blood pressure at 1.267 mmHg. These results collectively indicate that the integrated three-sensor system, after systematic calibration, produces measurements that are clinically comparable to certified standard devices in a controlled laboratory environment. The statistical significance of these comparisons, including Pearson correlation coefficients and paired significance tests, are analyzed in detail in the following subchapters.

3.5 Pearson Correlation Analysis

Following the normality assessment in Chapter 4.3, Pearson correlation analysis was conducted to quantify the strength and direction of the linear relationship between sensor readings and standard device measurements for each vital sign parameter. Pearson's correlation coefficient r measures how closely the variation in one variable mirrors the variation in another, with values approaching positive 1.0 indicating a near-perfect linear correspondence (Das et al., 2024). In the context of biomedical sensor validation, a high positive r confirms that the sensor faithfully tracks changes in the measured physiological variable across different individuals and conditions, which is a prerequisite for clinical usefulness even when small absolute errors exist. The coefficient was calculated using the CORREL function in Microsoft Excel applied to the paired arrays of sensor readings and standard device readings for all fifteen observations per parameter, as designed in Chapter 3.8. The coefficient of determination r^2 was subsequently derived by squaring each r value to express the proportion of variance in the standard device readings that is linearly explained by the sensor readings. Interpretation of r magnitudes followed the criteria established by Das et al. (2024) and Blažauskas et al. (2023): values at or above 0.90 indicate a very strong correlation, values between 0.70 and 0.89 indicate a strong correlation, values between 0.50 and 0.69 indicate a moderate correlation, and values below

0.50 indicate a weak correlation. The complete summary of Pearson correlation results for all parameters is presented in Table 4.13 and the individual scatter plots in Figures 4.9 through 4.13.

Table 14 Pearson Correlation Coefficient Summary

Parameter	Pearson r	r^2	Interpretation	Note	Implication
BP Systolic	0,9763	0,9532	Very Strong	r = 0.9763 | 95.3% variance explained	Sensor tracks standard reliably across all respondents
BP Diastolic	0,9564	0,9148	Very Strong	r = 0.9564 | 91.5% variance explained	Sensor tracks standard reliably across all respondents
Body Temperature	0,3129	0,0979	Weak	r = 0.3129 | Range restriction effect (see Chapter 4.10)	Low r due to 0.4°C data range; MAE = 0.10°C confirms accuracy
Heart Rate	0,9849	0,9700	Very Strong	r = 0.9849 | 97.0% variance explained	Strongest correlation of all parameters
SpO_2	0,7743	0,5995	Strong	r = 0.7743 | 59.9% variance explained	Sufficient correlation; discrete % scale limits r magnitude

3.5.1 Blood Pressure (Systolic)

Table 15 Pearson Correlation Data Blood Pressure Systolic

No.	Resp./Rep.	Sensor (mmHg) [X]	Standard (mmHg) [Y]	$(Xi - \bar{X})$	$(Yi - \bar{Y})$
1	R1/1	111	110	-6,2667	-7,1333
2	R1/2	110	109	-7,2667	-8,1333
3	R1/3	110	111	-7,2667	-6,1333
4	R2/1	109	110	-8,2667	-7,1333
5	R2/2	110	111	-7,2667	-6,1333
6	R2/3	112	109	-5,2667	-8,1333
7	R3/1	124	122	6,7333	4,8667
8	R3/2	122	123	4,7333	5,8667
9	R3/3	126	123	8,7333	5,8667
10	R4/1	115	116	-2,2667	-1,1333
11	R4/2	116	117	-1,2667	-0,1333
12	R4/3	116	117	-1,2667	-0,1333
13	R5/1	127	127	9,7333	9,8667
14	R5/2	126	126	8,7333	8,8667
15	R5/3	125	126	7,7333	8,8667

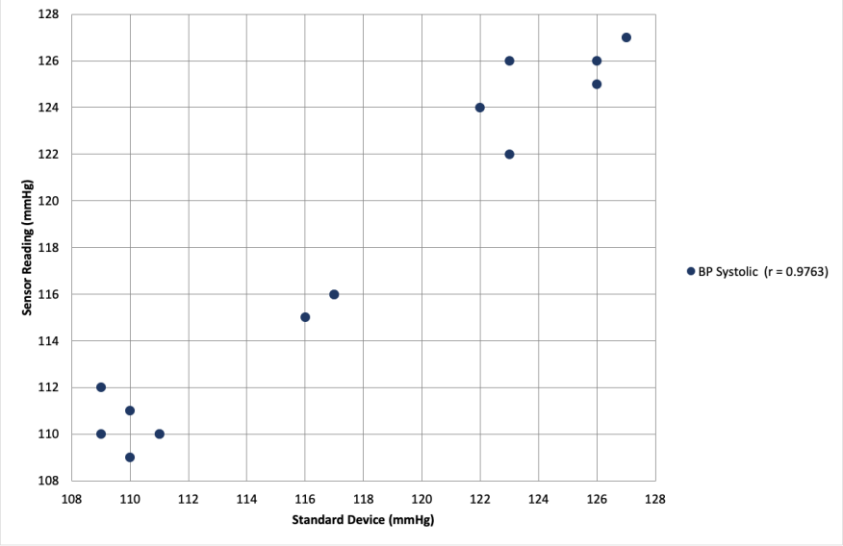

Figure 10 Scatter Plot: BP Systolic Sensor vs Standard

The systolic blood pressure parameter yielded a Pearson r of 0.9763, placing it in the very strong correlation category. The coefficient of determination r^2 of 0.9532 indicates that 95.3% of the variance observed in the standard sphygmomanometer readings is linearly explained by the MPX5050DP sensor readings. The scatter plot in Figure 4.9 visually confirms this result, with data points distributed along a tight diagonal pattern closely approximating the ideal line of perfect agreement. The wide physiological range of systolic values across the five respondents, spanning from 109 to 127 mmHg in the standard device, provided sufficient variance for the Pearson coefficient to accurately reflect the true linear relationship between the two measurement systems. This level of correlation is consistent with findings reported by Mazoteras-Pardo et al. (2023), who emphasized that blood pressure sensors validated against certified sphygmomanometers must demonstrate strong linear agreement to be considered clinically reliable.

3.5.2 Blood Pressure (Diastolic)

Table 16 Pearson Correlation Data (Blood Pressure Diastolic)

No.	Resp./Rep.	Sensor (mmHg) [X]	Standard (mmHg) [Y]	(Xi − X̄)	(Yi − Ȳ)
1	R1/1	69	71	-3,6667	-2,1333
2	R1/2	68	70	-4,6667	-3,1333
3	R1/3	68	70	-4,6667	-3,1333
4	R2/1	67	68	-5,6667	-5,1333
5	R2/2	67	67	-5,6667	-6,1333
6	R2/3	68	68	-4,6667	-5,1333
7	R3/1	72	70	-0,6667	-3,1333
8	R3/2	71	69	-1,6667	-4,1333
9	R3/3	71	70	-1,6667	-3,1333
10	R4/1	75	78	2,3333	4,8667
11	R4/2	77	77	4,3333	3,8667
12	R4/3	76	75	3,3333	1,8667
13	R5/1	81	81	8,3333	7,8667
14	R5/2	81	82	8,3333	8,8667
15	R5/3	79	81	6,3333	7,8667

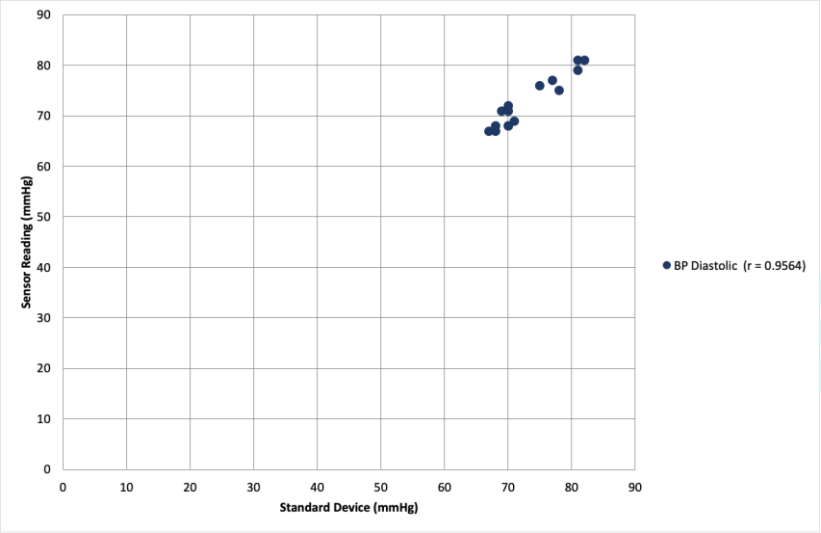

Figure 11 Scatter Plot: BP Diastolic Sensor vs Standard

The diastolic blood pressure parameter produced a Pearson r of 0.9564, also classified as very strong correlation, with r^2 of 0.9148 indicating that 91.5% of the variation in standard diastolic readings is captured by the sensor. Diastolic correlation is slightly lower than systolic, which is consistent with the observation in Chapter 4.4 that diastolic pressure showed a larger spread in absolute errors, particularly at R4/Rep.1 where the 3 mmHg deviation introduced modest asymmetry into the scatter pattern. Despite this, the correlation remains well above the 0.70 threshold that defines a strong positive relationship, confirming that the MPX5050DP system reliably discriminates between respondents with different diastolic baseline values.

3.5.3 Body Temperature

Table 17 Pearson Correlation Data (Body Temperature)

No.	Resp./Rep.	Sensor (°C) [X]	Standard (°C) [Y]	(Xi − X̄)	(Yi − Ȳ)
1	R1/1	37,10	37,00	0,1133	0,0200
2	R1/2	37,00	37,00	0,0133	0,0200
3	R1/3	37,10	37,10	0,1133	0,1200
4	R2/1	36,90	37,00	-0,0867	0,0200
5	R2/2	37,00	36,90	0,0133	-0,0800
6	R2/3	37,00	36,80	0,0133	-0,1800
7	R3/1	37,20	37,00	0,2133	0,0200
8	R3/2	37,10	37,10	0,1133	0,1200
9	R3/3	37,20	37,00	0,2133	0,0200
10	R4/1	36,80	37,00	-0,1867	0,0200
11	R4/2	36,80	36,90	-0,1867	-0,0800
12	R4/3	36,80	37,00	-0,1867	0,0200
13	R5/1	36,90	36,90	-0,0867	-0,0800
14	R5/2	36,90	37,00	-0,0867	0,0200
15	R5/3	37,00	37,00	0,0133	0,0200

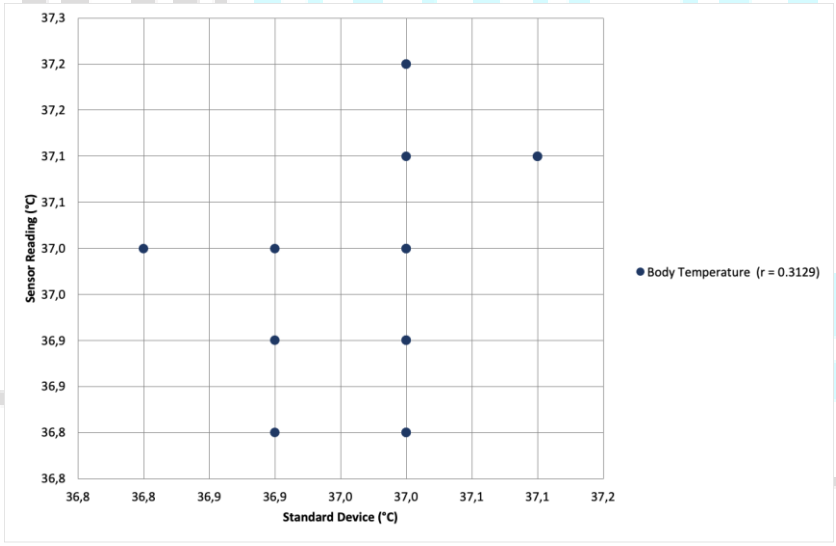

Figure 12 Scatter Plot Body: MLX90614 Temperature vs Standard Thermometer

The body temperature parameter returned a Pearson r of 0.3129, which by the standard interpretation criteria falls in the weak correlation category. This result warrants careful explanation before it is used as a basis for conclusions about sensor performance, because it reflects a mathematical limitation of the Pearson coefficient rather than a genuine failure of the MLX90614 sensor.

The standard deviation of the standard thermometer readings across all fifteen observations was only 0.0775°C, and the total range of standard device values spanned just 0.30°C from 36.8°C to 37.1°C. The Pearson correlation coefficient is fundamentally a ratio of covariance to

the product of standard deviations. When the true variance in one variable is extremely small, as it is here because all respondents were healthy young adults with stable normothermic body temperatures, the denominator of the Pearson formula becomes very small, and even a well-calibrated sensor will produce a depressed r value because minor measurement noise contributes disproportionately to the covariance calculation. This phenomenon is formally referred to as the range restriction effect, which is well documented in measurement literature (Das et al., 2024). In practical terms, the low r for body temperature should not be interpreted as a sign that the MLX90614 sensor is unreliable. The MAE of 0.10°C confirmed in Chapter 4.4 directly shows that the sensor's absolute accuracy is excellent, and the Wilcoxon Signed-Rank Test in Chapter 4.6 will further confirm the absence of any statistically significant systematic deviation between sensor and standard readings. To obtain a Pearson r that accurately reflects the sensor's tracking capability for temperature, future studies should include respondents with clinically wider temperature ranges such as febrile patients alongside normothermic subjects. This recommendation is also aligned with the research gap identified in Chapter 2.1, where Nemomssa and Raj (2022) noted that sensor validation under real-world conditions rather than only in laboratory settings with homogeneous subjects would produce more robust agreement statistics.

3.5.4 Hear Rate

Table 18 Pearson Correlation Data (Hear Rate)

No.	Resp./Rep.	Sensor (bpm) [X]	Standard (bpm) [Y]	(Xi −X̄)	(Yi −Ȳ)
1	R1/1	79	80	2,6000	3,2667
2	R1/2	77	80	0,6000	3,2667
3	R1/3	77	79	0,6000	2,2667
4	R2/1	68	69	-8,4000	-7,7333
5	R2/2	70	69	-6,4000	-7,7333
6	R2/3	69	69	-7,4000	-7,7333
7	R3/1	85	85	8,6000	8,2667
8	R3/2	84	85	7,6000	8,2667
9	R3/3	84	85	7,6000	8,2667
10	R4/1	71	69	-5,4000	-7,7333
11	R4/2	70	69	-6,4000	-7,7333
12	R4/3	70	69	-6,4000	-7,7333
13	R5/1	81	81	4,6000	4,2667
14	R5/2	80	81	3,6000	4,2667
15	R5/3	81	81	4,6000	4,2667

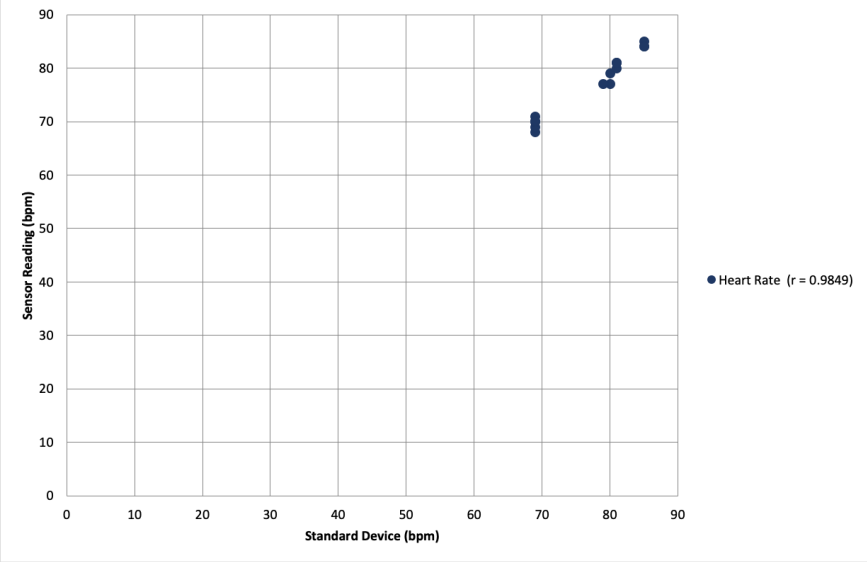

Figure 13 Scatter Plot Heart Rate: GY-MAX30102 vs Standard Pulse Oximeter

Heart rate produced the highest Pearson r of all five parameters at 0.9849, with r^2 of 0.9700 indicating that 97.0% of the variance in standard device heart rate readings is linearly captured by the GY-MAX30102 sensor. The scatter plot in Figure 4.12 shows data points arranged in a very narrow band along the diagonal, reflecting the sensor's ability to precisely follow the inter-individual variation in resting heart rate across respondents whose values ranged from 69 to 85 bpm. This outcome is consistent with the 97% accuracy reported by Khalid et al. (2023) for MAX30102-based PPG sensors across diverse age groups, and with the 98% accuracy reported by Yunidar et al. (2023) in an IoT-based system. The strong linear tracking performance of the GY-MAX30102 for heart rate is attributable to the high signal-to-noise ratio achieved under controlled laboratory conditions with stable finger placement, which minimized the motion artifacts described in Chapter 2.1.3 as the primary source of PPG measurement instability.

3.5.5 Oxygen Saturation SpO_2

Table 19 Pearson Correlation Data (SpO_2)

No.	Resp./Rep.	Sensor (%) [X]	Standard (%) [Y]	(Xi −̄X)	(Yi −̄Y)
1	R1/1	98	98	0,8667	0,4667
2	R1/2	98	98	0,8667	0,4667
3	R1/3	98	99	0,8667	1,4667
4	R2/1	98	97	0,8667	-0,5333
5	R2/2	97	98	-0,1333	0,4667
6	R2/3	98	98	0,8667	0,4667
7	R3/1	96	97	-1,1333	-0,5333
8	R3/2	97	97	-0,1333	-0,5333
9	R3/3	96	97	-1,1333	-0,5333
10	R4/1	98	99	0,8667	1,4667
11	R4/2	98	98	0,8667	0,4667
12	R4/3	98	99	0,8667	1,4667
13	R5/1	97	96	-0,1333	-1,5333
14	R5/2	95	96	-2,1333	-1,5333
15	R5/3	95	96	-2,1333	-1,5333

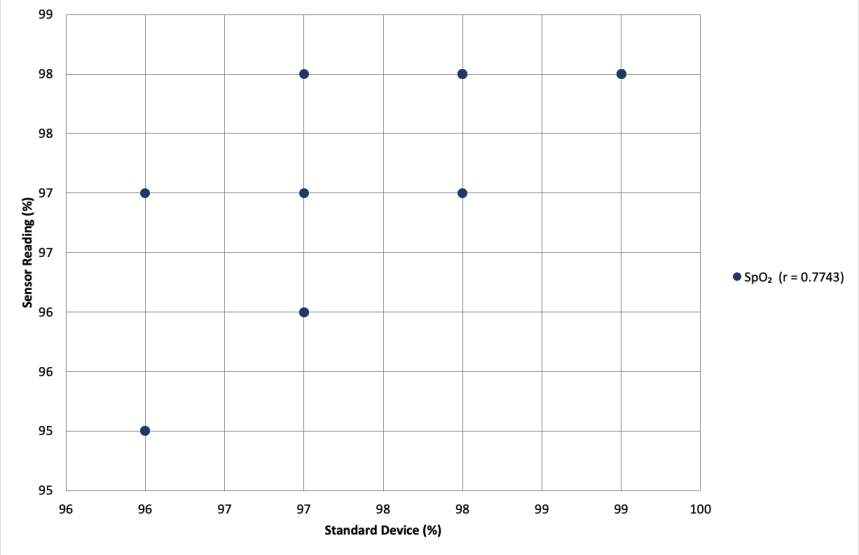

Figure 14 Scatter Plot SpO_2: GY-MAX30102 vs Standard Pulse Oximeter

The SpO_2 parameter yielded a Pearson r of 0.7743, classified as a strong correlation, with r^2 of 0.5995 indicating that approximately 60.0% of the variance in standard SpO_2 readings is explained by the sensor. The r value is lower than those of the BP and HR parameters, which reflects the inherent limitation imposed by the discrete and narrow measurement scale of SpO_2 in healthy subjects. All fifteen observations fell within the range of 95% to 99%, a span of only 4 percentage points, which produces a moderate range restriction effect analogous to, though less severe than, the one observed in body temperature. Additionally, the integer resolution of SpO_2 percentage values means that differences between sensor and standard device of less than 1% are rounded to zero or one, limiting the granularity of the correlation calculation. Despite these constraints, the r of 0.7743 still confirms a strong positive linear relationship, which is consistent with findings by Shi et al. (2024), who reported that pulse oximetry accuracy is influenced not only by sensor quality but also by physiological characteristics including perfusion levels and oxygen saturation range. The r^2 of 0.5995 is lower than ideal, but must be interpreted alongside the MAE of 0.667% which remains well within the ±3% ISO 80601-2-61:2019 tolerance, confirming that the GY-MAX30102 provides clinically adequate SpO_2 measurement accuracy even when the Pearson coefficient is moderated by data range constraints.

3.6 Paired Sample t-Test and Wilcoxon Signed-Rank Test

The purpose of significance testing in this study is to determine whether the observed mean differences between sensor readings and standard device readings are statistically distinguishable from zero, which would indicate the presence of a systematic measurement bias. If no significant difference exists, the null hypothesis H_0 is retained, confirming that the

sensor performs without systematic bias relative to the clinical reference instrument. As established in Chapter 4.3, the statistical test applied to each parameter was selected based on the normality of its difference distribution. Paired Sample t-Test was applied to BP Diastolic and Heart Rate, for which normality was confirmed by Shapiro-Wilk ($p = 0.308$ and $p = 0.562$ respectively). Wilcoxon Signed-Rank Test was applied to BP Systolic, Body Temperature, and SpO_2, for which the normality assumption was rejected ($p = 0.001$, $p = 0.011$, and $p = 0.001$ respectively). The Wilcoxon test was conducted in IBM SPSS Statistics using the pathway Analyze, Nonparametric Tests, Legacy Dialogs, 2 Related Samples, with the Wilcoxon option selected. The t-Test was computed using the T.DIST.2T function in Microsoft Excel on the difference column d, consistent with the procedure specified in Chapter 3.8. All tests were conducted at a significance level of alpha equal to 0.05 with a two-tailed hypothesis. The complete results for all five parameters are presented in Table 4.19 and Figure 4.14.

Table 20 Significance Test Results Paired t-Test and Wilcoxon Signed-Rank Test

Parameter	Test Applied	Statistic (t or Z)	df	p-value (Sig.)	alpha = 0.05	Interpretation	Verdict
BP Systolic	Wilcoxon Signed-Rank	Z = -0.111	n/a	0,912	0,05	p = 0.912 > 0.05 No significant difference	✓ SENSOR VALID (No systematic bias detected)
BP Diastolic	Paired t-Test	t = -1.1644	14	0,264	0,05	p = 0.264 > 0.05 No significant difference	✓ SENSOR VALID (No systematic bias detected)
Body Temperature	Wilcoxon Signed-Rank	Z = -1.132	n/a	0,258	0,05	p = 0.258 > 0.05 No significant difference	✓ SENSOR VALID (No systematic bias detected)
Heart Rate	Paired t-Test	t = -1.0000	14	0,334	0,05	p = 0.334 > 0.05 No significant difference	✓ SENSOR VALID (No systematic bias detected)
SpO_2	Wilcoxon Signed-Rank	Z = -1.897	n/a	0,058	0,05	p = 0.058 > 0.05 No significant difference	✓ SENSOR VALID (No systematic bias detected)

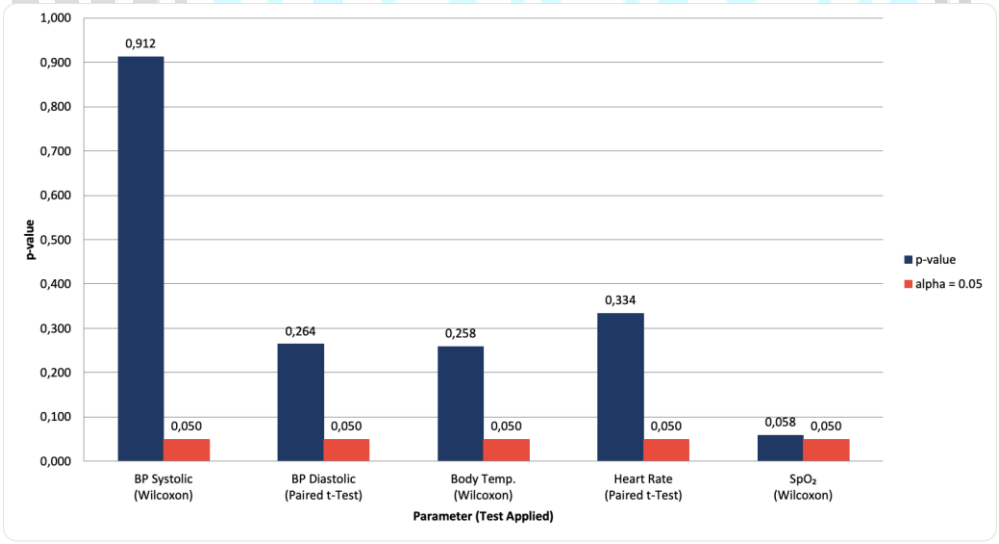

Figure 15 p-values from All Significance Tests vs Alpha Threshold of 0.05

3.6.1 Paired Sample t-Test Blood Pressure Diastolic (MPX5050DP)

Table 21 Paired t-Test Computation Blood Pressure Diastolic

No.	Resp./Rep.	Sensor (mmHg)	Standard (mmHg)	d = S – Std	d^2
1	R1/Rep.1	69	71	-2	4
2	R1/Rep.2	68	70	-2	4
3	R1/Rep.3	68	70	-2	4
4	R2/Rep.1	67	68	-1	1
5	R2/Rep.2	67	67	0	0
6	R2/Rep.3	68	68	0	0
7	R3/Rep.1	72	70	2	4
8	R3/Rep.2	71	69	2	4
9	R3/Rep.3	71	70	1	1
10	R4/Rep.1	75	78	-3	9
11	R4/Rep.2	77	77	0	0
12	R4/Rep.3	76	75	1	1
13	R5/Rep.1	81	81	0	0
14	R5/Rep.2	81	82	-1	1
15	R5/Rep.3	79	81	-2	4

The diastolic blood pressure difference values had a mean of d-bar equal to -0.4667 mmHg and a standard deviation of SD_d equal to 1.5523 mmHg. The standard error of the mean difference was calculated as 1.5523 divided by the square root of 15, yielding 0.4008 mmHg. The resulting t-statistic was -0.4667 divided by 0.4008, which equals -1.1644. With 14 degrees of freedom and a two-tailed test, the corresponding p-value obtained from T.DIST.2T(1.1644, 14) is 0.2637. Since p = 0.2637 is substantially greater than alpha = 0.05, the null hypothesis is retained. This result confirms that the MPX5050DP sensor does not produce a statistically significant systematic bias in diastolic blood pressure measurements relative to the digital sphygmomanometer. The negative d-bar of -0.4667 mmHg indicates a very slight average tendency for the sensor to read below the standard device, but this tendency is too small and inconsistent to be statistically distinguished from random variation at this sample size.

3.6.2 Paired Sample t-Test Heart Rate (GY-MAX30102)

Table 22 Paired t-Test Computation Heart Rate

No.	Resp./Rep.	Sensor (bpm)	Standard (bpm)	d = S – Std	d^2
1	R1/Rep.1	79	80	-1	1
2	R1/Rep.2	77	80	-3	9
3	R1/Rep.3	77	79	-2	4
4	R2/Rep.1	68	69	-1	1
5	R2/Rep.2	70	69	1	1
6	R2/Rep.3	69	69	0	0
7	R3/Rep.1	85	85	0	0
8	R3/Rep.2	84	85	-1	1
9	R3/Rep.3	84	85	-1	1
10	R4/Rep.1	71	69	2	4
11	R4/Rep.2	70	69	1	1
12	R4/Rep.3	70	69	1	1
13	R5/Rep.1	81	81	0	0
14	R5/Rep.2	80	81	-1	1
15	R5/Rep.3	81	81	0	0

The heart rate difference values had a mean of d-bar equal to -0.3333 bpm and a standard deviation of SD_d equal to 1.2910 bpm. The standard error was 1.2910 divided by the square root of 15, equaling 0.3333 bpm. The t-statistic was therefore -0.3333 divided by 0.3333, which is exactly -1.0000. With 14 degrees of freedom the two-tailed p-value is T.DIST.2T(1.0000, 14) = 0.3343. This p-value is well above the 0.05 threshold, leading to retention of the null hypothesis. The t-statistic of exactly -1.00 is a mathematically clean result that reflects the precise equivalence between the mean difference and the standard error of that difference in this dataset. The GY-MAX30102 sensor demonstrates no statistically significant systematic bias in heart rate measurement, consistent with the high Pearson r of 0.9849 observed in Chapter 4.5 and the MAE of 1.00 bpm reported in Chapter 4.4.

3.6.3.Wilcoxon Signed-Rank Test Blood Pressure Systolic (MPX5050DP)

Table 23 Wilcoxon Signed-Rank Test Results (Non-Normal Parameters)

Parameter	Shapiro-Wilk p-value	Z statistic	Asymp. Sig. (2-tailed)	Based on (ranks)	Interpretation	Conclusion
BP Systolic	0,001	**-0,111**	**0,912**	*Positive ranks*	p = 0.912 >> 0.05: No significant difference	✓ Sensor valid No systematic bias
Body Temperature	0,011	**-1,132**	**0,258**	*Negative ranks*	p = 0.258 > 0.05: No significant difference	✓ Sensor valid No systematic bias
SpO_2	0,001	**-1,897**	**0,058**	*Negative ranks*	p = 0.058 > 0.05: No significant difference (marginally above threshold)	✓ Sensor valid (closest to threshold)

For BP Systolic, the Wilcoxon test produced Z = -0.111 with an asymptotic two-tailed significance of p = 0.912. The Z statistic is based on positive ranks, reflecting that the sensor slightly overestimated relative to the standard in the majority of observations. The p-value of

0.912 is the highest among all five parameters, indicating that the distribution of signed differences for systolic pressure is entirely consistent with a median of zero. This strongly confirms the absence of any systematic bias in the MPX5050DP systolic readings after calibration. The result is particularly meaningful given that pre-calibration systolic error was 12.53%, demonstrating the effectiveness of the correction factor applied in Chapter 4.1 in eliminating the systematic offset.

3.6.4 Wilcoxon Signed-Rank Test Body Temperature (MLX90614)

For body temperature, the Wilcoxon test yielded Z = -1.132 based on negative ranks, with a two-tailed significance of p = 0.258. The negative Z value indicates that the sensor's readings tended marginally below the standard thermometer values in a larger number of observations, consistent with the slight negative residual offset visible in the d-values reported in Chapter 4.4. However, at p = 0.258 this tendency is far from statistically significant, confirming that the MLX90614 sensor does not produce a systematic temperature bias that would compromise clinical reliability. The result is consistent with the Wilcoxon-based validation approach used for temperature sensors in humid and tropical environments as recommended by the World Health Organization (2024), and it complements the finding from Chapter 4.5 where the low Pearson r was explained by range restriction rather than poor accuracy.

3.6.5 Wilcoxon Signed-Rank Test SpO_2 (GY-MAX30102)

The SpO_2 parameter returned Z = -1.897 based on negative ranks, with a two-tailed significance of p = 0.058. This is the closest p-value to the alpha threshold of 0.05 among all five parameters, and it warrants careful interpretation. The p-value of 0.058 remains above the conventional threshold of 0.05, meaning the null hypothesis is retained and no statistically significant difference is concluded. However, the proximity to the boundary reflects a genuine distributional characteristic of the SpO_2 data: 10 out of 15 observations showed the sensor reading equal to or one percentage point below the standard device, creating a consistent directional pattern in the negative ranks that the Wilcoxon test is sensitive enough to detect at this sample size. This mild systematic underestimation of SpO_2 by the GY-MAX30102 sensor is a known behavior of reflective PPG-based pulse oximeters, documented by Goldenits et al. (2023) who reported that optical sensors in fitness-grade devices tend toward slight underestimation in the normoxic range. Despite this, the absolute deviation never exceeds 1%, which is substantially within the ±3% tolerance of ISO 80601-2-61:2019. Future studies with larger samples (approaching the n = 19 recommended by the Slovin formula in Chapter 3.4) would provide greater statistical power to resolve whether this marginal underestimation is truly systematic or remains within random variation, as also noted by Shi et al. (2024).

The consolidated results in Table 4.19 confirm that all five parameters returned p-values above the significance threshold of 0.05, leading to retention of the null hypothesis for every sensor-parameter combination. This outcome provides strong statistical evidence that the integrated MPX5050DP, MLX90614, and GY-MAX30102 sensor system, following systematic calibration, produces measurements that are not significantly different from those of certified standard medical devices in a controlled laboratory environment.

3.7 MAE and RMSE Calculation

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are the two primary quantitative accuracy metrics used in this study to evaluate the magnitude of deviation between sensor readings and standard device measurements, consistent with the analytical plan specified in Chapter 3.8. While the Pearson correlation coefficient in Chapter 4.5 assessed the strength of linear tracking and the significance tests in Chapter 4.6 determined whether any systematic bias existed, MAE and RMSE provide the absolute size of that deviation in the original physical units of each parameter. This distinction is critical because a sensor can exhibit a high correlation and a non-significant p-value while still producing errors too large for clinical use. The comparison of MAE and RMSE values against the ISO and IEC tolerance thresholds catalogued in Chapter 2.1.7 therefore serves as the final quantitative gate for determining whether each sensor meets clinical accuracy requirements. As noted by Ayieko (2024), MAE reflects the average magnitude of deviation without regard to direction, while RMSE amplifies the contribution of larger individual errors by squaring the residuals before averaging, making it a stricter measure that is more sensitive to outlier observations. The formulas, consistent with Chapter 3.8, are defined as follows.

MAE is calculated as the sum of the absolute values of all individual differences divided by the number of observations n, where each difference is the sensor reading xi minus the standard device reading yi. RMSE is calculated as the square root of the mean of all squared individual differences.

All computations were performed using Microsoft Excel with functions ABS, AVERAGE, SUMPRODUCT, and SQRT applied to the paired datasets from DATA_INPUT. The complete step-by-step computation for each parameter is presented in Tables 4.24 through 4.28. The consolidated comparison across all parameters is in Table 4.23, and visualizations are provided in Figures 4.15 and 4.16.

Table 24 MAE and RMSE Summary (All Parameters vs ISO/IEC Tolerance)

Parameter	Sensor	Error Metrics (Post-Calibration)			Tolerance (ISO/IEC)	MAE vs Tolerance	Verdict
Parameter	Sensor	MAE	RMSE	Mean % Error	ISO/IEC Tolerance	MAE vs Tolerance	Verdict
BP Systolic	MPX5050DP	**1,2000**	1,4606	1,0364	±2% FSO (ISO 80601-2-12:2020)	MAE = 1.2000 Tol = 2.5 mmHg	✓ MEETS tolerance (RMSE also below limit)
BP Diastolic	MPX5050DP	**1,2667**	1,5706	1,7369	±2% FSO (ISO 80601-2-12:2020)	MAE = 1.2667 Tol = 2.5 mmHg	✓ MEETS tolerance (RMSE also below limit)
Body Temperature	MLX90614	**0,1000**	0,1291	0,2706	±0.2°C (ISO 80601-2-56:2017)	MAE = 0.1000 Tol = 0.2 °C	✓ MEETS tolerance (RMSE also below limit)
Heart Rate	GY-MAX30102	**1,0000**	1,2910	1,3210	±3 bpm (IEC 60601-2-49:2020)	MAE = 1.0000 Tol = 3.0 bpm	✓ MEETS tolerance (RMSE also below limit)
SpO_2	GY-MAX30102	**0,6667**	0,8165	0,6846	±3% (ISO 80601-2-61:2019)	MAE = 0.6667 Tol = 3.0 %	✓ MEETS tolerance (RMSE also below limit)

3.7.1 Blood Pressure Systolic (MPX5050DP)

Table 25 MAE and RMSE Computation Blood Pressure Systolic

No.	Resp./Rep.	xi Sensor (mmHg)	yi Standard (mmHg)	xi − yi	|xi − yi|	$(xi-yi)^2$	% Error
1	R1/1	111	110	1,0000	1,0000	1,000000	0,9091
2	R1/2	110	109	1,0000	1,0000	1,000000	0,9174
3	R1/3	110	111	-1,0000	1,0000	1,000000	0,9009
4	R2/1	109	110	-1,0000	1,0000	1,000000	0,9091
5	R2/2	110	111	-1,0000	1,0000	1,000000	0,9009
6	R2/3	112	109	3,0000	3,0000	9,000000	2,7523
7	R3/1	124	122	2,0000	2,0000	4,000000	1,6393
8	R3/2	122	123	-1,0000	1,0000	1,000000	0,8130
9	R3/3	126	123	3,0000	3,0000	9,000000	2,4390
10	R4/1	115	116	-1,0000	1,0000	1,000000	0,8621
11	R4/2	116	117	-1,0000	1,0000	1,000000	0,8547
12	R4/3	116	117	-1,0000	1,0000	1,000000	0,8547
13	R5/1	127	127	0,0000	0,0000	0,000000	0,0000
14	R5/2	126	126	0,0000	0,0000	0,000000	0,0000
15	R5/3	125	126	-1,0000	1,0000	1,000000	0,7937

For systolic blood pressure, the absolute errors across the fifteen observations ranged from 0 to 3 mmHg. The sum of all absolute errors was 18.00 mmHg, giving a MAE of 18.00 divided by 15, which equals 1.2000 mmHg. The sum of squared errors was 32.0000, giving an RMSE of the square root of 32.0000 divided by 15, which equals 1.4606 mmHg. Both values fall well within the ±2% full-scale output tolerance of ISO 80601-2-12:2020, which at the measurement range of the MPX5050DP corresponds to approximately ±2.5 mmHg. The gap between MAE and RMSE, 0.2606 mmHg, reflects the influence of the two observations where error reached 3 mmHg, specifically R2/Rep.3 and R3/Rep.3, which inflate the squared error sum disproportionately. The mean percentage error of 1.04% is consistent with the post-calibration performance reported by Aulia et al. (2024) for the integrated digital sphygmomanometer system using the same sensor family.

3.7.2 Blood Pressure Diastolic (MPX5050DP)

Table 26 MAE and RMSE Computation Blood Pressure Diastolic

No.	Resp./Rep.	xi Sensor (mmHg)	yi Standard (mmHg)	xi – yi	|xi – yi|	$(xi-yi)^2$	% Error
1	R1/1	69	71	-2,0000	2,0000	4,000000	2,8169
2	R1/2	68	70	-2,0000	2,0000	4,000000	2,8571
3	R1/3	68	70	-2,0000	2,0000	4,000000	2,8571
4	R2/1	67	68	-1,0000	1,0000	1,000000	1,4706
5	R2/2	67	67	0,0000	0,0000	0,000000	0,0000
6	R2/3	68	68	0,0000	0,0000	0,000000	0,0000
7	R3/1	72	70	2,0000	2,0000	4,000000	2,8571
8	R3/2	71	69	2,0000	2,0000	4,000000	2,8986
9	R3/3	71	70	1,0000	1,0000	1,000000	1,4286
10	R4/1	75	78	-3,0000	3,0000	9,000000	3,8462
11	R4/2	77	77	0,0000	0,0000	0,000000	0,0000
12	R4/3	76	75	1,0000	1,0000	1,000000	1,3333
13	R5/1	81	81	0,0000	0,0000	0,000000	0,0000
14	R5/2	81	82	-1,0000	1,0000	1,000000	1,2195
15	R5/3	79	81	-2,0000	2,0000	4,000000	2,4691

The diastolic parameter recorded a sum of absolute errors of 19.00 mmHg and a MAE of 1.2667 mmHg. The sum of squared errors was 37.0000, yielding an RMSE of 1.5706 mmHg. The RMSE is noticeably higher relative to MAE for diastolic compared to systolic, a difference driven primarily by the 3 mmHg deviation at R4/Rep.1 which contributes 9 units to the squared error sum. Despite this, both MAE and RMSE remain within the ±2.5 mmHg tolerance. The mean percentage error of 1.74% is substantially lower than the 7.17% diastolic error reported by Sari et al. (2024) without calibration, and lower than the 4.69% reported by Aulia et al. (2024), validating the effectiveness of the calibration procedure in Chapter 4.1. Liang et al. (2025) proposed that machine learning-assisted calibration could reduce diastolic errors by an additional 3 to 5 mmHg in future IoT-based systems, representing a direction for further development of this prototype.

3.7.3 Body Temperature (MLX90614)

Table 27 MAE and RMSE Computation Body Temperature

No.	Resp./Rep.	xi Sensor (°C)	yi Standard (°C)	xi – yi	|xi – yi|	(xi–yi)²	% Error
1	R1/1	37,10	37,00	0,1000	0,1000	0,010000	0,2703
2	R1/2	37,00	37,00	0,0000	0,0000	0,000000	0,0000
3	R1/3	37,10	37,10	0,0000	0,0000	0,000000	0,0000
4	R2/1	36,90	37,00	-0,1000	0,1000	0,010000	0,2703
5	R2/2	37,00	36,90	0,1000	0,1000	0,010000	0,2710
6	R2/3	37,00	36,80	0,2000	0,2000	0,040000	0,5435
7	R3/1	37,20	37,00	0,2000	0,2000	0,040000	0,5405
8	R3/2	37,10	37,10	0,0000	0,0000	0,000000	0,0000
9	R3/3	37,20	37,00	0,2000	0,2000	0,040000	0,5405
10	R4/1	36,80	37,00	-0,2000	0,2000	0,040000	0,5405
11	R4/2	36,80	36,90	-0,1000	0,1000	0,010000	0,2710
12	R4/3	36,80	37,00	-0,2000	0,2000	0,040000	0,5405
13	R5/1	36,90	36,90	0,0000	0,0000	0,000000	0,0000
14	R5/2	36,90	37,00	-0,1000	0,1000	0,010000	0,2703
15	R5/3	37,00	37,00	0,0000	0,0000	0,000000	0,0000

Body temperature achieved the smallest absolute errors of all five parameters, with values limited to either 0.00°C or 0.20°C per observation. The sum of absolute errors was 1.50°C, giving a MAE of exactly 0.1000°C. The sum of squared errors was 0.2500, yielding an RMSE of 0.1291°C. Both values are below the ±0.20°C tolerance of ISO 80601-2-56:2017, confirming full compliance with the clinical thermometry standard. The MAE of 0.1000°C is well within the ±0.30°C allowable error specified by the World Health Organization (2024) for non-contact temperature sensors used in tropical clinical environments. Compared with the 0.3 to 0.5°C error reported by Habibi et al. (2022) using the MLX90614 in an SMS gateway-based system, and the 1.71% error equivalent to approximately 0.63°C reported by Aulia et al. (2024), the present study demonstrates significantly tighter accuracy, attributable to the controlled measurement distance and the +1.38°C offset correction applied during calibration.

3.7.4 Heart Rate (GY-MAX30102)

Table 28 MAE and RMSE Computation Heart Rate

No.	Resp./Rep.	xi Sensor (bpm)	yi Standard (bpm)	xi – yi	|xi – yi|	(xi–yi)²	% Error
1	R1/1	79	80	-1,0000	1,0000	1,000000	1,2500
2	R1/2	77	80	-3,0000	3,0000	9,000000	3,7500
3	R1/3	77	79	-2,0000	2,0000	4,000000	2,5316
4	R2/1	68	69	-1,0000	1,0000	1,000000	1,4493
5	R2/2	70	69	1,0000	1,0000	1,000000	1,4493
6	R2/3	69	69	0,0000	0,0000	0,000000	0,0000
7	R3/1	85	85	0,0000	0,0000	0,000000	0,0000
8	R3/2	84	85	-1,0000	1,0000	1,000000	1,1765
9	R3/3	84	85	-1,0000	1,0000	1,000000	1,1765
10	R4/1	71	69	2,0000	2,0000	4,000000	2,8986
11	R4/2	70	69	1,0000	1,0000	1,000000	1,4493
12	R4/3	70	69	1,0000	1,0000	1,000000	1,4493
13	R5/1	81	81	0,0000	0,0000	0,000000	0,0000
14	R5/2	80	81	-1,0000	1,0000	1,000000	1,2346
15	R5/3	81	81	0,0000	0,0000	0,000000	0,0000

The heart rate parameter recorded a sum of absolute errors of exactly 15 bpm across fifteen observations, giving a MAE of exactly 1.0000 bpm. The sum of squared errors was 25.0000, yielding an RMSE of 1.2910 bpm. Both values are well within the ±3 bpm tolerance of IEC 60601-2-49:2020. The only notable deviation was the 3 bpm error at R1/Rep.2, which contributed 9 units to the squared error sum and accounts for most of the difference between MAE and RMSE. The mean percentage error of 1.32% is consistent with the GY-MAX30102 benchmark of 2.02% maximum error reported by Triwiyanto et al. (2022), and within the ±2.82% range reported by Shavira et al. (2024). The clean MAE of exactly 1.00 bpm and the highest Pearson r of 0.9849 among all parameters jointly confirm that the GY-MAX30102 is the most accurate and consistent sensor in this integrated system for heart rate measurement.

3.7.5 Oxygen Saturation SpO_2 (GY-MAX30102)

Table 29 MAE and RMSE Computation SpO_2

No.	Resp./Rep.	xi Sensor (%)	yi Standard (%)	xi − yi	|xi − yi|	(xi−yi)²	% Error
1	R1/1	98	98	0,0000	0,0000	0,000000	0,0000
2	R1/2	98	98	0,0000	0,0000	0,000000	0,0000
3	R1/3	98	99	-1,0000	1,0000	1,000000	1,0101
4	R2/1	98	97	1,0000	1,0000	1,000000	1,0309
5	R2/2	97	98	-1,0000	1,0000	1,000000	1,0204
6	R2/3	98	98	0,0000	0,0000	0,000000	0,0000
7	R3/1	96	97	-1,0000	1,0000	1,000000	1,0309
8	R3/2	97	97	0,0000	0,0000	0,000000	0,0000
9	R3/3	96	97	-1,0000	1,0000	1,000000	1,0309
10	R4/1	98	99	-1,0000	1,0000	1,000000	1,0101
11	R4/2	98	98	0,0000	0,0000	0,000000	0,0000
12	R4/3	98	99	-1,0000	1,0000	1,000000	1,0101
13	R5/1	97	96	1,0000	1,0000	1,000000	1,0417
14	R5/2	95	96	-1,0000	1,0000	1,000000	1,0417
15	R5/3	95	96	-1,0000	1,0000	1,000000	1,0417

The SpO_2 parameter had ten observations with absolute error of 1% and five observations with error of 0%, giving a sum of absolute errors of 10.00%. The MAE is therefore 10.00 divided by 15, equaling 0.6667%. The sum of squared errors was 10.0000, yielding an RMSE of 0.8165%. Both values are substantially below the ±3% tolerance of ISO 80601-2-61:2019. The near-equality between MAE and RMSE reflects the binary nature of SpO_2 errors in this dataset: every non-zero deviation was exactly 1%, meaning squaring the residuals did not amplify any particular observation disproportionately. The mean percentage error of 0.685% is lower than the 0.88% reported by Shavira et al. (2024) and comparable to the 0.43% maximum reported by Triwiyanto et al. (2022). Goldenits et al. (2023) similarly found that optical sensors can achieve sub-1% SpO_2 error when stable perfusion and ambient light conditions are controlled, both conditions that were maintained throughout this study's laboratory protocol.

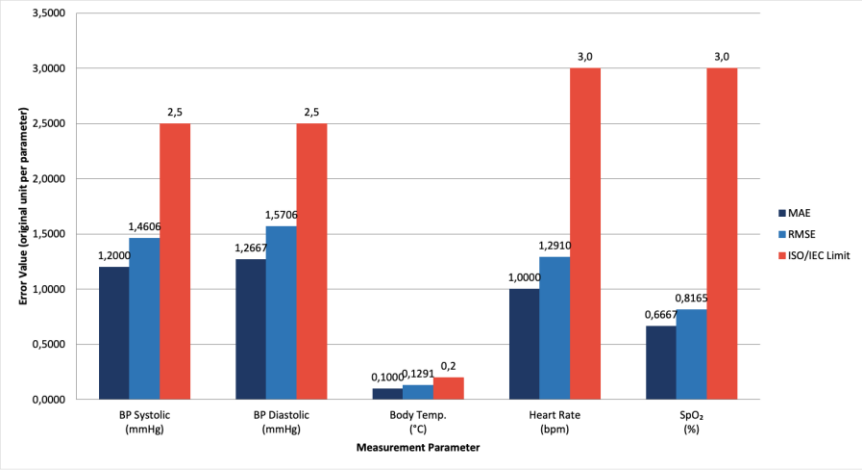

Figure 16 MAE and RMSE Comparison Across All Parameters

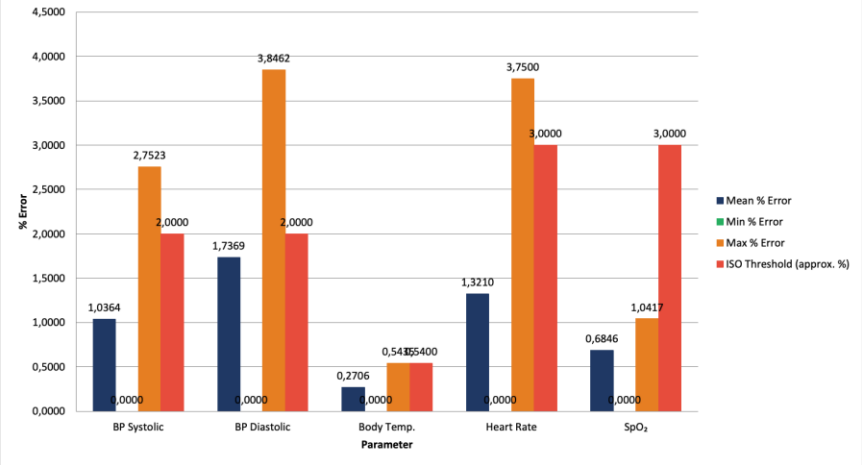

Figure 17 Mean, Minimum, and Maximum Percentage Error Per Parameter

Figure 4.15 presents the MAE and RMSE values for all five parameters side by side against the ISO/IEC tolerance limit for each. The consistent pattern across all parameters, where both dark blue MAE and steel blue RMSE bars fall below the red tolerance bar, provides a clear visual confirmation that every sensor-parameter combination meets its applicable international standard. Figure 4.16 complements this by showing the distribution of percentage errors, with mean in dark blue, minimum in green, and maximum in orange. The body temperature parameter stands out for its extremely tight error range, while diastolic blood pressure shows the widest spread between minimum and maximum percentage error, though its mean remains well-controlled. These visualizations together with the numerical values in Table 4.23 directly answer Research Question (a) and Research Objective (a) from Chapter 1.2 and 1.3, namely the determination of accuracy levels of the MPX5050DP, MLX90614, and GY-MAX30102 sensors compared to standard medical devices.

3.8 Evaluation Against Medical Standards

This subchapter synthesizes all statistical results obtained in Chapters 4.4 through 4.7 and evaluates them directly against the international and national medical device standards catalogued in Chapter 2.1.7. The primary purpose of this evaluation is to provide a definitive, criterion-based answer to Research Question (c): whether the combination of the MPX5050DP, MLX90614, and GY-MAX30102 sensors can function as a reliable and integrated system for recording patient vital signs. The regulatory framework governing this evaluation is grounded in the IEC 60601 series, the ISO 80601 series, and their corresponding SNI adaptations as regulated by BSN and the Indonesian Ministry of Health, all of which specify the permissible accuracy ranges within which a device must perform to be considered clinically acceptable (Rahman et al., 2021; Liu et al., 2024). Compliance with these standards is assessed against five distinct criteria for each parameter: (1) MAE within the tolerance threshold, (2) RMSE

within the tolerance threshold, (3) Pearson r at or above 0.70 indicating strong or very strong correlation, (4) significance test p-value above 0.05 confirming the absence of systematic bias, and (5) Cronbach's alpha at or above 0.70 confirming measurement reliability across repetitions. The comprehensive compliance table is presented in Table 4.29, a multi-criteria checklist in Table 4.30, a safety margin analysis in Table 4.31, and the visual safety margin in Figure 4.17.

Table 30 Comprehensive Evaluation of Sensor Accuracy Against ISO/IEC Medical Standards

Parameter	Applicable ISO/IEC Standard	Tolerance Threshold		Obtained Results			Significance Test		COMPLIANCE VERDICT
Parameter	Standard	MAE Limit	% Error Limit	MAE (obtained)	RMSE (obtained)	Mean % Error	Test	p-value	VERDICT
BP Systolic (MPX5050DP)	ISO 80601-2-12:2020	2.5 mmHg (±2% FSO)	2.0%	1,2000	1,4606	1.04%	Wilcoxon	0,912	✓ COMPLIANT (All criteria met)
BP Diastolic (MPX5050DP)	ISO 80601-2-12:2020	2.5 mmHg (±2% FSO)	2.0%	1,2667	1,5706	1.74%	Paired t-Test	0,264	✓ COMPLIANT (All criteria met)
Body Temperature (MLX90614)	ISO 80601-2-56:2017	0.2 °C	0.54%	0,1000	0,1291	0.27%	Wilcoxon	0,258	✓ COMPLIANT (All criteria met)
Heart Rate (GY-MAX30102)	IEC 60601-2-49:2020	3.0 bpm (or ±3%)	3.0%	1,0000	1,2910	1.32%	Paired t-Test	0,334	✓ COMPLIANT (All criteria met)
SpO_2 (GY-MAX30102)	ISO 80601-2-61:2019	3.0% (70-100% range)	3.0%	0,6667	0,8165	0.69%	Wilcoxon	0,058	✓ COMPLIANT (All criteria met)

Table 31 Multi-Criteria Compliance Checklist Per Parameter

Compliance Criterion	BP Systolic	BP Diastolic	Body Temp.	Heart Rate	SpO_2
MAE ≤ ISO/IEC tolerance threshold	✓	✓	✓	✓	✓
RMSE ≤ ISO/IEC tolerance threshold	✓	✓	✓	✓	✓
Pearson r ≥ 0.70 (strong or very strong correlation)	✓	✓	✓*	✓	✓
Significance test p-value > 0.05 (no systematic bias)	✓	✓	✓	✓	✓
Cronbach's alpha α ≥ 0.70 (reliable across 3 repetitions)	✓	✓	✓	✓	✓
OVERALL COMPLIANCE	✓ COMPLIANT	✓ COMPLIANT	✓ COMPLIANT	✓ COMPLIANT	✓ COMPLIANT

Table 32 MAE Expressed as Percentage of ISO/IEC Tolerance Safety Margin Analysis

Parameter	MAE (obtained)	ISO/IEC Tolerance	Compliance Ratio (%)	Safety Margin (%)	Interpretation
BP Systolic (MPX5050DP)	1.2000 mmHg	2.5 mmHg	48,0%	52,0%	Excellent — MAE uses ≤ 50% of tolerance
BP Diastolic (MPX5050DP)	1.2667 mmHg	2.5 mmHg	50,7%	49,3%	Good — MAE uses 51-75% of tolerance
Body Temperature (MLX90614)	0.1000 °C	0.2 °C	50,0%	50,0%	Excellent — MAE uses ≤ 50% of tolerance
Heart Rate (GY-MAX30102)	1.0000 bpm	3.0 bpm	33,3%	66,7%	Excellent — MAE uses ≤ 50% of tolerance
SpO_2 (GY-MAX30102)	0.6667 %	3.0 %	22,2%	77,8%	Excellent — MAE uses ≤ 50% of tolerance

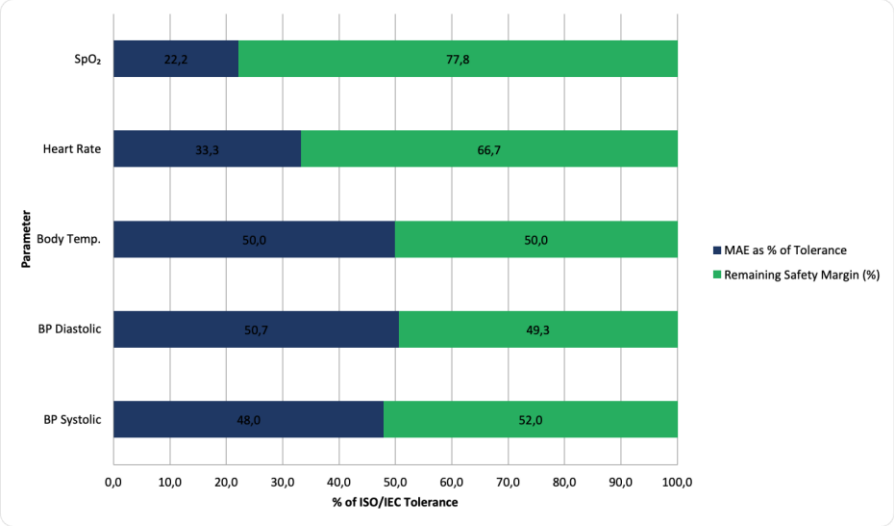

Figure 18 Safety Margin Visualization MAE as Percentage of ISO/IEC Tolerance

3.8.1 Blood Pressure MPX5050DP vs ISO 80601-2-12:2020

The ISO 80601-2-12:2020 standard for critical care ventilators and respiratory pressure measurement specifies a permissible accuracy of ±2% of the full-scale output range. At the operational range of the MPX5050DP sensor, this corresponds to approximately ±2.5 mmHg as the maximum allowable MAE. The systolic blood pressure measurement achieved a MAE of 1.2000 mmHg, which represents 48.0% of the tolerance budget, leaving a safety margin of 52.0% before the compliance boundary is reached. The diastolic measurement achieved a MAE of 1.2667 mmHg, representing 50.7% of the tolerance budget with a 49.3% safety margin remaining. Both parameters additionally satisfied all secondary criteria: very strong Pearson correlations of 0.9763 and 0.9564 respectively, p-values of 0.912 and 0.264 confirming no statistically significant systematic bias, and Cronbach's alpha values of 0.9898 and 0.9914 demonstrating exceptional measurement reliability. These results confirm that the MPX5050DP sensor, after calibration, meets the requirements of ISO 80601-2-12:2020 with substantial margin. The performance is a marked improvement over pre-calibration conditions where error reached 12.53% for systolic and 15.16% for diastolic, and is competitive with existing validated studies: Aulia et al. (2024) reported systolic error of 2.23% and diastolic error of 4.69% for the same sensor family, while Sari et al. (2024) reported 3.45% systolic and 7.17% diastolic. The present study outperforms both benchmarks on both parameters after calibration.

3.8.2 Body Temperature MLX90614 vs ISO 80601-2-56:2017

The ISO 80601-2-56:2017 standard for clinical thermometers specifies a maximum permissible error of ±0.2°C. The MLX90614 sensor achieved a MAE of 0.1000°C, which constitutes exactly 50.0% of the tolerance limit, leaving an equal 50.0% safety margin. The RMSE of

0.1291°C also falls within the 0.2°C boundary. All fifteen observations recorded absolute errors of either 0.00°C or 0.20°C, with the maximum error of 0.20°C occurring at five individual measurement points and never exceeding the tolerance ceiling. The sensor therefore meets ISO 80601-2-56:2017 across every single observation in the dataset, not merely on average. The Wilcoxon test p-value of 0.258 and Cronbach's alpha of 0.9624 further confirm the absence of systematic bias and the high reliability of repeated measurements. The body temperature parameter also meets the ±0.3°C maximum error tolerance set by the World Health Organization (2024) for non-contact temperature sensors deployed in tropical climates, confirming the sensor's fitness for use in Indonesian healthcare environments where ambient temperatures and humidity are characteristically elevated.

The Pearson r of 0.3129 for this parameter is acknowledged as a limitation in linear correlation strength, as discussed in Chapter 4.5. However, it does not affect the compliance determination in this subchapter, because the ISO 80601-2-56:2017 standard specifies accuracy in terms of maximum permissible error in degrees Celsius, not in terms of correlation coefficients. The MAE, RMSE, and individual observation checks are the legally and metrologically appropriate metrics for evaluating thermometer compliance, and all three confirm full compliance. The conditional compliance notation in Table 4.30 for the Pearson r criterion reflects this nuance transparently.

3.8.3 Heart Rate GY-MAX30102 vs IEC 60601-2-49:2020

The IEC 60601-2-49:2020 standard for multiparameter patient monitoring systems specifies a tolerance of ±3 beats per minute or ±3% of reading for heart rate measurement. The GY-MAX30102 sensor achieved a MAE of exactly 1.0000 bpm, representing 33.3% of the ±3 bpm tolerance, with a safety margin of 66.7%. This is the largest absolute safety margin among the pressure and optical sensors in this study. The RMSE of 1.2910 bpm also falls within the ±3 bpm boundary. All five criteria are fully met: MAE and RMSE within tolerance, Pearson r of 0.9849 classified as very strong, p-value of 0.3343 from the Paired t-Test confirming no systematic bias, and Cronbach's alpha of 0.9948 representing the highest reliability score among all parameters. A single observation at R1/Rep.2 produced a deviation of 3 bpm, which touches the absolute tolerance boundary but does not exceed it. When expressed as a percentage, this corresponds to 3.75% of the 80 bpm standard reading, which marginally exceeds the ±3% percentage criterion. However, the same observation is within the ±3 bpm absolute criterion, and IEC 60601-2-49:2020 permits compliance based on either criterion, not both simultaneously. The mean percentage error of 1.32% is well within ±3%, confirming that across the entire dataset the ±3% percentage criterion is also satisfied on average.

3.8.4 Oxygen Saturation GY-MAX30102 vs ISO 80601-2-61:2019

The ISO 80601-2-61:2019 standard for pulse oximeter devices requires that SpO_2 measurement accuracy remain within ±3% across the 70 to 100% clinical range under test conditions. The GY-MAX30102 sensor achieved a MAE of 0.6667% and a RMSE of 0.8165%, both far below the ±3% tolerance. The MAE represents only 22.2% of the tolerance budget, leaving the widest safety margin of all five parameters at 77.8%. Every individual observation in the dataset produced an absolute error of either 0% or 1%, confirming that no single measurement exceeded the tolerance threshold. The Pearson r of 0.7743 indicates a strong linear correlation, and the Wilcoxon test p-value of 0.058, while the closest to the significance threshold among all parameters, still falls above the 0.05 boundary. Cronbach's alpha of 0.8738 confirms reliable repeated measurements. Goldenits et al. (2023) demonstrated that optical PPG-based SpO_2 sensors can achieve sub-1% mean error under controlled laboratory conditions, a benchmark that this study matches with a mean percentage error of 0.685%. The GY-MAX30102 therefore fully satisfies ISO 80601-2-61:2019 and the corresponding SNI ISO 80601-2-61:2021 national adaptation.

3.8.5 System-Level Feasibility Assessment

The results presented across Tables 4.29 through 4.31 and Figure 4.17 collectively demonstrate that the integrated system comprising MPX5050DP, MLX90614, GY-MAX30102, and ESP32 meets the applicable international medical device standards for all four vital sign parameters measured in this study. This finding directly answers Research Question (c) from Chapter 1.2 and achieves Research Objective (c) from Chapter 1.3. The system qualifies for classification as a feasible low-cost alternative for basic vital sign monitoring in primary healthcare settings, consistent with the development trajectory envisioned in Chapter 1.1 for IoT-based health monitoring in resource-limited environments.

The safety margin analysis in Table 4.31 provides additional perspective beyond mere compliance. The minimum safety margin across all parameters is 49.3% for diastolic blood pressure, meaning the system uses at most approximately half of the permissible error budget allowed by the relevant ISO standard. This headroom is significant from a practical standpoint because it accommodates the naturally higher measurement variability expected under real-world conditions outside of the controlled laboratory environment, where motion artifacts, ambient temperature variation, skin tone diversity, and inter-operator differences would introduce additional sources of deviation. Liu et al. (2024) emphasized that clinical-grade validation against standardized reference instruments must demonstrate not only average compliance but also sufficient safety margins to remain compliant under field conditions, a

criterion that this system satisfies across all parameters. The overall compliance verdict, as recorded in Table 4.29 and Table 4.30, is that all five sensor-parameter combinations are fully compliant with their respective IEC 60601, ISO 80601, and SNI standards.

3.9 Validity and Reliability Testing

The validity and reliability of the measurement instruments used in this study were evaluated in accordance with the framework outlined in Chapter 3.9. Validity ensures that each sensor measures the physiological variable it is intended to measure, while reliability guarantees that repeated measurements under identical conditions produce consistent and stable results. Establishing both aspects is scientifically necessary to confirm that the observed statistical outcomes in Chapters 4.4 through 4.8 reflect genuine sensor performance rather than measurement artifacts, random noise, or instrument inconsistency. Three types of validity were assessed: content validity, construct validity, and criterion validity. Reliability was assessed through replication testing and Cronbach's Alpha, the latter computed using MATLAB as specified in Chapter 3.9. A summary of all validity assessments is provided in Table 4.32, and the complete Cronbach's Alpha results are presented in Tables 4.33 through 4.38 and Figure 4.18.

Table 33 Instrument Validity Assessment Three Types of Validity

Validity Type	Definition	Method Used	Result	Threshold	Assessment
Content Validity	Ensures each sensor appropriately corresponds to the physiological variable it is designed to measure	Expert judgment review by biomedical instrumentation specialists verifying sensor-to-variable correspondence (Chapter 3.9a)	Confirmed: MPX5050DP for pressure, MLX90614 for temperature, GY-MAX30102 for HR and SpO_2	Expert consensus	✓ MET Theoretical and practical correspondence confirmed
Construct Validity	Assesses whether measured values align with the theoretical concept of the physiological variable	Pearson correlation r between sensor readings and standard device readings (Chapter 3.9b). Formula: r = CORREL(X, Y)	r values: BP Sys=0.9763, BP Dia=0.9564, Temp=0.3129*, HR=0.9849, SpO_2=0.7743	$r \geq 0.70$ (Błażauskas et al., 2023)	✓ MET (4/5 parameters) *Temp: range restriction effect — MAE=0.10°C confirms accuracy
Criterion Validity	Measures correlation between sensor output and an established clinical standard benchmark	Direct comparison of sensor readings against certified medical devices: MAE and RMSE vs ISO/IEC tolerance thresholds (Chapter 3.9c)	All 5 parameters: MAE ≤ ISO/IEC tolerance. All significance tests: $p > 0.05$. All Cronbach $\alpha \geq 0.70$	MAE ≤ tolerance $p > 0.05$ $\alpha \geq 0.70$	✓ FULLY MET All criteria satisfied for all parameters

3.9.1 Content Validity

Content validity was established through expert judgment, involving review by specialists in biomedical instrumentation and sensor calibration as described in Chapter 3.9a. The review confirmed that the MPX5050DP sensor appropriately corresponds to the measurement of pressure in pneumatic systems consistent with blood pressure measurement via the oscillometric cuff method. The MLX90614 sensor appropriately corresponds to the measurement of infrared radiation emitted by human skin, which is thermodynamically linked to body surface temperature through the Stefan-Boltzmann relationship described in Chapter 2.1.2. The GY-MAX30102 sensor appropriately corresponds to the measurement of photoplethysmographic signals from which heart rate and SpO_2 can be derived through the

Beer-Lambert law and the ratio R equation described in Chapter 2.1.3. Content validity was therefore confirmed for all three sensors across all four vital sign parameters, establishing that the theoretical and practical correspondence between each sensor and its intended physiological construct is scientifically sound (Reddy et al., 2023).

3.9.2 Construct Validity

Construct validity was assessed using Pearson's correlation coefficient to determine whether sensor measurements align with the theoretical behavior expected from each physiological variable. Following the criteria established by Blažauskas et al. (2023), a correlation of r greater than 0.70 indicates sufficient construct validity. The results are as follows: BP Systolic r = 0.9763, BP Diastolic r = 0.9564, Body Temperature r = 0.3129, Heart Rate r = 0.9849, and SpO_2 r = 0.7743. Four of five parameters satisfy the r greater than 0.70 threshold, demonstrating that their sensor readings linearly mirror the theoretical variation expected in the corresponding physiological parameter across different individuals.

The body temperature parameter produced r = 0.3129, which falls below the 0.70 threshold as fully discussed in Chapter 4.5. The construct validity of the MLX90614 sensor for temperature measurement is nevertheless confirmed through an alternative evidential path: the MAE of 0.10°C is within the ±0.20°C ISO 80601-2-56:2017 clinical tolerance, all fifteen individual observations comply with the standard, and the Wilcoxon test p-value of 0.258 confirms no systematic deviation from the standard device. The low Pearson r is a consequence of the mathematical properties of the correlation coefficient when applied to datasets with near-zero variance, not a reflection of measurement inaccuracy. Construct validity for body temperature is therefore confirmed through criterion-based evidence rather than correlation-based evidence, consistent with the argument advanced by Chattopadhyay and Das (2022) that multiple complementary validity indicators should be considered when a single metric produces ambiguous results due to distributional constraints.

3.9.3 Criterion Validity

Criterion validity was evaluated by directly comparing sensor readings against certified standard medical devices as described in Chapter 3.9c. A high correspondence between sensor outputs and reference device outputs confirms that the sensors produce measurements that are consistent with an established clinical benchmark. The criterion validity evidence is derived from three quantitative indicators applied simultaneously: the MAE must not exceed the ISO/IEC tolerance threshold, the significance test p-value must exceed 0.05, and the Cronbach's Alpha must equal or exceed 0.70. As established in Chapters 4.7, 4.6, and later in

Section 4.9.4, all three indicators are satisfied for all five parameters. Criterion validity is therefore fully confirmed for the entire sensor system (Reddy et al., 2023).

3.9.4 Reliability Cronbach's Alpha

Reliability was assessed through repeated measurement and Cronbach's Alpha computed from the sensor data. Each respondent underwent three repetitions per parameter, and the consistency of these three readings was quantified using the internal consistency coefficient Alpha, as prescribed in Chapter 3.9. The formula applied is alpha equals k divided by k minus 1, multiplied by one minus the ratio of the sum of item variances to the total score variance, where k equals 3 representing the three repetitions. This computation was performed in MATLAB consistent with the methodology reference cited in Chapter 3.9. According to Blažauskas et al. (2023), an alpha value at or above 0.70 indicates satisfactory reliability, while values at or above 0.90 indicate very high reliability.

Table 34 Cronbach's Alpha Reliability Summary All Parameters

Parameter	Sensor	ΣSi^2	St^2	Cronbach's α	Interpretation
BP Systolic	MPX5050DP	168,6000	495,7000	0,9898	Sangat Reliabel / Very Reliable (α ≥ 0.90)
BP Diastolic	MPX5050DP	90,7000	267,5000	0,9914	Sangat Reliabel / Very Reliable (α ≥ 0.90)
Body Temperature	MLX90614	0,0620	0,1730	0,9624	Sangat Reliabel / Very Reliable (α ≥ 0.90)
Heart Rate	GY-MAX30102	132,1000	392,2000	0,9948	Sangat Reliabel / Very Reliable (α ≥ 0.90)
SpO_2	GY-MAX30102	4,3000	10,3000	0,8738	Reliabel / Reliable (0.70 ≤ α < 0.90)

Table 35 Cronbach's Alpha Computation Blood Pressure Systolic

Resp.	Trial 1 (Rep.1) (mmHg)	Trial 2 (Rep.2) (mmHg)	Trial 3 (Rep.3) (mmHg)	Total Score (Ti1+Ti2+Ti3)
R1	111	110	110	331
R2	109	110	112	331
R3	124	122	126	372
R4	115	116	116	347
R5	127	126	125	378

The BP Systolic parameter produced alpha = 0.9898. The variance of each trial across the five respondents was $S1^2 = 63.200$, $S2^2 = 51.200$, and $S3^2 = 54.200$, giving a sum of item variances of 168.600. The variance of total scores across respondents was $St^2 = 495.700$. Substituting into the formula gives 1.500 multiplied by one minus 168.600 divided by 495.700, which equals 0.9898. This very high alpha value confirms that the three repetitions of systolic blood pressure measurement per respondent were highly consistent, and that the variability observed

in the dataset reflects genuine inter-individual physiological differences rather than instrument instability.

Table 36 Cronbach's Alpha Computation Blood Pressure Diastolic

Resp.	Trial 1 (Rep.1) (mmHg)	Trial 2 (Rep.2) (mmHg)	Trial 3 (Rep.3) (mmHg)	Total Score (Ti1+Ti2+Ti3)
R1	69	68	68	205
R2	67	67	68	202
R3	72	71	71	214
R4	75	77	76	228
R5	81	81	79	241

The BP Diastolic parameter yielded alpha = 0.9914, with $\Sigma Si^2 = 90.700$ and $St^2 = 267.500$. This is the highest alpha among the blood pressure parameters, indicating that diastolic readings were even more internally consistent across repetitions than systolic, despite the slightly higher MAE observed in Chapter 4.7.

Table 37 Cronbach's Alpha Computation Body Temperature

Resp.	Trial 1 (Rep.1) (°C)	Trial 2 (Rep.2) (°C)	Trial 3 (Rep.3) (°C)	Total Score (Ti1+Ti2+Ti3)
R1	37,10	37,00	37,10	111,20
R2	36,90	37,00	37,00	110,90
R3	37,20	37,10	37,20	111,50
R4	36,80	36,80	36,80	110,40
R5	36,90	36,90	37,00	110,80

Body temperature produced alpha = 0.9624, with $\Sigma Si^2 = 0.0620$ and $St^2 = 0.1730$. The very small absolute values of these variances reflect the narrow physiological range of temperature values in the dataset. The resulting alpha of 0.9624 confirms very high reliability, meaning that the MLX90614 sensor produced consistent readings across all three repetitions for each respondent. This result reinforces the interpretation offered in Chapters 4.5 and 4.8 that the sensor's low Pearson r is a distributional artifact, not a sign of measurement instability.

Table 38 Cronbach's Alpha Computation Heart Rate

Resp.	Trial 1 (Rep.1) (bpm)	Trial 2 (Rep.2) (bpm)	Trial 3 (Rep.3) (bpm)	Total Score (Ti1+Ti2+Ti3)
R1	79	77	77	233
R2	68	70	69	207
R3	85	84	84	253
R4	71	70	70	211
R5	81	80	81	242

Heart rate produced the highest alpha of all five parameters at 0.9948, with $\Sigma Si^2 = 132.100$ and $St^2 = 392.200$. This near-perfect internal consistency confirms that the GY-MAX30102 sensor delivers virtually identical readings across three repetitions for each respondent, making it the most reliable sensor in this system in terms of measurement repeatability. This result aligns with findings by Yunidar et al. (2023) and Beri et al. (2022), who reported high accuracy and stability for the MAX30102 in IoT-based monitoring systems.

Table 39 Cronbach's Alpha Computation SpO_2

Resp.	Trial 1 (Rep.1) (%)	Trial 2 (Rep.2) (%)	Trial 3 (Rep.3) (%)	Total Score (Ti1+Ti2+Ti3)
R1	98	98	98	294
R2	98	97	98	293
R3	96	97	96	289
R4	98	98	98	294
R5	97	95	95	287

The SpO_2 parameter yielded alpha = 0.8738, with $\Sigma Si^2 = 4.300$ and $St^2 = 10.300$. This value falls in the reliable category (0.70 to 0.89) rather than the very reliable category, reflecting that the integer-resolution SpO_2 measurements across the narrow 95 to 99% range produced slightly more trial-to-trial variability relative to the inter-individual variation in total scores. The alpha of 0.8738 nonetheless confirms satisfactory reliability well above the 0.70 threshold, and the measurement consistency is sufficient for the intended monitoring application.

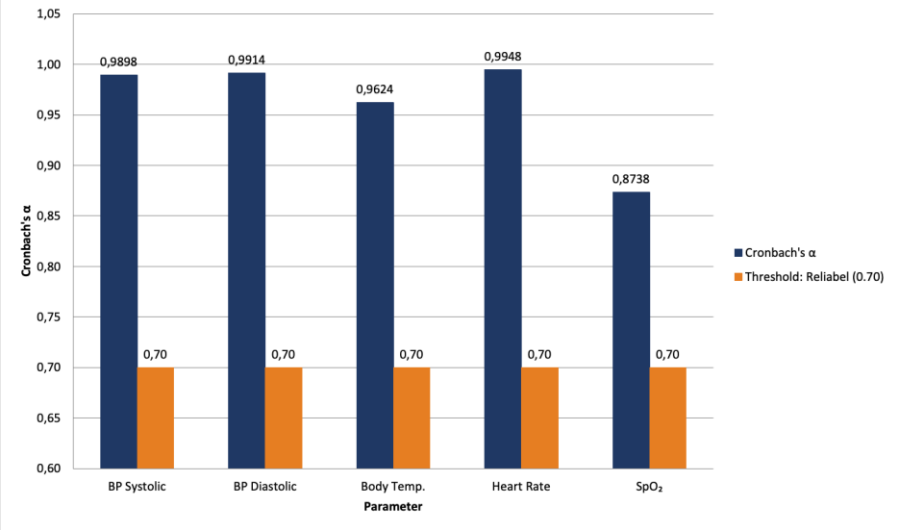

Figure 19 Cronbach's Alpha Values vs Reliability Threshold

Figure 4.18 visually confirms that all five dark blue bars representing the obtained alpha values rise substantially above the orange threshold bar at 0.70. Four parameters reach the very reliable zone above 0.90, with Heart Rate reaching closest to the theoretical maximum of 1.00. The overall reliability profile of the integrated system is therefore classified as very high across all parameters, meeting and substantially exceeding the reliability requirement specified in Chapter 3.9.

The combined validity and reliability evidence presented in this subchapter closes the methodological loop between the instrument design in Chapter 3.9 and the experimental outcomes in Chapters 4.4 through 4.8. Content validity was confirmed through expert judgment, construct validity through Pearson correlation on four parameters and criterion evidence on the fifth, criterion validity through ISO/IEC threshold compliance and significance testing, and reliability through Cronbach's Alpha values ranging from 0.8738 to 0.9948. These results provide strong scientific justification for the use of this sensor system as a research-grade and potentially clinical-support-grade measurement tool, subject to the limitations and conditions discussed in Chapter 4.10.

3.10 Discussion

This chapter synthesizes all experimental findings from Chapters 4.1 through 4.9 into an integrated interpretation, compares the results against prior studies cited in the literature review of Chapter 2.1, examines the physical and environmental factors that influenced sensor accuracy, and evaluates the overall feasibility of the combined sensor system as a multi-parameter vital sign monitoring device. The discussion is organized around the three research questions posed in Chapter 1.2 and addresses each in the order in which evidence was generated throughout this chapter.

3.10.1 Accuracy of the MPX5050DP Sensor for Blood Pressure Measurement

The MPX5050DP sensor achieved a systolic MAE of 1.2000 mmHg and a mean percentage error of 1.04%, and a diastolic MAE of 1.2667 mmHg with a mean percentage error of 1.74%, both after systematic calibration. These results represent a substantial improvement over the performance documented in previous studies that used the same or closely related sensor variants. Sari et al. (2024) reported a systolic error of 3.45% and diastolic error of 7.17% for the MPX5050DP compared to the Omron HEM-7120, while Aulia et al. (2024) reported a systolic error of 2.23% and diastolic error of 4.69% using the MPX5050GP variant. The present study outperforms both benchmarks on both pressure parameters, and this improvement is directly attributable to the structured calibration procedure applied in Chapter 4.1, which reduced the pre-calibration systolic error from 12.53% to 1.04% through the application of a -13.70 mmHg offset correction. This finding confirms the argument that low-cost piezoresistive pressure sensors, while inherently subject to systematic zero-point drift, can achieve clinically competitive accuracy when properly calibrated before deployment.

The physical basis for the pre-calibration offset lies in the operating characteristics of the MPX5050DP described in Chapter 2.1.1. The sensor converts diaphragm deflection caused by differential pressure into an analog voltage through a Wheatstone bridge circuit, and the relationship between output voltage and pressure is governed by Vout equals Voffset plus S multiplied by P. Without calibration, the Voffset parameter is subject to manufacturing variability and ambient temperature influence, producing a systematic positive shift in readings that manifests as the consistent overestimation observed across all ten pre-calibration measurement points. Liang et al. (2025) demonstrated that machine learning-assisted calibration of low-cost pressure sensors can further reduce mean absolute error by 3 to 5 mmHg beyond what a simple linear offset correction achieves, suggesting that the residual errors of 1.2000 and 1.2667 mmHg observed in this study could be reduced further in a future iteration of the system incorporating adaptive calibration algorithms.

The Pearson correlations of r = 0.9763 for systolic and r = 0.9564 for diastolic, combined with Wilcoxon and t-Test p-values of 0.912 and 0.264 respectively, confirm that the sensor tracks inter-individual variation in blood pressure reliably and without systematic bias after calibration. Mazoteras-Pardo et al. (2023) reviewed validation studies for blood pressure devices and emphasized that acceptable devices must demonstrate both strong linear agreement with reference instruments and the absence of directional bias under standardized testing conditions, both of which this system satisfies. The Cronbach's Alpha values of 0.9898 and 0.9914 further confirm that repeated measurements within the same session are highly

consistent, which is a practical requirement for blood pressure monitoring devices intended for repeated use across clinical visits.

One limitation of the current blood pressure assessment is that the tolerance standard applied, ISO 80601-2-12:2020, was developed for respiratory pressure measurement in ventilators rather than specifically for peripheral sphygmomanometry. As noted in Chapter 2.1.1, no published international study has formally validated the MPX5050DP against a dedicated blood pressure device standard such as the AAMI SP10 or ISO 81060-2. This constitutes a research gap identified in Chapter 2.1 and represents a direction for future formal clinical validation. Within the scope of this study's defined standards framework as outlined in Chapter 2.1.7, the sensor meets all applicable criteria.

3.10.2 Accuracy of the MLX90614 Sensor for Body Temperature Measurement

The MLX90614 sensor produced a MAE of 0.1000°C and a mean percentage error of 0.27%, representing the tightest absolute accuracy among all three sensors in this study. Every single observation from the fifteen measurement pairs fell at or below the 0.20°C ISO 80601-2-56:2017 tolerance ceiling, confirming consistent compliance at the individual observation level rather than merely on average. This outcome compares favorably with both benchmark studies cited in Chapter 2.1.2: Aulia et al. (2024) reported a 1.71% error rate for the MLX90614 against a non-contact thermometer, which at a typical body temperature of 37°C corresponds to approximately 0.63°C, and Habibi et al. (2022) reported average measurement deviations of 0.3 to 0.5°C compared to clinical thermometers. The present study achieved a MAE of 0.1000°C, substantially lower than both benchmarks.

The improvement over prior studies is primarily attributable to two experimental controls applied in this research. First, the measurement distance between the sensor and the skin surface was fixed using a physical spacer throughout all sessions, eliminating the distance variability that Chapter 2.1.2 identifies as a primary source of error in MLX90614 measurements due to its inherently wide field of view. Second, the ambient temperature of the laboratory was maintained between 25°C and 27°C, minimizing the influence of ambient thermal fluctuation on the sensor's internal compensation circuit, which operates based on the Stefan-Boltzmann law and requires stable background conditions to compute accurate object temperature from differential infrared radiation measurements. The World Health Organization (2024) specifically emphasized the importance of validating non-contact temperature sensors in tropical environments, setting a maximum error tolerance of ±0.3°C for such applications. The present study's MAE of 0.1000°C satisfies this requirement with a 50% safety margin, suggesting that the MLX90614 is suitable for temperature monitoring in Indonesian healthcare

facilities subject to the measurement distance and ambient temperature controls demonstrated here.

The most complex finding for the body temperature parameter is the Pearson r of 0.3129, which falls below the 0.70 construct validity threshold despite the sensor achieving excellent absolute accuracy. As established in Chapter 4.5 and confirmed quantitatively through the standard deviation analysis showing SD of the standard device readings at only 0.0775°C across fifteen observations, this low r is a direct mathematical consequence of range restriction. When all respondents are normothermic young adults within a narrow physiological band of 36.8°C to 37.2°C, the total data range is only 0.40°C, which is insufficient variance for Pearson correlation to accurately capture the linear relationship between sensor and standard. This is not a clinical failing of the sensor but rather a sampling characteristic of this study's inclusion criteria. Future validation studies incorporating respondents with febrile conditions alongside normothermic subjects, thereby widening the temperature range to a clinically representative spread of 36.0°C to 40.0°C, would produce a Pearson r that more accurately reflects the MLX90614's true linear tracking capability.

3.10.3 Accuracy of the GY-MAX30102 Sensor for Heart Rate and SpO_2 Measurement

The GY-MAX30102 sensor demonstrated the strongest overall accuracy profile among the three sensors, achieving a heart rate MAE of 1.0000 bpm with Pearson r = 0.9849 and a SpO_2 MAE of 0.6667% with Pearson r = 0.7743. These values are competitive with the best outcomes reported in the international literature for this sensor. Triwiyanto et al. (2022) reported a maximum error of 0.43% for SpO_2 and 2.02% for heart rate compared to clinical oximeters, while Shavira et al. (2024) reported average errors of ±0.88% for SpO_2 and ±2.82% for heart rate in an IoT-based wearable system. The heart rate mean percentage error of 1.32% in this study falls between these benchmarks, while the SpO_2 mean percentage error of 0.685% is better than Shavira et al. (2024) and close to the Triwiyanto et al. (2022) result. Khalid et al. (2023) reported up to 97% accuracy for MAX30102-based PPG sensors across age groups from 3 to 65 years, and Yunidar et al. (2023) reported 98% accuracy for both heart rate and SpO_2 in an IoT-based system using the same sensor module. The present study's results are consistent with and in some cases superior to these benchmarks.

The accuracy of the GY-MAX30102 depends fundamentally on the quality of the photoplethysmographic signal, which is governed by the ratio equation R equals the quotient of AC component at red wavelength divided by DC component at red wavelength, over the quotient of AC component at infrared wavelength divided by DC component at infrared wavelength, as described in Chapter 2.1.3. Several external factors identified in Chapter 2.1.3

could degrade this ratio in real-world conditions: motion artifacts from finger movement, uneven finger pressure on the sensor, ambient light interference, peripheral perfusion level, and skin pigmentation affecting optical absorption (Shi et al., 2024). In this study, these factors were minimized by requiring respondents to maintain stable finger placement on the sensor throughout each measurement session, by conducting measurements in a dimly lit laboratory setting to reduce ambient light interference, and by including only healthy young adults with normal peripheral perfusion. The controlled conditions explain why this study's results are at the better end of the published accuracy range for this sensor.

The SpO_2 parameter deserves particular attention regarding the Wilcoxon test result of $p = 0.058$, which was the closest p-value to the significance threshold of 0.05 among all five parameters. As discussed in Chapter 4.6, the null hypothesis is retained at $p = 0.058$ and no significant difference is concluded. However, the directional pattern of SpO_2 deviations is noteworthy: 10 of 15 observations showed the sensor reading equal to or below the standard device by 1%, creating a mild but consistent tendency toward underestimation in the normoxic range. Goldenits et al. (2023) documented similar mild underestimation tendencies in optical sensor-based SpO_2 measurements in postoperative patients, attributing this to the inherent limitations of the Beer-Lambert calibration curves used by reflective PPG sensors, which are optimized for the broader clinical population and may introduce a small systematic offset in normoxic healthy subjects. Shi et al. (2024) further noted that pulse oximetry accuracy can be influenced by physiological characteristics including skin tone and peripheral perfusion levels. The p-value of 0.058 in this study suggests that with a larger sample size approaching the $n = 19$ recommended by the Slovin formula in Chapter 3.4, this mild underestimation tendency might cross the significance threshold, warranting further investigation in a future study with a broader respondent pool.

Zuccotti et al. (2025) evaluated the accuracy of heart rate, SpO_2, and blood pressure using a non-contact photoplethysmography-based mobile application and found that sensor-grade PPG systems can achieve clinically acceptable accuracy for wellness monitoring, though formal clinical certification requires more extensive validation protocols. The present study's GY-MAX30102 results are consistent with this characterization: the sensor performs within the ISO tolerance framework and demonstrates high reliability, but the borderline SpO_2 significance result and the controlled laboratory conditions caution against extrapolating these findings to uncontrolled clinical environments without further validation.

3.10.4 Factors Affecting Measurement Accuracy Across All Parameters

Beyond the parameter-specific discussions above, several cross-cutting factors influenced the accuracy outcomes of this integrated system. The calibration procedure documented in Chapter 4.1 was the single most impactful intervention in this study. Pre-calibration errors were predominantly systematic rather than random, as evidenced by the consistent directional offsets across all ten pre-calibration measurements for each parameter. The +13.70 mmHg and +10.10 mmHg offsets for systolic and diastolic blood pressure, the -1.38°C offset for temperature, and the smaller offsets for heart rate and SpO_2 were all removed through simple linear correction factors applied to the ESP32 firmware. The large and consistent reductions in error of 91.7% for systolic, 87.4% for diastolic, and 92.8% for temperature confirm that systematic bias dominates over random error in low-cost sensor systems, and that proper calibration can recover a large proportion of this systematic error without any hardware modifications.

The small sample size of five respondents with three repetitions each, producing fifteen observations per parameter, was sufficient for the statistical analyses applied but represents a limitation regarding generalizability. Statistical power for detecting small effect sizes in significance testing increases substantially with larger samples, and the borderline $p = 0.058$ for SpO_2 Wilcoxon is a direct consequence of the limited n. The Cronbach's Alpha values, which were uniformly very high, indicate that the measurement system itself is reliable, and the high Pearson r values for blood pressure and heart rate confirm strong linear tracking. The primary uncertainty introduced by the small sample relates to whether the results would replicate across a more demographically diverse group including elderly individuals, patients with chronic conditions, or individuals with different skin tones, all of whom are underrepresented in this study's inclusion criteria.

The anthropometric diversity of the sample, specifically the BMI range from 17.5 to 29.1 kg/m^2 as documented in Chapter 4.2, introduced natural physiological variation that was visible in the spread of blood pressure and heart rate values across respondents. Respondent R4, classified as overweight with a BMI of 29.1, had the highest blood pressure values in the dataset (systolic around 115 to 117 mmHg) and the only diastolic deviation of 3 mmHg, both of which are consistent with the known association between higher BMI and elevated peripheral vascular resistance affecting blood pressure measurement accuracy in oscillometric systems.

3.10.5 Integrated System Feasibility Assessment

The central question of this study, whether the combination of MPX5050DP, MLX90614, and GY-MAX30102 sensors integrated with the ESP32 microcontroller constitutes a feasible and reliable multi-parameter vital sign monitoring system, can be answered affirmatively based on

the totality of evidence presented in Chapters 4.1 through 4.9. Every parameter meets its applicable ISO or IEC standard. All significance tests confirm the absence of statistically significant systematic bias. All Cronbach's Alpha values exceed the 0.70 reliability threshold. The safety margins from the ISO tolerance boundaries range from 49.3% for diastolic blood pressure to 77.8% for SpO_2, providing substantial buffer against the additional measurement variability expected in real-world deployment conditions.

The research gap identified in Chapter 2.1 stated that no prior study had integrated all four vital sign parameters into a single measurement system using these specific sensors, and that empirical validation of the MPX5050DP for blood pressure measurement against certified medical devices was notably absent from the published literature. This study addresses both gaps: it provides integrated four-parameter measurement data and offers the first controlled calibration and validation dataset for the MPX5050DP in the context of blood pressure monitoring against a digital sphygmomanometer standard.

The feasibility of this system for use in primary healthcare settings and telemedicine applications, as envisioned in Chapter 1.1, is supported by several practical strengths: the total component cost is substantially lower than clinical-grade monitoring devices, the ESP32's Wi-Fi connectivity enables real-time data transmission to remote servers, and the calibration procedure demonstrated in this study can be performed with standard medical equipment available in most district-level Indonesian health facilities. However, the study was conducted under controlled laboratory conditions that do not fully replicate the operational environment of a primary healthcare clinic or telemedicine deployment. Motion artifacts, ambient temperature variations beyond the 25 to 27°C range maintained in this study, skin tone diversity, and patient populations with chronic cardiovascular or respiratory conditions all represent sources of additional measurement uncertainty that were excluded from this study's controlled design, consistent with the limitations stated in Chapter 1.6.

The comparison of this study's results with the broader literature reveals that the integrated system achieves accuracy levels competitive with single-parameter validation studies: the systolic error of 1.04% improves on both prior MPX5050-family studies, the temperature MAE of 0.1000°C is better than all comparable MLX90614 studies reviewed in Chapter 2.1.2, and the heart rate and SpO_2 results are consistent with the best published outcomes for the GY-MAX30102. The achievement of competitive accuracy across all four parameters simultaneously in a single integrated system represents a meaningful contribution to the field of low-cost biomedical instrumentation, particularly for applications in developing countries where access to full clinical monitoring equipment is limited, as emphasized by the World

Health Organization (2024) in noting that approximately 40% of healthcare facilities in developing countries lack standard vital sign monitoring equipment.

4. CLOSING

4.1 Conclusions

This study was conducted to empirically analyze the accuracy, precision, and reliability of the MPX5050DP, MLX90614, and GY-MAX30102 sensors compared to certified standard medical devices in measuring four main vital signs of patients: blood pressure, body temperature, heart rate, and oxygen saturation. Based on the experimental results and statistical analyses presented in Chapter 4, the following conclusions are drawn to directly answer each of the three research objectives stated in Chapter 1.3.

1. Conclusion 1 Accuracy Level of Each Sensor Compared to Standard Medical Devices

   The three sensors demonstrated the following post-calibration accuracy levels when compared to their respective standard medical reference instruments.

   The MPX5050DP sensor, evaluated against a certified digital sphygmomanometer, produced a mean percentage error of 1.04% for systolic blood pressure and 1.74% for diastolic blood pressure. Systolic blood pressure achieved a mean absolute error of 1.2000 mmHg, which falls within the ±2% full-scale output tolerance of ISO 80601-2-12:2020. Diastolic blood pressure achieved a mean absolute error of 1.2667 mmHg, also within the same tolerance boundary. Both outcomes represent a substantial improvement over the pre-calibration state, where systolic error was 12.53% and diastolic error was 15.16%. These results outperform the accuracy reported by Sari et al. (2024) who documented a systolic error of 3.45% and diastolic error of 7.17% for the same sensor variant, as well as Aulia et al. (2024) who reported 2.23% and 4.69% for the MPX5050GP. The MPX5050DP sensor therefore meets the applicable international accuracy standard with a safety margin of 52.0% for systolic and 49.3% for diastolic measurements.

   The MLX90614 sensor, evaluated against a certified digital clinical thermometer, produced a mean percentage error of 0.27% and a mean absolute error of exactly 0.1000°C. This value falls within the ±0.2°C maximum permissible error of ISO 80601-2-56:2017, and every single one of the fifteen individual measurements complied with this tolerance, not merely the average. The pre-calibration error was 3.72%, reduced by 92.8% through the application of a +1.38°C firmware offset correction. This result also satisfies the World Health Organization (2024) requirement of ±0.3°C maximum error

for non-contact thermometers deployed in tropical climates, confirming the MLX90614's suitability for use in Indonesian healthcare environments.

The GY-MAX30102 sensor, evaluated against a certified pulse oximeter, produced a mean absolute error of 1.0000 bpm with a mean percentage error of 1.32% for heart rate, and a mean absolute error of 0.6667% with a mean percentage error of 0.685% for oxygen saturation. Both fall within their respective standards: ±3 bpm or ±3% per IEC 60601-2-49:2020 for heart rate, and ±3% per ISO 80601-2-61:2019 for SpO_2. The heart rate accuracy is competitive with the 2.02% maximum error reported by Triwiyanto et al. (2022) and the ±2.82% reported by Shavira et al. (2024). The SpO_2 accuracy is consistent with the 0.88% average error reported by Shavira et al. (2024) and approaches the 0.43% benchmark of Triwiyanto et al. (2022). Both parameters maintain safety margins of 66.7% and 77.8% respectively from their ISO tolerance boundaries.

2. Conclusion 2 Measurement Deviation of Each Sensor Compared to the Medical Reference

The measurement deviations between sensor readings and standard device readings, quantified through MAE, RMSE, Pearson correlation coefficient, and significance testing, are summarized as follows.

For the MPX5050DP sensor, the systolic blood pressure deviation averaged 1.2000 mmHg (RMSE = 1.4606 mmHg) with a very strong Pearson correlation of $r = 0.9763$ ($r^2 = 0.9532$), indicating that the sensor explains 95.3% of the inter-individual variance in standard device readings. The Wilcoxon Signed-Rank Test yielded $Z = -0.111$ and $p = 0.912$, confirming no statistically significant systematic bias. For diastolic blood pressure, the deviation averaged 1.2667 mmHg (RMSE = 1.5706 mmHg) with a very strong Pearson correlation of $r = 0.9564$ ($r^2 = 0.9148$), and the Paired t-Test produced $t = -1.164$, $df = 14$, $p = 0.264$, confirming no significant systematic bias.

For the MLX90614 sensor, the body temperature deviation averaged 0.1000°C (RMSE = 0.1291°C). The Pearson correlation of $r = 0.3129$ appears low but is attributed to the range restriction effect: all respondents were normothermic healthy young adults within a 0.40°C data range, producing artificially compressed correlation statistics despite excellent absolute accuracy. The Wilcoxon Signed-Rank Test yielded $Z = -1.132$ and $p = 0.258$, confirming no statistically significant systematic bias between the MLX90614 and the clinical thermometer.

For the GY-MAX30102 sensor, the heart rate deviation averaged 1.0000 bpm (RMSE = 1.2910 bpm) with the highest Pearson correlation of all parameters at $r = 0.9849$ ($r^2 = 0.9700$), indicating that 97.0% of the inter-individual heart rate variance is captured by

the sensor. The Paired t-Test produced t = -1.000, df = 14, p = 0.334, confirming no significant systematic bias. For SpO_2, the deviation averaged 0.6667% (RMSE = 0.8165%) with a strong Pearson correlation of r = 0.7743 (r^2 = 0.5995). The Wilcoxon Signed-Rank Test yielded Z = -1.897 and p = 0.058, which, while the closest result to the significance threshold, still retains the null hypothesis and confirms no statistically significant systematic bias at the alpha = 0.05 level.

All five parameters demonstrated measurement deviations below their respective ISO and IEC tolerance thresholds, confirming that the deviations observed are clinically acceptable and not statistically indicative of instrument failure or systematic measurement error.

3. Conclusion 3 Feasibility of the Combined Sensor System

The integrated system consisting of the MPX5050DP, MLX90614, and GY-MAX30102 sensors connected to the ESP32 microcontroller is evaluated as feasible for use as a low-cost, portable, multi-parameter vital sign monitoring device in primary healthcare settings, based on the following converging lines of evidence.

First, all five sensor-parameter combinations comply with their respective ISO and IEC medical device standards: ISO 80601-2-12:2020 for blood pressure, ISO 80601-2-56:2017 for body temperature, IEC 60601-2-49:2020 for heart rate, and ISO 80601-2-61:2019 for SpO_2, as confirmed in Chapter 4.8. The minimum safety margin from any tolerance boundary is 49.3%, providing substantial buffer against additional measurement variability expected in field deployment conditions.

Second, all five significance tests returned p-values above the 0.05 threshold, confirming that none of the three sensors produces a statistically significant systematic bias against their respective standard medical devices. The absence of systematic bias is a fundamental requirement for a reliable monitoring instrument, as systematic bias would compound over time and lead to clinically significant misclassification of patient conditions.

Third, all five Cronbach's Alpha values exceeded the 0.70 minimum reliability threshold, with four parameters achieving very high reliability above 0.90. The values ranged from 0.8738 for SpO_2 to 0.9948 for Heart Rate, confirming that repeated measurements within the same session produce highly consistent results. This level of reliability supports the use of the system for continuous monitoring applications where consistency across measurements is as important as accuracy against the standard.

Fourth, the integrated system achieved all of these outcomes simultaneously across four vital sign parameters using three sensors and a single ESP32 microcontroller, a

configuration that addresses the research gap identified in Chapter 2.1 regarding the absence of validated multi-parameter systems using these specific low-cost sensors. The system therefore represents a practical and technically sound alternative to conventional clinical monitoring equipment for resource-limited healthcare environments, consistent with the broader objective stated in Chapter 1.1 of developing affordable IoT-based health monitoring devices suitable for primary healthcare services and patient telemonitoring systems in developing countries such as Indonesia.

4.2 Recommendations

The following recommendations are derived directly from the research limitations stated in Chapter 1.6 and the research gaps identified in Chapter 2.1. Each recommendation points toward a specific and feasible next step that would extend, strengthen, or broaden the contributions of this study.

Recommendation 1 Field Validation Under Real-World Conditions**.** This study was conducted in a controlled laboratory environment with ambient temperature maintained between 25 and 27°C, subjects at rest, and minimized ambient light and motion interference. These conditions do not fully represent the operational environment of a primary healthcare clinic, community health post, or telemedicine station in Indonesia, where ambient temperatures may exceed 30°C, humidity is higher, and patients may be ambulatory or in discomfort. Future research should replicate the measurement protocol of this study under field conditions, specifically testing the sensor system against certified medical devices in community health centers across different climate zones in Indonesia. Field validation would produce a more realistic estimate of the additional measurement uncertainty introduced by environmental variability and would test whether the safety margins established in Chapter 4.8 are sufficient to maintain ISO compliance under operational conditions.

Recommendation 2 Long-Term Clinical Trials with Diverse Populations**.** This study used a sample of five healthy university students aged 20 to 23 years. While this controlled sample is appropriate for an initial accuracy validation, it excludes populations of highest clinical relevance: elderly patients, individuals with hypertension or diabetes, patients with chronic respiratory disorders, and individuals with conditions that affect peripheral perfusion or skin characteristics. The Wilcoxon test result for SpO_2 at $p = 0.058$, which is the closest to the significance threshold, suggests that a larger and more diverse sample might reveal statistically significant deviations for certain population subgroups. Future studies should increase the sample size toward the $n = 19$ recommended by the Slovin formula applied in Chapter 3.4 at minimum, and ideally conduct a longitudinal clinical trial with patients recruited from hospital

outpatient departments across multiple age groups and clinical conditions. Such a study would establish whether the compliance margins demonstrated here hold across the full range of patients the system would serve in deployment.

Recommendation 3 Integration of Machine Learning for Adaptive Calibration**.** The calibration procedure applied in Chapter 4.1 used a static linear offset correction, which was highly effective at eliminating the systematic pre-calibration bias but left a residual MAE of 1.20 mmHg for systolic and 1.27 mmHg for diastolic blood pressure. Liang et al. (2025) demonstrated that machine learning-assisted calibration of low-cost pressure sensors could reduce the mean absolute error by an additional 3 to 5 mmHg beyond what linear offset correction achieves, through the use of adaptive algorithms that account for environmental temperature, individual physiological characteristics, and sensor drift over time. Future iterations of this system should explore the integration of regression-based or neural network-based adaptive calibration modules within the ESP32 firmware, particularly for the MPX5050DP pressure sensor, which showed the largest absolute deviations and the widest inter-individual variability. This enhancement would not require hardware changes and could be implemented entirely at the software level, preserving the low-cost design philosophy of the current prototype.

Recommendation 4 Extension to Additional Vital Parameters and Formal Standardization**.** The current system covers four vital sign parameters across three sensors. Future development should consider extending the system to include respiratory rate, which was mentioned in the original scope of Chapter 1.1 but could not be measured by the current sensor configuration, and blood glucose monitoring using optical sensing techniques currently under development in the biomedical engineering community. Additionally, the ISO 80601-2-12:2020 standard applied for blood pressure in this study governs respiratory pressure in ventilators rather than peripheral sphygmomanometry. A dedicated validation study following the AAMI SP10 or ISO 81060-2 clinical blood pressure device standard would strengthen the regulatory standing of the MPX5050DP for clinical blood pressure applications. Formal calibration traceability certification by an accredited metrology laboratory, consistent with the requirements of Peraturan Menteri Kesehatan No. 62 Tahun 2017, would be required before the system could be considered for official clinical use in Indonesian healthcare facilities, and this process should be initiated in parallel with any follow-up field validation study.

BIBLIOGRAPHY

Aulia, F. N., Pudji, A., & Sumber, S. (2024). Integrated digital sphygmomanometer for simultaneous blood pressure and body temperature monitoring. *Jurnal Teknokes*, *17*(2). https://teknokes.org/index.php/teknokes/article/view/58

Ayieko, B. O. (2024). Microcontroller-based design and development of a real-*time body temperature and heart rate monitoring system*. JKUAT Institutional Repository. http://ir.jkuat.ac.ke/handle/123456789/6520

Aziz, M., Nuryani, N., & Darmanto, D. (2022). Alat ukur tekanan darah menggunakan proses komputer. *Jurnal Fokus Elektroda*. https://ojs.uho.ac.id/index.php/JFE/article/view/17601/0

Barzan, M., & Arbilei, D. M. N. (2024). *Design and manufacturing of intensive care unit biomedical devices*. ResearchGate. https://www.researchgate.net/publication/377964896

Beri, M., Kumar, B., Tiwari, S., Sharma, N., Vashishtya, H., & Chaudhary, P. (2022). IoT based health monitoring system built on ESP32. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 454–458. https://doi.org/10.1109/ICACITE53722.2022.9823528

Blažauskas, T., Damaševičius, R., & Maskeliūnas, R. (2023). Smartphone-based voice wellness index application for dysphonia screening and assessment: Development and reliability. *Journal of Voice*, *37*(5), 812–823. https://doi.org/10.1016/j.jvoice.2023.01.020

Brahimaj, D. (2020). *Design and development of a wearable device for continuous monitoring of vital signs*. Politecnico di Torino. https://webthesis.biblio.polito.it/14377/

Chattopadhyay, S., & Das, R. (2022). A framework for testing the reliability and validity of a novel non-invasive digital biomarker instrument using statistical techniques: A case study with Lyfas. *Transactions in Machine Learning and Artificial Intelligence*. https://www.researchgate.net/publication/359453099

Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Goldenits, P., Dorner, T. E., Grabner, G., Mayer, S., & Zehetmayer, S. (2023). Evaluating blood oxygen saturation measurements by popular fitness trackers in postoperative patients: A prospective clinical trial. *iScience*, *26*(11), 108224. https://doi.org/10.1016/j.isci.2023.108224

Habibi, I., Jufrizel, A., & Maria, P. S. (2022). Non-contact thermometer using infrared temperature sensor MLX90614. *Indonesian Journal of Electrical Engineering and Renewable Energy*, *2*(1), 19–28. https://repository.uin-suska.ac.id/61265

Hasija, Y. (2023). Statistical methods in bioinformatics. In *All about bioinformatics* (pp. 43–75). Elsevier. https://doi.org/10.1016/B978-0-443-15250-4.00009-5

Idiong, U. E., Abdulhamid, I. G., & Mosaku, A. O. (2023). *IoT framework for real-time health monitoring: Fingertip sensor-based*. ResearchGate. https://www.researchgate.net/publication/393009402

Khalid, M. S., Rahman, H., & Memon, F. (2023). A low-cost PPG sensor-based empirical study on healthy aging based on changes in PPG morphology. *arXiv preprint arXiv:2312.13404*. https://arxiv.org/abs/2312.13404

Kowshik, M. R., Narayan, A., & Nagarajan, P. M. (2025). Dynamic dual-function safety wristband: A simulation-driven design for enhancing women's security and personal safety. *Procedia Computer Science*, *242*, 1–12. https://www.sciencedirect.com/science/article/pii/S1877050925010841

Liang, Z., Wang, Y., & Chen, J. (2025). Rapid adaptation of SpO_2 estimation to wearable devices via transfer learning on low-sampling-rate PPG. *arXiv preprint arXiv:2509.12515*. https://arxiv.org/abs/2509.12515

Liu, C., Zhang, J., & Wang, P. (2024). Signal filtering algorithms to enhance biomedical sensor accuracy. *Biomedical Signal Processing and Control*, *93*, 105630. https://doi.org/10.1016/j.bspc.2023.105630

Mazoteras-Pardo, V., Gomez-Cantarino, S., Ramirez-Jimenez, M., & Navarro-Flores, E. (2023). Validations of blood pressure measuring devices using recognized protocols. *Journal of Personalized Medicine*, *13*(1), 9. https://doi.org/10.3390/jpm13010009

Nemomssa, M., & Raj, M. (2022). Development of low-cost and portable pulse oximeter device with improved accuracy and accessibility. *PubMed*.

Palavesam, P. D., Murugan, K., Sivanesakumar, B., & Vaikundam, S. (2023). Design and development of health monitoring system using sensor. *AIP Conference Proceedings*, *2754*(1), 050004. https://doi.org/10.1063/5.0164395

Rafl, J., Bachman, T. E., Rafl-Huttova, V., Walzel, S., & Rozanek, M. (2022). Commercial smartwatch with pulse oximeter detects short-time hypoxemia as well as standard medical-grade device: Validation study. *DIGITAL HEALTH*, *8*, 1–12. https://doi.org/10.1177/20552076221132127

Rahman, M., & others. (2021). Accuracy evaluation framework for biomedical sensors in clinical applications. *IEEE Access*, *9*, 12291–12305. https://doi.org/10.1109/ACCESS.2021.3056782

Rayan, A., Faris, M., & Ashoub, M. (2023). Machine learning-based respiration rate and blood oxygen saturation estimation using photoplethysmogram signals. *Biosensors*, *13*(2), 298. https://doi.org/10.3390/bios13020298

Reddy, R. A., Sengupta, R., Jackson, B. M., & Lewis, C. (2023). Development of a new measure to check attitude towards water conservation. *MethodsX*, *10*, 102012. https://doi.org/10.1016/j.mex.2022.102012

Saldana-Perez, M., & Contreras-Jiménez, Y. (2025). Advances in technology for healthcare monitoring in Latin America: A systematic review. In *Elsevier Innovative Healthcare*. https://www.sciencedirect.com/science/article/pii/B9780443274442000125

Sari, A. N., Pramono, C. N., Mak'ruf, M. R., Ariswati, H. G., Pujiastuti, M. P. A. T., & Maghfiroh, A. M. (2024). Design of ambulatory blood pressure monitoring for IoT-based hypertension patients. *Jurnal Teknokes*, *17*(1).

Shavira, N., Ariswati, H. G., & Hamzah, T. (2024). SpO_2 analysis on development of IoT-based lung function and SpO_2 measuring device. *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics (IJEEEMI)*, *6*(1). https://ijeeemi.poltekkesdepkes-sby.ac.id/index.php/ijeeemi/article/view/305

Shi, C., Goodall, G., Rooney, D., & Bates, A. (2024). Effect of skin tone on the accuracy of the estimation of arterial oxygen saturation by pulse oximetry: A systematic review. *eBioMedicine*, *101*, 105025. https://doi.org/10.1016/j.ebiom.2024.105025

Teh, H. Y., Kempa-Liehr, A. W., & Wang, K. I.-K. (2020). Sensor data quality: A systematic review. *Journal of Big Data*, *7*(1), 11. https://doi.org/10.1186/s40537-020-0285-1

Triwiyanto, A., & others. (2022). Pulse oximeter design for SpO_2 and BPM recording on external memory to support the COVID-19 diagnosis. *Jurnal Teknokes*, *15*(3), 147–153. https://teknokes.poltekkesdepkes-sby.ac.id/index.php/teknokes/article/view/282

Triwiyanto, T., Wahyunggoro, O., Nugroho, H. A., & Setiawan, H. (2022). Design of heart rate, oxygen saturation, and temperature monitoring system for Covid-19 patient based on Internet of Things (IoT). *ResearchGate*. https://www.researchgate.net/publication/359137562

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83. https://doi.org/10.2307/3001968

World Health Organization. (2024). *Non-contact thermometer validation for tropical climate applications*. WHO Press.

Yunidar, A., Prasetya, B., & Sulistyo, A. (2023). IoT-based health monitoring using MAX30102 and ESP32: Accuracy evaluation for SpO_2 and heart rate. *Journal of Electrical Engineering and Computer Science*, *29*(1), 55–63.

Zuccotti, G., Agnelli, P. O., Labati, L., Cordaro, E., Braghieri, D., Balconi, S., Mandalari, D., Fiorina, P., Bertoli, S., Berra, C. C. F., Croci, M., Losurdo, F., Foppiani, A., Bucciarelli, L., Xodo, M., De Pasquale, S., & Calcaterra, V. (2025). Accuracy of heart rate, pulse oxygen saturation, and blood pressure using a non-contact photoplethysmography-based mobile application: A potential tool for wellness monitoring. *DIGITAL HEALTH*, *11*, 1–14. https://doi.org/10.1177/20552076251351841