

Analyzing the correlation between body composition variables and cellular phase angle via computerized bioelectrical impedance analysis (BIA)

Rodrigo Alvaro B. Lopes-Martins^{1,4}, Andria C. Quiuli², Gabriela M. Gomes³, Paulo Ricardo P. França², Pablo Miguel de J. Pereira², Nicolle B. Souza², Osvaldo S. de Araújo Junior², Larissa G. dos Santos², Pedro S. L. Lopes-Martins², Maria Eduarda N. Silva¹, Patrícia Sardinha Leonardo²

¹Laboratory of Biophotonics and Experimental Therapeutics (LABITEX), Evangelical University of Goiás (UniEvangélica), Anápolis (GO), Brazil.

²Laboratory of Health Technologies (LATES), Evangelical University of Goiás (UniEvangélica), Anápolis (GO), Brazil.

³Scientific Initiation Program of the Medicine Course, Evangelical University of Goiás (UniEvangélica), Anápolis (GO), Brazil.

⁴Post-graduate Program in Bioengineering, Brazil University, Itaquera - São Paulo (SP), Brazil.

Abstract:

Background: Various diseases and pathological conditions, as well as inflammatory processes, either in isolation or in combination with advanced age, can lead to significant alterations in body composition. This study investigates the correlation between various body composition metrics and the cellular Phase Angle (PhA) through Computerized Bioelectrical Impedance Analysis (BIA) in a diverse age group of individuals ranging from 18 to 80 years. **Objective:** Our aim was to explore the interplay between body composition variables such as Body Mass Index, Lean Body Mass, Skeletal Muscle Mass, and their potential influence on the cellular integrity as indicated by PhA. **Methods:** Utilizing a dataset of 199 participants, we employed both linear and advanced machine learning models, including Random Forest regression, to analyze the predictive relationships and variable importance within our body composition metrics. **Results:** Initial analyses revealed strong correlations between mass-related measures and suggested complex relationships with the PhA, an indicator of cellular health and membrane integrity. The Random Forest model significantly outperformed simple linear regression in predicting PhA, emphasizing the nonlinear nature of these relationships and the importance of a comprehensive approach in body composition assessment. **Conclusion:** Our findings highlight the nuanced interactions between body composition variables and their collective impact on cellular health as measured by PhA. This study underscores the potential of utilizing PhA alongside traditional metrics for a more nuanced understanding of body composition and its implications for health. Future research should continue to leverage advanced statistical and machine learning techniques to further elucidate these complex relationships, with implications for nutritional, sports medicine, and biomedical fields.

Keywords: Body composition; cellular phase angle; bioelectrical impedance analysis; body mass index.

Corresponding author: Rodrigo Alvaro Brandão Lopes-Martins

E-mail: ralopesmartins@gmail.com

Received: 03 Apr, 2024.

Accepted: 10 May, 2024.

Published: 14 May, 2024.

Copyright © 2023. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided article is properly cited.



BACKGROUND

Various diseases and pathological conditions, as well as inflammatory processes, either in isolation or in combination with advanced age, can lead to significant alterations in body composition⁽¹⁾. Often, these alterations manifest as weight loss and reduced muscle mass, along with imbalances in fluid volume. These challenging characteristics need to be explored in order to reduce the risk of worse functional and clinical

outcomes^(1,2). However, a key focus of therapeutic interventions should be on restoring/preserving muscle mass and function, not merely body weight⁽³⁾, as muscle mass reduction - regardless of age and body weight - can influence the onset, progression, treatment response, and health outcomes⁽⁴⁾.

As first-line screening tools, weight loss and body mass index (BMI) lack sensitivity for detecting muscle mass loss and can even lead to misleading conclusions about patients' nutritional status and physical composition, as well as being uninformative for distinguishing between body water compartments and their distribution^(1,5). This issue becomes especially pronounced in elderly patients, due to conditions like sarcopenia and frailty, which are difficult to diagnose.

In humans, muscle mass remains relatively stable during early life, but after the age of 30, a natural process of muscle mass reduction begins at a rate of 0.5 to 1.0% per year. With aging, the impaired balance between protein synthesis and proteolysis in skeletal muscle results in a progressive decline in mass, strength, and function of the skeletal muscle, defined as sarcopenia⁽⁶⁾. In terms of human health, sarcopenia is associated with an increase in adverse outcomes, including falls, functional decline, frailty, and mortality.

In this context, the bioelectrical impedance analysis (BIA) method has become extremely popular for assessing body composition because it is an easy-to-use, portable, quick, relatively inexpensive, and non-invasive technology. Consequently, BIA is widely used in hospitals, clinics, and other healthcare facilities. However, BIA does not directly measure body fat and lean mass but measures the whole-body impedance, i.e., the body's opposition to an alternating current. Impedance consists of two components, resistance (R) and reactance (Xc)⁽⁶⁾. The R is derived from the amount of intracellular and extracellular water (ECW), while Xc depends on the integrity of cell membranes. Body fat and lean body mass are calculated using R and Xc in a regression equation created from a reference method, such as isotope dilution or dual-energy X-ray absorptiometry (DXA), used to measure body composition in healthy individuals without fluid imbalance, body shape abnormalities, or apparent disease^(7,8). Therefore, BIA can provide reliable information on body composition in healthy individuals but may not be accurate in patients with diseases, as they may have altered levels of intracellular or ECW, abnormal body weight, or other dysfunctions⁽⁹⁾. Consequently, body composition assessed by BIA in patients with acute or chronic disease, systemic edema, malnutrition, or pathological obesity should be evaluated with caution.

Specifically, the Phase Angle (PhA) is a useful indicator of cell membrane integrity, the distribution of water between intracellular and extracellular spaces, and the prediction of body cell mass, as they are described by the components of electrical impedance (Z), R, (function of the volume of intracellular and extracellular fluid) and Xc (function of the dielectric material of tissue cells). The PhA is geometrically calculated from R and Xc measured at 50 kHz.

The PhA can be simply calculated as an arctangent using the raw data of R and Xc at a frequency of 50 kHz, as follows: $(Xc/R) \times 180^\circ/\pi$. Thus, the PhA is obtained directly from the BIA without using a regression equation⁽⁷⁾. Previous studies have shown that PhA can be an indicator of cellular health and that a higher value reflects greater cellularity, integrity of cell membranes, and better cellular function^(10,11).

In healthy adults, the main determinants of the phase angle appear to be age, gender, and BMI⁽¹²⁾. According to some studies, the phase angle decreases with increasing age due to a reduction in X_c because of muscle mass loss. Additionally, men have a higher PhA than women due to their greater body muscle mass. The PhA increases with an increase in BMI due to the change in the amount of muscle mass and fat. A recent systematic review showed that PhA is associated with total body protein, muscle mass, and grip strength and suggested that PhA is a useful marker of muscle mass and function⁽¹³⁾.

Evidence indicates that malnutrition, inflammation, and oxidative stress impair the electrical tissue properties, leading to a lower PhA^(5,14). PhA has a prognostic impact in various disease scenarios. Compared to healthy individuals, a low PhA has been shown to be predictive of decreased survival in various chronic diseases and acute stress conditions. Furthermore, as muscle strength is related to intracellular water (ICW) and ICW with PhA, it has been found that PhA is directly associated with muscle strength from childhood to old age and in various conditions^(14,15).

The objectives of the present study were to assess the body composition of individuals aged 18 to 80 years through BIA and to characterize the studied population, according to different age groups. Additionally, we aimed to evaluate the degree of correlation between the variables analyzed by BIA, identify determinant variables of body composition that may influence or determine the behavior of others using complex statistical methods, and identify possible correlations among the analyzed variables, and identify one or more that may influence the cellular PhA.

METHODS

A total of 199 body composition records obtained through BIA (Tera Science, SJCampos - SP, Brazil) from September to December 2023 were analyzed. The study adhered to Resolution 466/2012 of the National Health Council and was approved by the Research Ethics Committee (CAAE: 63398622.8.0000.5076; N^o. 5.736.112). All participants were recruited through a public call in an extension project currently underway at the university and were provided with a consent form detailing the experimental procedures, potential risks, and discomforts involved in the study. This information was also thoroughly explained verbally. After selection, acceptance, and understanding of the physical exercise's inherent risks, all participants signed the consent form.

This study aimed to explore the complex relationships between various body composition metrics, including but not limited to BMI, PhA, Lean Body Mass, and skeletal muscle mass. Initial exploratory data analysis involved calculating Pearson correlation coefficients to identify linear relationships between variables. Subsequently, recognizing the limitations of linear models in capturing the nuanced interactions among the studied variables, we employed more sophisticated machine learning techniques, specifically Random Forest regression, to delve deeper into the predictive relationships and variable importance.

Data Preparation

The dataset comprised measurements of body composition metrics from a cohort of individuals. Each entry included demographic information (age and height), alongside specific body composition measurements such as BMI, PhA, lean body mass, skeletal muscle mass, among others. Preliminary data cleaning involved handling missing values by imputing them with the mean of their respective variables to maintain the integrity and distribution of the dataset.

Exploratory Data Analysis

The exploratory data analysis EDA phase began with calculating Pearson's correlation coefficients between all pairs of variables, visualized through a heat map. This step aimed to identify initial linear relationships and guide the subsequent modeling efforts.

Linear Regression Analysis

A simple linear regression model was first employed to understand the predictive power of the selected independent variables on the PhA, considered a crucial indicator of cellular health and nutritional status. The dataset was split into training (80%) and testing (20%) subsets to validate the model's predictive performance. However, the linear model's limited capability to capture complex relationships was evident through low R-squared values, indicating a poor fit.

Random Forest Regression

Given the limitations observed with linear regression, we advanced to a Random Forest regression model. Random Forest, an ensemble learning method that utilizes multiple decision trees to improve prediction accuracy and control over-fitting, was deemed suitable for our data's complexity. The model was trained on the same training subset, with hyperparameters tuned to optimize performance based on cross-validation scores within the training data. The Random Forest model's performance was evaluated through the R-squared metric on the testing subset, demonstrating a significant improvement in predictive accuracy compared to the linear model. Furthermore, the feature importance generated by the Random Forest model provided insights into which variables most significantly influenced the PhA, offering a deeper understanding of the underlying relationships within the body composition data.

Cross-Validation

To ensure the robustness and generalizability of our Random Forest model, we employed k-fold cross-validation with k set to 5. This method further validated the model's effectiveness across different subsets of the data, providing a comprehensive view of its predictive stability and accuracy. The methodology employed in this study illustrates the importance of utilizing advanced statistical and machine learning techniques to uncover the intricate relationships in body composition data. While initial linear analysis offered foundational insights, the application of Random Forest regression unveiled deeper predictive relationships and variable importance, contributing valuable knowledge towards the understanding of body composition metrics and their impact on health.

RESULTS

In this study, we studied the variables of 199 participants in the research project. Figure 01 shows the histogram of age frequencies of the subjects assessed at the Health Technologies Laboratory.

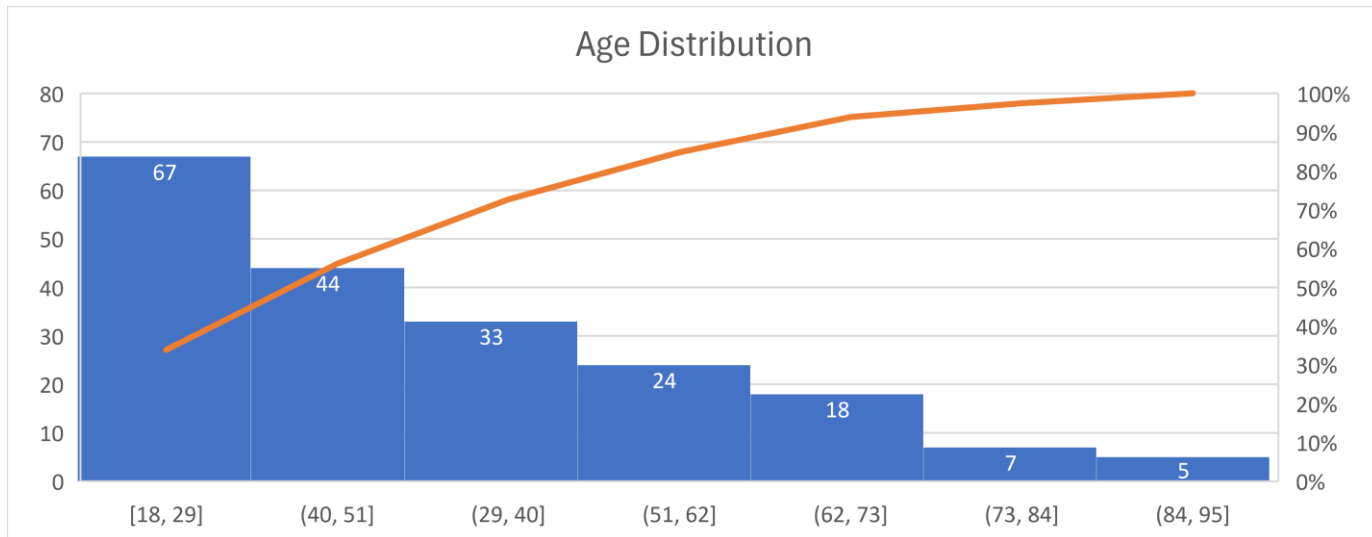


Figure 01. Age distribution of the 199 subjects.

Note: The numbers inside the columns indicate the number of participants for each age interval

1st Phase Analysis

The body composition data sheet contains various measurements and indexes crucial for understanding an individual's physical and health condition. Table 01 displays the utilized data, which includes the following analyzed variables and linear correlations one to one of all data.

Linear Regression demonstrated that there is a strong positive correlation between measures related to mass (such as Lean Mass, Skeletal Muscle Mass, ICW, Total Body Water), indicating that, as expected, individuals with more lean mass tend to have a higher volume of body water and greater skeletal muscle mass. Figure 02 demonstrate the strong positive correlation between Muscle mass (Kg) and the contents of ICW, with a $R^2 = 0.9317$.

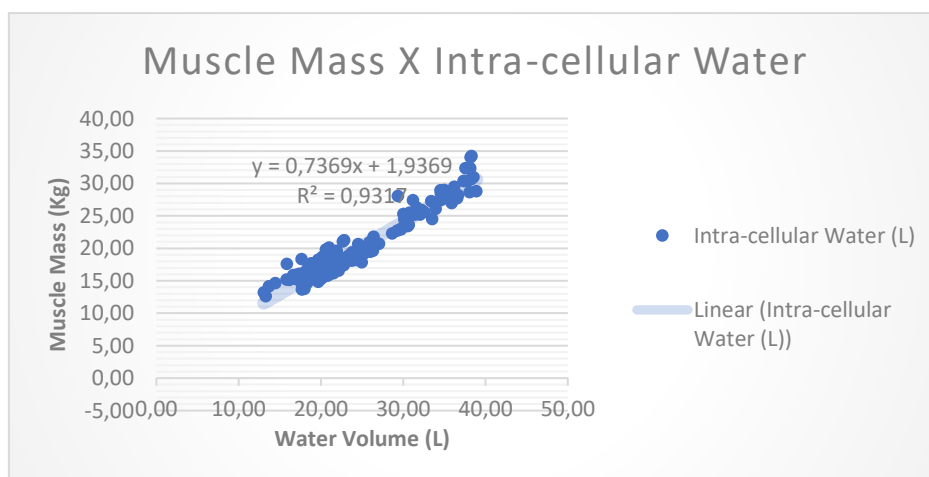


Figure 02. Linear Regression of Muscle mass and Intra-cellular Water of 199 subjects

Table 01. Correlation of all data

	Age	Height	BMI	Phase Angle	Lean Mass	Muscle Mass	Muscle Mass Index	Extra cellular Water	Intra cellular Water	Total Water	Water in Lean mass	Fat weight	%fat	Basal Matabolic Index	muscle/fat ratio
Age	1	-0,3	0,3	-0,48	-0,07	-0,29	-0,26	0,04	-0,09	-0,04	0,17	0,25	0,27	-0,03	-0,28
Height	-0,3	1	-0,18	0,37	0,77	0,78	0,49	0,6	0,74	0,72	0,04	-0,23	-0,56	0,72	0,54
BMI	0,3	-0,18	1	0,06	0,36	0,15	0,32	0,58	0,24	0,39	0,55	0,9	0,68	0,47	-0,45
Phase Angle	-0,48	0,37	0,06	1	0,49	0,6	0,62	0,33	0,57	0,5	0,38	-0,11	-0,3	0,46	0,34
Lean Mass	-0,07	0,77	0,36	0,49	1	0,95	0,85	0,93	0,98	1	0,56	0,11	-0,34	0,99	0,42
Muscle Mass	-0,29	0,78	0,15	0,6	0,95	1	0,92	0,81	0,97	0,94	0,54	-0,12	-0,54	0,91	0,6
Muscle Mass Index	-0,26	0,49	0,32	0,62	0,85	0,92	1	0,77	0,89	0,88	0,73	-0,03	-0,41	0,83	0,5
Extra-celular Water	0,04	0,6	0,58	0,33	0,93	0,81	0,77	1	0,85	0,94	0,66	0,41	-0,24	0,95	0,2
Intra-celular Water	-0,09	0,74	0,24	0,57	0,98	0,97	0,89	0,85	1	0,98	0,58	0,01	-0,54	0,95	0,51
Total Water	-0,04	0,72	0,39	0,5	1	0,94	0,88	0,94	0,98	1	0,63	0,12	-0,33	0,99	0,41
Water in Lean mass	0,17	0,04	0,55	0,38	0,56	0,54	0,73	0,66	0,58	0,63	1	0,22	-0,04	0,58	0,14
Fat weight	0,25	-0,23	0,9	-0,11	0,11	-0,12	-0,03	0,41	0,01	0,12	0,22	1	0,65	0,24	-0,68
%fat	0,27	-0,56	0,68	-0,3	-0,34	-0,54	-0,41	-0,24	-0,54	-0,33	-0,04	0,65	1	-0,22	-0,87
Basal Matabolic Index	-0,03	0,72	0,47	0,46	0,99	0,91	0,83	0,95	0,95	0,99	0,58	0,24	-0,22	1	0,32
muscle/fat ratio	-0,28	0,54	-0,45	0,34	0,42	0,6	0,5	0,2	0,51	0,41	0,14	-0,68	-0,88	0,32	1

Note: Different colors indicate the values from weak to strong correlations, positive or negative. All data combinations, one to one, are demonstrated in the table

The PhA shows varied correlations with other metrics, suggesting that this measure has a complex relationship with body composition that may not be directly linear or varies significantly between individuals. The Basal Metabolic Rate also shows interesting correlations, especially with mass measures, which makes sense since larger bodies require more energy to maintain basic functions as demonstrated in Figure 03.

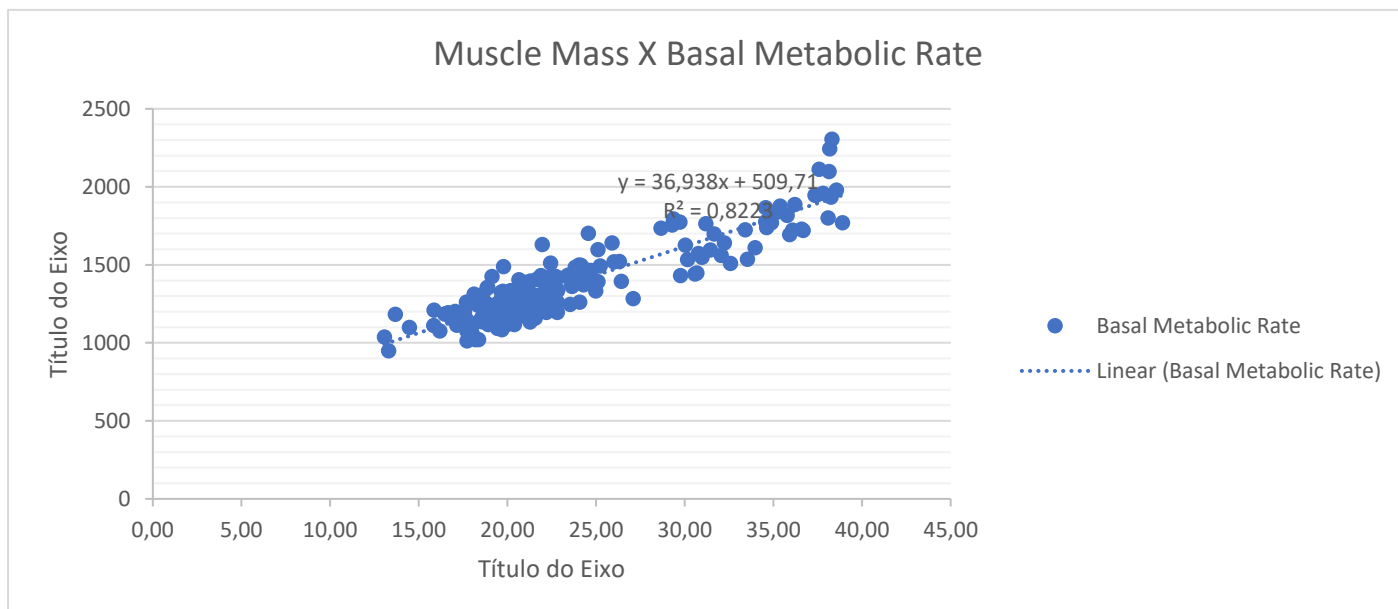


Figure 03. Linear Regression of Muscle mass and Basal Metabolic Rate of 199 subjects

The BMI has some correlations with fat and mass measures, but not as strong as one might expect, highlighting BMI's limitation in differentiating between fat mass and lean mass. These correlations can help understand how different aspects of body composition are interconnected and how changes in one area (like lean mass gain) can affect others (such as an increase in total body water). However, it is important to remember that correlations do not imply causality, and individuals can vary significantly.

2nd Phase Analysis

The Random Forest model demonstrated a significantly better ability to predict the PhA compared to simple linear models, with an R^2 of 0.52 on the test set. Cross-validation reinforced the model's robustness, indicating consistent performance across different data partitions. The use of the Random Forest model proved to be an effective approach to analyze the complex relationship between the PhA and body composition variables, highlighting its utility in contexts where non-linear relationships and complex interactions are present.

The coefficient of determination (R^2) for the test set improved to 0.52, indicating that the model can now explain approximately 52% of the variance in the PhA based on the provided variables. This is a significant improvement compared to the poor fit observed with the linear model. Cross-validation, which helps to assess the model's ability to generalize to new data, resulted in an average R^2 of 0.41 with a standard deviation of 0.14. This suggests that the model performs reasonably consistently across different subsets of the data.

DISCUSSION

In a brief, our analysis highlights the complexity of the relationships between different body composition variables and the value of using advanced models to understand these relationships. This may have significant practical implications for nutrition, sports medicine, and biomedical research, offering insights into how different components of body composition interact and affect overall health.

The data evaluation indicate that the Random Forest model is significantly more effective at capturing the complex relationships between the PhA and other body composition variables than the simple linear regression model. This demonstrates the importance of choosing the correct model for the nature of the data and the type of problem being solved. From this analysis using a Random Forest model (Random Forest Regressor), we can conclude several important things about the relationship between the Phase Angle and other body composition variables:

Complexity of Relationships

The relationships between the PhA and other body composition variables are complex and non-linear. The simple linear regression model was not able to adequately capture these relationships, as evidenced by the poor fit. However, the Random Forest model, which can capture non-linear relationships and interactions between variables, performed significantly better.

Importance of Variables

Although the Random Forest model does not provide simple regression coefficients like a linear model, it allows for the evaluation of variable importance. This means that some variables, such as Total Body Water, % Fat, Lean Mass, and others, likely have a significant impact on the PhA. These variables reflect important aspects of body composition that are associated with cellular integrity and nutritional status, which is consistent with what the PhA represents.

Predictive Capacity

The Random Forest model demonstrated a reasonable ability to predict the PhA based on the provided variables, with an R^2 of approximately 0.52 for the test set. This indicates that, while there is still a significant amount of unexplained variance, the model manages to capture a substantial portion of the relationship between the PhA and body composition variables.

Generalization

Cross-validation showed that the model has a reasonably good ability to generalize to new data, with an average R^2 of 0.41 and a standard deviation of 0.14. This suggests that the model, while not perfect, is robust across different data subsets. This analysis had practical implications that reinforce the importance of considering multiple aspects of body composition when assessing nutritional status and health. The PhA, being influenced by various dimensions of body composition, can serve as a useful indicator of overall health and cellular integrity.

Limitations

The main limitation of this study is convenience sampling. In this case, the age frequency histogram shows that around 50% of the individuals analyzed were aged between 18 and 40, including students and university staff. However, this is the first study to carry out a complex analysis involving BIA (Tera Science, SJC Campos - SP, Brazil), and to establish possible correlations between variables that could explain or indicate a relationship with the cell phase angle.

CONCLUSION

Based on the analysis provided, several key conclusions can be drawn regarding the relationships between the PhA and other body composition variables. Our findings underscore the intricate and non-linear nature of the relationships between the PhA and other body composition variables. Traditional linear regression models proved insufficient in capturing these complexities, highlighting the necessity of employing advanced modeling techniques such as Random Forest. While the Random Forest model does not yield straightforward regression coefficients, it emphasizes the significance of variables such as Total Body Water, % Fat, Lean Mass, among others, in influencing the Phase Angle. These variables reflect crucial aspects of body composition associated with cellular integrity and nutritional status, aligning with the physiological significance of the PhA. Despite the presence of unexplained variance, the Random Forest model demonstrates reasonable predictive capability for the PhA, as evidenced by an R^2 of approximately 0.52 for the test set. This suggests the model's effectiveness in capturing a substantial portion of the relationship between the PhA and body composition variables.

These findings underscore the importance of comprehensive body composition assessment in nutritional and health evaluations. By incorporating various dimensions of body composition, including the PhA, practitioners can gain valuable insights into overall health and cellular integrity, thereby informing targeted interventions and monitoring strategies.

It's important to acknowledge the study's limitations, notably the convenience sampling method employed. The predominance of individuals aged between 18 and 40, including students and university staff, may introduce bias and limit the generalizability of the findings. However, this study represents a pioneering effort in employing advanced bioimpedance analysis techniques to elucidate complex relationships between body composition variables, laying the groundwork for future research in this area. In summary, this analysis underscores the importance of employing advanced modeling techniques to unravel the complexities of body composition relationships, offering valuable insights with significant implications for nutrition, sports medicine, and biomedical research.

Author Contributions: A.C.Q., G.M.G., P.R.F., P.M.J.P., N.B.S., O.S.A.J., L.G., P.S.L.L.M., M.E.N.S., participate in data collection; R.A.B.L.M. was responsible for data analysis; and P.S.L. was responsible for study coordination.

Financial Support: The authors thank TheraScience for supporting this study.

Conflict of Interest: The authors declare no conflict of interests for this manuscript.

REFERENCES

- 1- Cederholm T, Barazzoni R, Austin P, Ballmer P, Biolo G, Bischoff SC, et al. ESPEN guidelines on definitions and terminology of clinical nutrition. *Clin Nutr*. 2017;36(1):49–64.
- 2- Lukaski HC, Vega Diaz N, Talluri A, Nescolarde L. Classification of Hydration in Clinical Conditions: Indirect and Direct Approaches Using Bioimpedance. *Nutrients*. 2019;11(4):809.
- 3- Cederholm T, Jensen GL, Correia MITD, Gonzalez MC, Fukushima R, Higashiguchi T, et al. GLIM criteria for the diagnosis of malnutrition - A consensus report from the global clinical nutrition community. *Clin Nutr*. 2019;38(1):1–9.
- 4- Prado CM, Purcell SA, Alish C, Pereira SL, Deutz NE, Heyland DK, et al. Implications of low muscle mass across the continuum of care: a narrative review. *Ann Med*. 2018;50(8):675–93.
- 5- Johnson Stoklossa CA, Sharma AM, Forhan M, Siervo M, Padwal RS, Prado CM. Prevalence of Sarcopenic Obesity in Adults with Class II/III Obesity Using Different Diagnostic Criteria. *J Nutr Metab*. 2017;2017:7307618.
- 6- Baumgartner RN, Chumlea WC, Roche AF (1988) Bioelectrical impedance phase angle and body composition. *Am J Clin Nutr* 48: 16–23.
- 7- Dey DK, Bosaeus I (2003) Comparison of bioelectrical impedance prediction equations for fat-free mass in a population-based sample of 75 y olds: the NORA study. *Nutrition* 19: 858–864.
- 8- Genton L, Karsegard VL, Kyle UG, Hans DB, Michel JP, et al. (2001) Comparison of four bioelectrical impedance analysis formulas in healthy elderly subjects. *Gerontology* 47: 315–323.
- 9- Norman K, Stobäus N, Pirlich M, Bosy-Westphal A (2012) Bioelectrical phase angle and impedance vector analysis—clinical relevance and applicability of impedance parameters. *Clin Nutr* 31: 854–861.
- 10- Kyle UG, Bosaeus I, De Lorenzo AD, Deurenberg P, Elia M, et al. (2004) Bioelectrical impedance analysis— part I: review of principles and methods. *Clin Nutr* 23: 1226–1243.
- 11- Bosy-Westphal A, Danielzik S, Dörhöfer RP, Later W, Wiese S, et al. (2006) Phase angle from bioelectrical impedance analysis: population reference values by age, sex, and body mass index. *J Parenter Enteral Nutr* 30: 309–316.
- 12- Dittmar M (2003) Reliability and variability of bioimpedance measures in normal adults: effect of age, gender, and body mass. *Am J Phys Anthropol* 122: 361–370.
- 13- Martins PC, de Lima TR, Silva AM, Silva DAS (2022) Association of phase angle with muscle strength and aerobic fitness in different populations: a systematic review. *Nutrition* 93: 111489.
- 14- Garlini LM, Alves FD, Ceretta LB, Perry IS, Souza GC, Clausell NO. Phase angle and mortality: a systematic review. *Eur J Clin Nutr*. 2019;73(4):495–508.
- 15- Custódio Martins P, de Lima TR, Silva AM, Santos Silva DA. Association of phase angle with muscle strength and aerobic fitness in different populations: A systematic review. *Nutrition*. 2022;93:111489.