Original Article

Technical Error Measurements, Reliability, and Validity of Customized Anthropometric Grid

*Siti Nur Sarah binti Zubir,[a] Mohd Nur Ikhwan bin Shafiee, Siti Zaharah binti Abd. Rahman, Hazwan Adli bin Hamadan, Nor Shahira binti Mohd Salim, Mohd Esa bin Baruji [a], Evelyn Tan [a], Raemy bin Md Zein [a]

[a] *Consultation Research & Development Department, National Institute of Occupational Safety and Health (NIOSH), Lot 1, Jalan 15/1, Section 15, 43650 Bandar Baru Bangi, Selangor, Malaysia.*

*Corresponding author: sitinursarah@hotmail.com

Article history

Received 20/5/2021
Accepted (Panel 1) 11/4/2023
Accepted (Panel 2) 20/4/2023

ABSTRACT: *The National Institute of Occupational Safety and Health (NIOSH) Malaysia aimed to study anthropometric measurements for the working-age population. However, there is limited literature on the reliability and technical error of measurement (TEM) in anthropometric measurement. This study assessed the properties of stature, eye height, tibial height, bideltoid breadth, elbow span, and waist circumference of 10 volunteers. Four observers measured the mentioned dimensions using a Martin anthropometer and designated anthropometric grid. Findings showed high inter- and intra-observer reliability using 'mean absolute difference' and 'reliability coefficient.' Hence, measurements of the above dimensions using the test instruments may be considered reliable and valid within the limits of their error. We recommend that special attention be given to improving reliability and validity of anthropometric measurements to ensure the accuracy of data for ergonomic design.*

Keywords – *Reliability, Technical Error of Measurement, Validity*

1.0 INTRODUCTION

Anthropometric evaluation of a group of people is critical for developing ergonomic products, tools, and workstations for all populations (Mohd Nur Ikhwan et al., 2021). Anthropometry is the measurement of humans, encompassing any physiological, psychological, or anatomical traits (Ulijaszek & Kerr, 1999). Measurements are essential for the analysis and design of human use and optimizing working and living conditions. These measurements provide data on the engineering and personal protective equipment that should be used or designed. This would ensure a proper match between the assigned work and the individual worker, comfort and ease of equipment use, and effective control of the equipment. Most equipment used in Malaysia are made in other countries and the design is based on the country's nationals' anthropometric data, which may not be the same as that of Malaysian nationals. All anthropometric measurements require measurement tools. Any equipment or observers are liable to errors that must be considered as they are propagated in subsequent statistical analyses and influence interpretations and conclusions (Arroyo, Freire, Ansotegui, & Rocandio, 2010). Variations in human traits due to biological differences and experimental environments may result in measurement errors owing to fluctuations in biological or mechanical factors. Although obtaining an error-free true value is ideal, it is practically impossible to achieve this with continuous measurements. However, it is crucial to minimize such errors as accurately as possible to detect actual variations and improve performance (Ryan-Stewart, Faulkner, & Jobson, 2022). Measurement errors owing to technical, personal, or measurement factors can be minimized through appropriate equipment calibration and control of observer bias.

Estimating the technical error of measurement (TEM) in anthropometric studies is essential for ensuring accuracy and reproducibility. Studies that do not estimate the TEM are prone to significant errors (Krishan & Kanchan, 2016). Various terms are used to describe anthropometric measurement errors, including coefficient of reliability (R), technical error measurement (TEM), and validity (Ulijaszek & Kerr, 1999). In the anthropometric literature, precision refers to the degree of variability between repeated measurements on a subject by the same observer (intra-observer precision) or by different observers (inter-observer precision) (Arroyo et al., 2010). Reliability is the degree to which within-subject variability is present and is caused by factors other than the variance of measurement error or physiological variation. The second type of measurement error, validity, is the extent to which the 'true' value of a measurement is attained. The technical error of measurement (TEM) is another accuracy index used to express error margins in anthropometry (Jamaiyah et al., 2010). Furthermore, there is a lack of research on the technical error of measurement (TEM) of anthropometric measurements in Malaysia. Considering these research gaps, the purpose of the present study was to describe the standardization process of the inter- and intra-observer reliability of the anthropometric measurements used and the validity of the designated anthropometric grid.

2.0 MATERIAL AND METHOD

Four beginner observers were analyzed using the coefficient of reliability (R) and TEM calculations for their measurement results to interpret the proportion of between-subject variance free from measurement error and to verify the intra- and inter-evaluator variation, respectively. All observers were trained during the study phase to familiarize themselves with the anthropometric techniques. Ideally, duplicate measurements of at least 10 individuals should be performed to calculate the intra- and interobserver TEM and R (Ulijaszek & Kerr, 1999). Many previous studies have used similar methods, such as those of Carsley et al. (2019) and Hardy et al. (2018).

2.1 Anthropometric methods

Anthropometric measurements are non-invasive quantitative assessments of the human body used to analyze various physical characteristics. These measurements are widely recognized as some of the most reliable tools available for evaluating and quantifying human dimensions (Casadei & Kiel, 2023). Anthropometric measurements were performed using a Martin Anthropometer (TTM, Japan) to the nearest 1.0 mm and a designated anthropometric grid designed (Figure 1) from a previous study (Nurani et al., & Hari Krishnan (2015) to the nearest 0.1 cm. All dimensions were measured in centimeters (cm) to the nearest 0.1 cm. To obtain the TEM, R values, and validity, each observer measured ten volunteers on two different days. On the first day, the observers measured the subjects in the morning and evening, whereas on the second day, all subjects were measured only in the evening. The method used in this study was adapted from Perini et al. (2005).

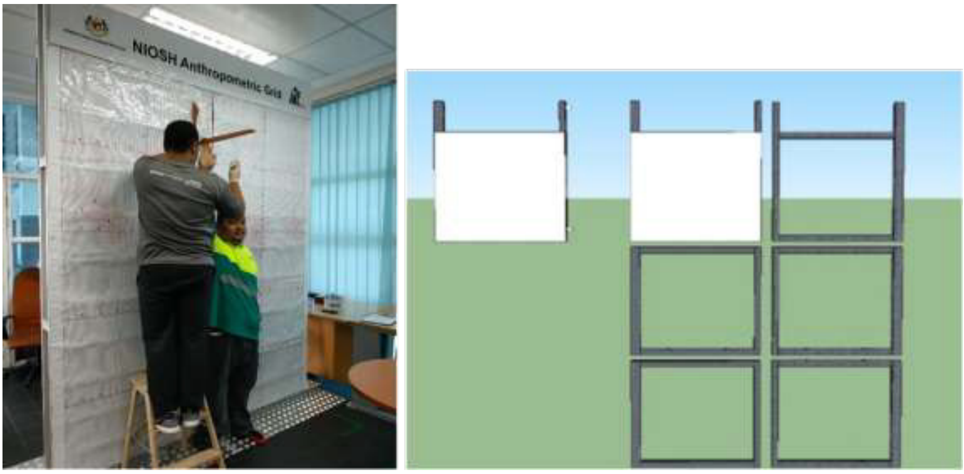

Figure 1: Designated anthropometric grid

2.3 Statistical analysis

For reliability, the findings of the statistical analyses were reported using the mean absolute difference (MAD) and correlation coefficient (R). The absolute mean difference is a crude method to check for differences or agreement between two readings. The correlation coefficient (R) was used to assess the reliability more objectively; the values for the reliability coefficient ranged from 0 to 1. A coefficient of below 0 indicates "no reliability", >0 to <0.2 is slight reliability, 0.2 - <0.4 is fair reliability, 0.4 - <0.6 is moderate, 0.6 - < 0.8 is substantial, and 0.8 – 1.0 is almost perfect reliability (Carsley et al., 2019; Haniff Jamaiyah et al., 2008). TEM is obtained by performing several repeated measurements on the same subject, either by the same observer or by two or more observers, taking the differences and entering them into an appropriate equation. The calculations for intra- and inter-observer errors were broadly the same. For intra-observer analysis, TEM involves two measurements, and inter-observer TEM involves two measurements using Equation (1).

D is the sum of the deviations and N is the number of volunteers measured. When more than two observers are involved, the equation for estimating the inter-observer TEM is more complex using Equation (2).

Where N is the number of participants, K is the number of observers (assuming one determination per observer) for the variable for each participant, and M is the measurement. The units of TEM were the same as those of the anthropometric measurements in this study, centimeters (cm).

The positive association between TEM and the measurement size is problematic because the relative imprecision of different measurements cannot be assessed. To compare the TEM of different variables or populations, Perini et al. (2005) recommended the conversion of absolute TEM to relative TEM (% TEM) using Equation (3).

Despite making a direct comparison of different anthropometric measures possible, % TEM provides no information for the comparison of studies in which more than one observer is used, and in which both intra- and inter-observer TEM are reported. There are two methods to overcome this problem. The first is to square the TEM, turning them into variances, summing them, and then taking the square root to obtain the total TEM. For the two observers and two measurements per observer, where, TEM (intra1) is the intraobserver TEM for the first observer, TEM (intra2) is the intraobserver TEM for the second observer, and TEM (inter) is the interobserver TEM between the two observers. TEM (intra1), TEM (intra2), and TEM (inter) were calculated using Equation (1). A value for TEM (intra) for each observer, calculated using Equation (1), is incorporated in Equation (4). All TEM (intra-observer) values were squared, summed, and divided by the number of observers. Furthermore, with more than two observers, TEM (inter) was calculated using Equation (2):

This value can then be used to compare measurement errors across studies, regardless of the number of observers used. Another approach to obtain comparability of anthropometric measurement errors is to use the coefficient of reliability (R), which ranges from 0 to 1 and can be calculated using the following equation:

SD2 is the total intersubject variance for the study, including measurement error. All analyses were performed using Microsoft Excel. The coefficient of reliability (R) ranges from 0 to 1 and can be calculated using the following equation:

$$R = 1\text{-}\left(\frac{(\text{total TEM})^2}{SD^2}\right)$$

SD^2 is the total intersubject variance for the study, including measurement error. Mueller and Martorell (1988) reported that this coefficient is the most commonly used measure of anthropometric precision in population studies. All analyses were performed using Microsoft Excel.

3.0 RESULT

3.1 Sample characteristics

The anthropometrical measurements were extracted from a sample of 10 NIOSH workers (age: 27.70 ± 2.26; female: 6 and male: 4) who volunteered to participate in the study.

3.2 Reliability

3.2.1 *Intra-observer reliability*

Table 1 Intra-observer reliability

	Mean of First Measurement (cm)	Mean of Second Measurement(cm)	Mean Absolute Difference (cm)	% TEM
Stature				
Observer 1	160.4 ± 8.6	160.4 ± 8.3	0	0.3
Observer 2	160.7 ± 8.6	160.6 ± 8.6	0.10	0.4
Observer 3	157.7 ± 7.7	157.8 ± 7.5	0.05	0.4
Observer 4	159.1 ± 8.1	159.5 ± 7.9	0.46	0.4
Eye height				
Observer 1	150.0 ± 8.6	149.7 ± 8.1	0.23	0.4
Observer 2	149.8 ± 8.1	149.6 ± 8.1	0.20	0.3
Observer 3	148.0 ± 8.1	147.4 ± 8.1	0.60	0.6
Observer 4	148.5 ± 7.9	148.9 ± 7.7	0.38	0.4
Tibial height				
Observer 1	41.7 ± 3.3	41.6 ± 3.2	0.10	0.6
Observer 2	41.3 ± 3.3	41.7 ± 3.0	0.34	1.4
Observer 3	41.4 ± 3.2	41.6 ± 3.2	0.16	0.7
Observer 4	41.6 ± 2.9	41.7 ± 3.1	0.04	0.4
Bideltoid breadth				
Observer 1	41.9 ± 3.9	41.8 ± 3.8	0.10	0.7
Observer 2	41.9 ± 4.2	42.1 ± 4.2	0.20	0.8
Observer 3	42.2 ± 4.2	41.9 ± 3.9	0.3	0.8
Observer 4	41.9 ± 3.7	41.8 ± 3.7	0	0.9
Elbow span				
Observer 1	83.4 ± 4.9	83.0 ± 4.8	0.36	0.5
Observer 2	83.3 ± 4.9	83.2 ± 4.7	0.06	0.5
Observer 3	83.2 ± 4.0	83.3 ± 4.3	0.09	0.5
Observer 4	82.7 ± 4.2	83.1 ± 4.5	0.46	1.0
Waist circumference				
Observer 1	15.4 ± 1.3	15.3 ± 1.5	0.10	2.5
Observer 2	15.3 ± 1.3	15.1 ± 1.4	0.20	1.6
Observer 3	15.2 ± 1.6	15.0 ± 1.2	0.10	2.7
Observer 4	15.2 ± 1.6	15.0 ± 1.2	0.20	2.5

Anthropometric measurements are presented in Table 1. The mean absolute difference (MAD), which measures the average difference between two sets of measurements, ranged from 0 to 0.60. The MAD indicate that the differences between the measurements were minimal. The technical error of measurement (TEM), which measures measurement precision, ranged from 0.3% to 2.7%. This suggests that the measurements were relatively precise with some variability depending on the measurement.

3.2.2 Inter-observer reliability

Table 2 Inter-observer reliability

	Mean (cm)	SD	Total TEM	% Total TEM	R
Stature	159.1	7.8	1.3	0.8	0.972
Eye height	148.9	7.6	1.3	0.9	0.969
Tibial height	41.6	3.0	0.5	1.3	0.968
Bideltoid breadth	42.3	3.7	1.5	3.6	0.832
Elbow span	83.4	4.3	0.7	0.9	0.970
Waist circumference	15.3	1.3	0.4	2.9	0.888

The results of the inter-observer variability in anthropometric measurements are presented in Table 2, which includes the percentage of total technical error of measurement (TEM) and the coefficient of reliability (R) for each measurement. The total TEM values ranged from 0.8 to 3.6, indicating the error observed for each measurement. The coefficients of reliability (R) for all measurements were > 0.8, suggesting good reliability and consistency of the measurements obtained by different observers.

3.3 Validity of Designated Anthropometric Grid

Table 3 Inter-instrument (Validity)

	Mean (cm)	SD	Total TEM	% Total TEM	R
Stature	159.1	7.8	2.5	1.5	0.900
Eye height	149.0	7.9	1.9	1.2	0.940
Tibial height	41.6	3.0	0.5	1.0	0.972
Bideltoid breadth	42.0	3.7	0.6	1.2	0.976
Elbow span	83.3	4.2	1.0	1.1	0.946

Table 3 presents the comparison results between the Martin anthropometer and the designated anthropometric grid. The technical error of measurement (TEM) for all measurements ranged from 1.0 to 2.5, indicating a relatively small margin of error. Moreover, all correlation coefficients (R values) were greater than 0.8, indicating a strong correlation between the Martin anthropometer and the designated anthropometric grid.

4.0 DISCUSSION

This study reports the intra- and interobserver reliabilities of stature, eye height, bideltoid breadth, tibial height, elbow span, and waist circumference using multiple reliability statistics, mean absolute difference, and technical error measurement of the coefficient of reliability (R). Similar analyses were performed for intra-instrument (validity) between the Martin anthropometer and designated anthropometric grid. All mean absolute differences were minimal. Almost all intra- and inter-observer > 1.5 and > 2.0, respectively. According to Table 1, Table 2, and Table 3, some technical measurement errors (TEM) are above the acceptable range. The TEM value may be attributed to all the observers being beginner anthropometrists who were unfamiliar with the body landmarks used in the measurements. However, all R values in this study were greater than 0.8, indicating acceptable reliability. TEM can be considered as an evaluation session for observers to improve anthropometric data quality (Conkle, Ramakrishnan, Flores-Ayala, Suchdev, & Martorell, 2017).

The validity of the designated anthropometric grid as a test instrument compared to the reference instrument, the Martin anthropometer, was also assessed in the present study. The technical error of measurement (TEM) for the designated anthropometric grid was within an acceptable limit, whereas the correlation coefficients (R values) between the two instruments were above 0.9. These results indicated the high accuracy of the designated anthropometric grid as a measurement tool for anthropometric studies. The designated anthropometric grid readings were remarkably close to those obtained using the Martin anthropometer. Therefore, the designated anthropometric grid can be considered as a suitable alternative measurement instrument for anthropometric studies.

5.0 CONCLUSION

In conclusion, our results showed acceptable precision for the evaluated anthropometric dimensions considering that the observers were beginners. Although some TEM values were above the acceptable limit, the R values for all measurements were close to 1.0. Thus, the designated anthropometric grid is relatively reliable and valid for use in anthropometric development. In conclusion, this study aimed to assess the reliability and validity of a designated anthropometric grid as a measurement instrument for anthropometric studies. The study found that the designated anthropometric grid demonstrated high accuracy compared with the Martin anthropometer, with a technical error of measurement (TEM) within the acceptable limit and correlation coefficients (R-values) above 0.9.

Meanwhile, the intra- and inter-observer reliabilities for the measured parameters were acceptable, despite some technical errors of measurement being above the acceptable range. These results suggest that the designated anthropometric grid can be considered a suitable alternative to the Martin anthropometer for use in anthropometric studies and that the reliability of measurements can be improved with the training and experience of observers. Overall, this study provided valuable insights into the validity and reliability of anthropometric measurements using a designated anthropometric grid, which could improve the accuracy and quality of anthropometric data collection in future studies.

ACKNOWLEDGEMENTS

We thank the research assistants who participated in the study. We also express our gratitude to the NIOSH, Malaysia. for funding. This research was funded by NIOSH Malaysia under project code (03.08/03/NG06-01/ANTHROPOMETRIC/2018/01) which was approved in June 2018.

REFERENCES

Arroyo, M., Freire, M., Ansotegui, L., & Rocandio, A. M. (2010). Intraobserver error associated with anthropometric measurements made by dietitians. Nutricion Hospitalaria, 25(6), 1053–1056. https://doi.org/10.3305/nh.2010.25.6.4854

Carsley, S., Parkin, P. C., Tu, K., Pullenayegum, E., Persaud, N., Maguire, J. L., & Birken, C. S. (2019). Reliability of routinely collected anthropometric measurements in primary care. BMC Medical Research Methodology, 19(1), 1–8. https://doi.org/10.1186/s12874-019-0726-8

Conkle, J., Ramakrishnan, U., Flores-Ayala, R., Suchdev, P. S., & Martorell, R. (2017). Improving the quality of child anthropometry: Manual anthropometry in the Body Imaging for Nutritional Assessment Study (Bina). PLOS ONE, 12(12), e0189332. https://doi.org/10.1371/journal.pone.0189332

Hardy, J., Kuter, H., Campbell, M., & Canoy, D. (2018). Reliability of anthropometric measurements in children with special needs. Archives of Disease in Childhood, 103(8), 757–762. https://doi.org/10.1136/archdischild-2017-314243

Jamaiyah, H., Geeta, A., Safiza, M. N., Khor, G. L., Wong, N. F., Kee, C. C., Rahmah, R., Ahmad, A. Z., Suzana, S., Chen, W. S., Rajaah, M., & Adam, B. (2010). Reliability, technical error of measurements and validity of length and weight measurements for children under two years old in Malaysia. Medical Journal of Malaysia, 65, Suppl A, 131–137

Jamaiyah, H., Geeta, A., Safiza, M. N., Wong, N. F., Kee, C. C., Ahmad, A. Z., Suzana, S., Rahmah, R., Khor, G. L., Ruzita, A. T., Chen, W. S., Rajaah, M., & Faudzi, A. (2008). Reliability and technical error of Calf Circumference and Mid-half Arm Span measurements for nutritional status assessment of elderly persons in Malaysia. Malaysian Journal of Nutrition, 14(2), 137–150

Krishan, K., & Kanchan, T. (2016) Measurement Error in Anthropometric Studies and its Significance in Forensic Casework. Annals of Medical and Health Sciences Research, 6(1), 62–63. https://doi.org/10.4103/2141-9248.180277

Ikhwan, Mohd. N., S., Raemy, M. Z., Mohd Esa, B., Evelyn, T., Sarah, Siti N., Z., Siti Zaharah, A. R., … Nor Sahira, S. (2021), Digital Health: Stepping into the Future of Medicine. Development of Malaysian anthropometric portal. In, 8th National Health seminar. Journal of Public Health Medicine. https://doi.org/https Maizatul Azma Masri, Ruhi Fadzlyana Jailani, & Liyana Azmi (Eds.). Nilai: Malaysian, (15). https://doi.org/10.37268/mjphm/vol.22/no.Suppl.1

Perini, T. A., de Oliveira, G. L., Ornelia, J. dos S., & de Oliveira, F. P. (2005). Technical error of measurement in anthropometry. Revista Brasileira de Medicina do Esporte, 11(1), 86–90. https://doi.org/http. https://doi.org/10.1590/S1517-86922005000100009

Ryan-Stewart, H., Faulkner, J., & Jobson, S. (2022). The impact of technical error of measurement on somatotype categorization. Applied Sciences, 12(6). https://doi.org/10.3390/app12063056

Siti Nurani, H., Rosnah, M., Raemy, M., Mohd Rizal, H., & Hari Krishnan, T. (2015). Anthropometric data of Malaysian workers. New Ergonomics Perspective – Selected Papers of the 10th Pan-Pacific Conference on Ergonomics (January), 353–360. https://doi.org/10.1201/b17990-61

Ulijaszek, S. J., & Kerr, D. A. (1999). Anthropometric measurement error and the assessment of nutritional status. British Journal of Nutrition, 82(3), 165–177. https://doi.org/10.1017/s0007114599001348