

RAE - CEA - 09P04

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE
O PROJETO: ANÁLISE MULTICOMPARTIMENTAL
DA COMPOSIÇÃO CORPORAL EM JOVENS
ESPORTISTAS E NÃO ESPORTISTAS**

Profª. Drª. Silvia Nagib Elian

Carolina Miguel Lopes

Tatiana Martorano Bona

ESTATÍSTICA APLICADA – CEA
RELATÓRIO DE ANÁLISE ESTATÍSTICA – CÓDIGO 09P04

TÍTULO: Relatório de análise estatística sobre o projeto “Análise multicompartimental da composição corporal em jovens esportistas e não esportistas”.

PESQUISADOR: Dalmo Roberto Lopes Machado.

ORIENTADOR: Prof. Dr. Valdir José Barbanti.

INSTITUIÇÃO: Escola de Educação Física e Esporte – USP

FINALIDADE: Tese de doutorado.

RESPONSÁVEIS PELA ANÁLISE: Prof^a Dr^a Silvia Nagib Elian
Carolina Miguel Lopes
Tatiana Martorano Bona

REFERÊNCIA DESTE TRABALHO:

ELIAN, S. N., LOPES, C. M. e BONA, T. M. **Relatório de análise estatística sobre o projeto: Análise multicompartimental da composição corporal em jovens esportistas e não esportistas.** São Paulo, IME-USP, 2009 (RAE - CEA – 09P04).

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

BALLOR, D. L. (1996). Exercise training and body composition changes. In: **Human Body Composition**, [TG Lohman, editor]. Champaign, IL: Human Kinetics Publishers, p.287–304.

BARROSO, L. P. e ARTES, R.(2003). **Análise Multivariada**. Minicurso apresentado no 10º Simpósio de Estatística Aplicada à Experimentação Agronômica e 48ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria, Lavras, MG, 151p.

BEDRICK, E. J.; TSAI, C. (1994). Model Selection for Multivariate Regression in Small Samples. In: **Biometrics**, v.50, p.226-231.

BUSSAB, W. O. e MORETTIN, P. A. (2005). **Estatística Básica**. 5.ed. São Paulo: Atual. 526p.

DRAPER, N. R. & SMITH, H. (1981). **Applied Regression Analysis**. 2.ed. New York: John Wiley. 709p.

HEALY, M. J. R.; LOVAIC, J. A.; MANDEL, S. P. H. et al. (1981). The individual and the group. In.: WEINER, J. S.; LONNIE, J. A. (eds.) **Practical Human Biology**. New York, Academic Press, p.11-23.

JOHNSON, R. A. e WICHERN, D. W. (2007). **Applied Multivariate Statistical Analysis**, 6.ed. New York: Prentice Hall. 773p.

KIM, J.; WANG, Z.; HEYMSFIELD, S. B.; BAUMGARTNER, R. N. e GALLAGHER, D. (2002). Total-body skeletal muscle mass: estimation by a new dual-energy X-ray

absorptiometry method. **The American Journal of Clinical Nutrition**, v.76, n.2, Aug, p.378-83.

MIRWALD, R. L.; BAXTER-JONES, A. D. G.; BAILEY, D. A. e BEUNEN, G. P. (2002) An assessment of maturity from anthropometric measurements. **Medicine Science of Sports and Exercise**, v.34, n.4, p.689-694.

MONTGOMERY, D. C.; PECK, E. A. e VINING, G. G. (2001) **Introduction to Linear Regression Analysis**. 3.ed. New York: John Wiley. 641p.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Excel *for Windows* ® , versão 2003.

Microsoft Word *for Windows* ® , versão 2003.

MINITAB *for Windows* ® , versão 15.0.

TÉCNICAS ESTATÍSTICAS UTILIZADAS: (entre parênteses a classificação ISI)

Análise Descritiva Unidimensional (03:010)

Análise Descritiva Multidimensional (03:020)

Análise de Regressão Clássica (07:020)

Estimação Paramétrica Multidimensional (04:160)

Testes de Hipóteses Paramétricas (05:010)

ÁREA DE APLICAÇÃO

Ciências Biomédicas (14:990)

ÍNDICE

Resumo	6
1. Introdução	7
2. Objetivos	7
3. Descrição do Estudo	8
4. Descrição das Variáveis	11
5. Análise Descritiva Unidimensional	13
6. Análise Descritiva Multidimensional	13
6.1. Relações entre as Variáveis Dependentes	13
6.2. Relações entre as Variáveis Dependentes e as Variáveis Independentes	14
6.3. Relações entre as Variáveis Independentes	16
7. Análise Inferencial	16
7.1. Modelo de Regressão Linear Multivariada	16
7.2. Seleção de Variáveis Independentes	17
7.3. Análise do Primeiro Modelo	19
7.4. Análise do Segundo Modelo	21
7.5. Análise do Terceiro Modelo	23
7.6. Comparação entre os Modelos	24
7.7. Intervalo de Predição para Novas Observações	25
8. Considerações Finais	26
Apêndice A - Tabelas	28
Apêndice B - Gráficos	37
Apêndice C - Matriz dos Coeficiente de Correlação Linear de Pearson	48
Apêndice D - Multicolinearidade	54
Apêndice E - Análise de Resíduos	57

RESUMO

A análise da composição corporal tem se mostrado uma ferramenta de grande importância na melhoria das condições de saúde e de aptidão física. Atualmente, o sofisticado método de absorptometria de raios-X de dupla energia (DXA) prevê as componentes corporais, via técnica multicompartimental, de forma precisa. Porém, seu uso é limitado e restrito.

Este trabalho tem como objetivo analisar as componentes da composição corporal (massa de gordura, massa muscular e massa óssea) com base na análise multicompartimental, considerando idade, maturação e prática esportiva; e propor equações preditivas para essas componentes a partir de variáveis antropométricas. Para isto, foram utilizados quatrocentos e oito indivíduos do sexo masculino com idade entre oito e dezoito anos classificados em esportistas e não esportistas.

A análise descritiva indicou que só parece haver diferenças entre esportistas e não esportistas para a variável massa de gordura, sendo essa a única que não se modifica consideravelmente com a idade e a maturação. Além disso, as componentes corporais são altamente correlacionadas entre si e com as variáveis explicativas, justificando a utilização de um modelo de regressão multivariada. Notou-se também que as variáveis explicativas possuem alta correlação entre si, gerando o problema da multicolinearidade.

Selecionadas as variáveis explicativas, ajustaram-se três modelos de regressão linear multivariada. Todos satisfaziam às suposições básicas, o grau de multicolinearidade entre as variáveis explicativas era aceitável e tais variáveis explicativas apresentaram contribuição estatisticamente significativa.

1. Introdução

No ambiente escolar, clínico e esportivo, a avaliação da composição corporal (CC) em crianças e adolescentes constitui uma importante ferramenta em programas relacionados à melhoria da aptidão física e das condições de saúde, bem como na prevenção e tratamento de diversas doenças como o sobrepeso e obesidade. Atualmente, os métodos disponíveis para prever a composição corporal (CC), além de específicos para determinadas populações, em geral assumem modelos bicompartimentais a partir da densidade corporal (massa magra e massa gorda), limitando as generalizações e a precisão na previsão dos componentes corporais (massa de gordura corporal, massa óssea e massa muscular).

Uma análise mais precisa da dinâmica (interdependência) da composição corporal (CC) com a maturação na adolescência e o envolvimento em práticas esportivas pode fornecer valiosas informações de aspecto educacional, esportivo e de saúde, fundamentando intervenções e orientações que venham a favorecer tanto o rendimento esportivo como a promoção da saúde.

Os métodos mais sofisticados e considerados mais precisos para quantificar os componentes corporais, como a absorptometria de raios-X de dupla energia (DXA), possuem uso limitado e precisão dos softwares desconhecida, além de utilização restrita a laboratórios e a situações clínicas muito específicas.

O objetivo do presente estudo é prever a composição corporal via análise multicompartimental, que fraciona o peso corporal ($\text{Peso Corporal} = \text{Massa de Gordura} + \text{Massa Muscular} + \text{Massa Óssea} + \text{Resíduo}$), considerando a associação com a maturação e práticas esportivas, a partir de variáveis antropométricas, em crianças e adolescentes do sexo masculino com idades entre 8 e 18 anos.

2. Objetivos

Os objetivos deste trabalho são:

- Analisar os componentes da composição corporal, considerando maturação, idade e prática esportiva, com base na análise multicompartimental da composição corporal;
- Propor equações preditivas para massa de gordura (MG), massa muscular (MM) e massa óssea (MO) a partir das variáveis antropométricas e avaliar a qualidade dos modelos obtidos.

3. Descrição do estudo

Propuseram-se a participar voluntariamente do estudo 408 jovens do sexo masculino recrutados em centros esportivos, escolas públicas e particulares de Presidente Prudente/SP. Os jovens deveriam ser aparentemente saudáveis, sem partes do corpo amputadas, sem uso declarado de medicamentos ou algum tratamento que pudesse influenciar em seu metabolismo, apetite ou crescimento e sem restrição médica.

Tais jovens deveriam ter entre 8 e 18 anos, de acordo com a chamada idade milesimal, que não é simplesmente a idade cronológica do indivíduo. O cálculo da idade milesimal, de acordo com o método desenvolvido por Healy (Healy et al. (1981)), baseia-se na obtenção de frações do ano equivalentes à data de avaliação (F1) e a data de nascimento do avaliado (F2), através das operações:

$$F1 = \frac{(\text{dia da avaliação} - 1) + 30,3(\text{mês da avaliação} - 1)}{365}$$

e

$$F2 = \frac{(\text{dia de nascimento} - 1) + 30,3(\text{dia da avaliação} - 1)}{365}.$$

A partir desses valores foi obtida a fração da idade (F), com base nas operações $(F1 - F2)$, quando $F1 > F2$; ou $(1 - F2 + F1)$, quando $F1 < F2$. Essa fração leva em conta se a data de avaliação ocorreu antes ou depois do aniversário do indivíduo, variando entre

zero e um. Quanto mais próximo de zero, significa que o indivíduo fez aniversário recentemente e quanto mais próximo de um, mais perto está do aniversário seguinte.

O cálculo de anos foi obtido pelas operações (Ano da Avaliação – Ano de Nascimento, se $F1 > F2$; ou (Ano de avaliação – Ano de Nascimento) – 1, quando $F1 < F2$. A idade milesimal foi obtida pela soma da quantidade de anos e a fração idade (F). Uma vez determinada a idade milesimal, o intervalo para cada idade foi determinado para aqueles sujeitos que estivessem entre -0,500 a 0,499 do valor inteiro de cada idade. Por exemplo, aqueles que tivessem idade milesimal entre 9,500 e 10,499 anos, eram classificados com idade de 10 anos. Adotou-se esse cálculo de idade, uma vez que as crianças modificam-se fisicamente de um ano para o outro; por exemplo, um menino com 15 anos e 11 meses possui características biológicas de um menino de 16 anos e não mais de 15. Assim, a partir desse cálculo, ele seria considerado com 16 anos ao invés de 15.

Os meninos foram agrupados em esportistas (ES; $n=181$) e não esportistas (NE; $n=227$). Os esportistas eram aqueles envolvidos em programas de práticas esportivas no mínimo há dois anos, com pelo menos seis horas semanais de treinamento e com experiência em competições no último ano na modalidade praticada. As modalidades escolhidas foram futebol de campo, futsal, judô e atletismo, porém, após a seleção, tais características das modalidades não foram levadas em consideração. Os jovens não esportistas eram aqueles não engajados em nenhum programa de atividade esportiva.

Os indivíduos foram classificados por níveis de maturação, baseados no pico de velocidade de crescimento (PVC) proposto por Mirwald et al. (2002). O conceito de maturação tenta associar o tempo biológico ao tempo cronológico, pois apesar de duas crianças terem a mesma idade cronológica, elas poderão estar com diferentes níveis de maturidade biológica.

A literatura sugere indicadores para o pico de velocidade de crescimento para meninos e para meninas. O pico de velocidade de crescimento dos meninos, utilizado no estudo, segue a fórmula:

$$PVC = -9,236 + 0,0002708 (CP \times TC) - 0,001663 (IxCP) + 0,007216 (IxTC) + 0,02292 (P/E),$$

em que: CP = comprimento de perna (em cm), TC = altura tronco-cefálica (em cm), I = idade (em anos), P = peso (em kg), E = estatura (em cm).

Esse modelo para estimativa do pico de velocidade do crescimento fornece valores contínuos. Assim, esse resultado será utilizado como variável independente na proposição dos modelos preditivos das componentes corporais. Já para efeito descritivo, foi adotado um agrupamento etário do PVC, através da transformação para o valor inteiro mais próximo, resultando em nove níveis definidos na forma de escala: -5, -4, -3, -2, -1, 0, 1, 2 e 3 anos para atingir o pico de velocidade de crescimento. Dessa forma, a análise descritiva de gráficos torna-se mais clara. O nível de maturação "0" indica que o indivíduo estava em seu pico máximo de crescimento. O nível 3, por exemplo, indica que o pico de crescimento foi atingido há 3 anos, e o nível -5 indica que faltam 5 anos para atingi-lo. É importante ressaltar que a idade em que ocorre o pico de crescimento varia entre indivíduos, estando a idade média em torno de 14 anos para indivíduos do sexo masculino. Foi coletada uma amostra de tal forma que todos os níveis de maturação fossem contemplados.

A Absortometria de raios-X de dupla energia (DXA) é uma técnica multicompartimental e foi utilizada para medir a composição corporal (CC) regional e total, não só pela baixa exposição radiológica e pela praticidade e rapidez na estimativa dos multicomponentes corporais, mas por permitir melhores interpretações da composição corporal (CC) da criança e adolescente durante o crescimento. Para o seu funcionamento, o indivíduo se deita sobre a mesa do scanner e tem seus membros inferiores atados com fitas de velcro. Os braços ficam ao longo do corpo com as mãos voltadas para baixo. Um braço mecânico desliza por sobre o corpo e, utilizando a energia transmitida de dois fótons que incidem através dos ossos e tecidos moles, faz uma varredura do corpo para a obtenção das componentes.

Os exames de Absortometria de Raios X de Dupla Energia (DXA) de corpo total e a coleta das medidas antropométricas de cada avaliado foram realizados em uma única sessão e pelos mesmos avaliadores em laboratório com temperatura constante controlada (23°C), de outubro a dezembro de 2008. Os voluntários eram trazidos em grupos de 3 a 5 sujeitos por sessão de medidas.

O estudo possui algumas limitações: a) não foram controlados fatores nutricionais, sociais, econômicos ou hereditários; b) não foram controladas a intensidade e periodicidade das práticas de treinamento e c) estimativas de referência da composição corporal pela DXA se limitavam ao equipamento e softwares disponíveis.

4. Descrição das variáveis

As variáveis do estudo estão divididas em preditivas e resposta. Como preditivas ou independentes temos as variáveis de caracterização, de dimensão corporal, a espessura das dobras cutâneas, os perímetros e os diâmetros ósseos. Como variáveis resposta ou dependentes, temos os componentes da composição corporal.

▪ Caracterização

- Idade (em anos).
- Categoria (esportista/não esportista).
- Pico de velocidade do crescimento (em anos).

▪ Dimensões corporais

- Peso (em kg).
- Estatura (em cm).
- Altura tronco-cefálica (em cm).
- Espessura das dobras cutâneas (em mm): bicipital, tricipital, subescapular, peitoral, axilar média, supra-ilíaca, abdominal vertical, abdominal horizontal, coxa medial e panturrilha.
- Perímetros (em cm): tórax meso-external, braço-relaxado, braço-contraído, antebraço, punho, cintura, abdome, quadril, coxa e perna-medial.
- Diâmetros ósseos (em cm): biacromial, bicrista-ilíaco, tórax transverso, biepicondilar do úmero, bitrocantérico, biestiloidal de rádio e ulna, biepicondilar do fêmur e bimaleolar.

- **Composição corporal** (obtida pela DXA)

- Massa de gordura (em kg).
- Massa muscular (em kg).
- Massa mineral óssea (em kg).

Todas as medidas antropométricas foram tomadas seguindo as padronizações da Internacional Society for Advancement in Kinanthropometry –ISAK. O Peso corporal foi medido com uma balança eletrônica com precisão de 100 gramas. A estatura foi medida com um estadiômetro de alumínio fixo de parede com escala de precisão 0,1cm. A altura tronco-cefálica foi obtida pelo mesmo estadiômetro a partir da estatura do indivíduo sentado em um banco de 50 cm.

A espessura das dobras cutâneas foram medidas em milímetros através de um plicômetro. Os perímetros, medidos em centímetros, foram obtidos por uma fita metálica de 2 m. Os diâmetros, por sua vez, foram medidos utilizando paquímetros de alumínio anodizado com precisão de 0,01 m e tolerância de 0,01 mm.

O exame de composição corporal foi realizado utilizando a DXA, com o procedimento descrito anteriormente. A partir dos valores obtidos, foi possível mensurar as componentes da composição corporal. A massa de gordura (MG) foi obtida diretamente pela DXA. A massa muscular (MM) foi obtida a partir de um cálculo que leva em consideração o tecido magro apendicular (TMA) dos braços e pernas obtidos pela DXA e o peso corporal (PC), como descrito a seguir:

$$MM \text{ (kg)} = 1,003.TMA + 0,039.PC - 1,315.$$

A massa óssea (MO), por sua vez, também foi obtida através de um cálculo que leva em conta o conteúdo mineral ósseo (CMO), obtido pela DXA:

$$MO \text{ (kg)} = 1,0436.CMO .$$

A equação de predição da massa muscular e da massa óssea foram desenvolvidas e validadas para crianças e adolescentes, respectivamente por Kim et al. (2006) e Ballor (1996).

5. Análise descritiva unidimensional

Esta seção apresenta uma análise exploratória unidimensional das variáveis resposta em estudo.

Inicialmente, pelo Gráfico B1, observa-se como se distribui, em média, o peso corporal dos meninos de 8 a 18 anos analisados: desse peso corporal, 45% correspondem à massa muscular, 21% à massa de gordura, 5% à massa óssea e 29% a resíduos (vísceras, água, sangue, etc).

Os Gráficos B2, B3 e B4 representam os *box-plots* (Bussab e Morettin, 2005) das variáveis massa de gordura, massa muscular e massa óssea, respectivamente. As medidas descritivas das três variáveis resposta encontram-se na Tabela A1. A partir do Gráfico B2 é possível observar que a massa de gordura possui acentuada assimetria, resultando em vários pontos distantes da mediana, enquanto as massas muscular (Gráfico B3) e óssea (Gráfico B4) parecem ser simétricas. Para a variável massa de gordura, os pontos distantes da mediana são de meninos de no mínimo 12 anos, sugerindo que com o aumento da idade pode aumentar também a variabilidade dessa variável. Dessa forma, seria mais conveniente se basear na mediana (6,62) do que na média (9,34) como medida de tendência central do grupo, uma vez que a média é mais influenciada por esses pontos.

Observando o coeficiente de variação da massa de gordura (80,14) nota-se que essa variável possui a maior dispersão relativa à média se comparada com as variáveis massa muscular (36,99) e massa óssea (37,14). Essas duas, por sua vez, possuem coeficientes de variação próximos. Porém, ao analisar o desvio padrão, observa-se que o da massa óssea é bem baixo (0,79), até porque ela varia apenas de 0,70 a 4,90 kg nesta amostra e, como visto no Gráfico B1, esta quantidade constitui apenas 5% do peso total.

6. Análise descritiva multidimensional

6.1. Relações entre as variáveis dependentes

A partir da matriz de correlações lineares de Pearson (Apêndice C), é possível verificar que as variáveis massas muscular e massa óssea são altamente correlacionadas (0,947), o que justifica o uso de um modelo de regressão multivariado, considerando as três variáveis resposta simultaneamente. A variável massa de gordura não possui correlação tão alta com as variáveis massa muscular (0,314) e massa óssea (0,406), mas ainda assim o uso de um modelo de regressão multivariado é oportuno.

6.2. Relações entre as variáveis dependentes e as variáveis independentes

Na Tabela A2 e nos Gráficos B5, B6 e B7, as componentes corporais estão sendo analisadas em função da categoria: esportistas e não esportistas. Observa-se que apenas a variável massa de gordura (Gráfico B5) tem distribuição assimétrica para as duas categorias, sugerindo que a mediana (ES = 5,98; NE = 7,76) pode ser uma medida descritiva mais representativa do que a média (ES = 7,56; NE = 10,75).

O resultado do coeficiente de variação da variável massa de gordura para a categoria dos esportistas (75,25) é próximo ao da categoria dos não esportistas (78,10), porém se excluíssemos os dois esportistas com valor de massa de gordura mais discrepante — um menino de 13 anos e outro de 17 — o coeficiente de variação da variável massa de gordura para essa categoria cairia consideravelmente (61,65), mostrando como esses indivíduos, considerados como discrepantes em relação à massa de gordura, impactam nesse coeficiente.

Nota-se que, para a categoria dos esportistas, a variável massa de gordura parece apresentar valores mais baixos em comparação com a categoria dos não esportistas (Gráfico B5). Esse resultado era esperado, uma vez que, quando um indivíduo realiza atividades esportivas, ele tende a perder massa de gordura.

Em relação as variáveis massa muscular (Gráfico B6) e massa óssea (Gráfico B7), não há indícios de grandes diferenças entre as categorias. O mesmo pode ser verificado pelas medidas descritivas que apresentam valores muito próximos (Tabela A2).

As variáveis resposta são altamente correlacionadas com as explicativas, conforme pode ser observado na matriz de correlação do Apêndice C. A variável massa

de gordura apresenta alta correlação principalmente com as variáveis referentes à espessura das dobras cutâneas e com os perímetros (0,77 em média). As variáveis massa muscular e óssea, uma vez que são altamente correlacionadas entre si (0,947), possuem alta correlação com as mesmas variáveis explicativas, principalmente com idade, PVC, as variáveis de dimensão corporal, perímetros e diâmetros (correlação de 0,836 e 0,845, em média, para respectivamente massa muscular e massa óssea com essas variáveis citadas). Quanto mais correlacionadas forem as variáveis resposta com as explicativas, mais justificado está o uso do modelo de regressão na futura análise inferencial.

Quando analisada em função da idade, a massa de gordura apresentou um comportamento de oscilação mais acentuado a partir dos 13 anos de idade (Gráfico B8). Entretanto, quando avaliada por níveis de maturação (Gráfico B9), observa-se uma tendência de comportamento mais uniforme. Os Gráficos B8 e B9 sugerem que a variável massa de gordura aumenta em média conforme aumenta a idade e os níveis de maturação. Porém, analisando os gráficos de dispersão (Gráficos B14 e B15), nota-se uma grande variabilidade da massa de gordura em cada idade (principalmente a partir dos 13 anos) e em cada nível de maturação. Além disso, os coeficientes de correlação obtidos, respectivamente, 0,279 e 0,222, sugerem inicialmente a inexistência de relação entre massa de gordura e nível de maturação (0,279) e entre massa de gordura e idade (0,222).

Por outro lado, a massa muscular média quando analisada por idade (Gráfico B10), apresenta um comportamento bastante similar ao obtido por nível de maturação (Gráfico B11). Os gráficos de dispersão dão indícios da existência de relação entre a massa muscular e idade (Gráfico B16) e massa muscular e nível de maturação (Gráfico B17). Analisando os coeficientes de correlação, nota-se que essa variável parece estar um pouco mais correlacionada com o PVC (0,926) do que com a idade (0,876).

O mesmo ocorre com a massa óssea. Seu comportamento é semelhante tanto quando avaliada em função da idade (Gráfico B12), como quando observada em função do nível de maturação (Gráfico B13). Os gráficos de dispersão também sugerem a existência de relação da massa óssea com a idade (Gráfico B18) e com o nível de

maturação (Gráfico B19), o que também pode ser observado pelos altos coeficientes de correlação entre massa óssea e idade (0,869) e massa óssea e PVC (0,924).

6.3. Relações entre as variáveis independentes

As variáveis explicativas também são altamente correlacionadas entre si (ver Apêndice C). Dentro de cada grupo de variáveis (dobras cutâneas, perímetros e diâmetros) a correlação é muito alta (0,843, 0,880 e 0,777 em média, respectivamente). O PVC, idade e as variáveis de dimensão corporal também apresentam forte relação linear entre si (coeficientes de correlação, em média, de 0,893).

A alta correlação entre as variáveis independentes evidencia a existência de relações lineares entre elas. Tal fato, conhecido como multicolinearidade (Apêndice D), traz efeitos indesejáveis para a análise inferencial. Segundo Hair et al. (2005), além dos efeitos na explicação, a multicolinearidade pode ocasionar sérios efeitos nas estimativas dos coeficientes de regressão e na aplicabilidade geral do modelo estimado.

A principal fonte de multicolinearidade neste projeto vem das relações naturais existentes entre as variáveis independentes. Por exemplo, as variáveis espessura das dobras cutâneas e diâmetros ósseos são todas medidas do tecido adiposo subcutâneo e estão altamente relacionadas.

Utilizando o método descrito no apêndice D para verificar a existência de multicolinearidade entre as variáveis independentes do estudo, obteve-se um resultado que indicou alta multicolinearidade ($L \approx 6441$).

7. Análise inferencial

7.1. Modelo de regressão linear multivariada

Com objetivo de relacionar as variáveis resposta — massa de gordura, massa muscular e massa óssea — com as variáveis independentes descritas na Seção 4, foram utilizados modelos de regressão linear multivariada (Johnson e Wichern, 2007), da forma:

$$\mathbf{Y}_{(n \times m)} = \mathbf{Z}_{(n \times (r+1))} \boldsymbol{\beta}_{((r+1) \times m)} + \boldsymbol{\varepsilon}_{(n \times m)},$$

em que:

\mathbf{Y} é a matriz na qual cada coluna contém os valores de cada variável resposta;

\mathbf{Z} é a matriz de variáveis explicativas;

$\boldsymbol{\beta}$ é a matriz na qual cada coluna contém os parâmetros de regressão associados a cada variável resposta;

n é o tamanho da amostra, $n=408$;

m é o número de variáveis resposta, $m=3$ e

r o número de variáveis explicativas.

Supõe-se que os vetores aleatórios $\boldsymbol{\varepsilon}_{(i)}$, $i=1, 2, 3$, correspondentes às colunas de $\boldsymbol{\varepsilon}$, possuem distribuição normal multivariada com vetor de médias zero e matriz de covariância $\sigma_{ii} \mathbf{I}_n$, sendo \mathbf{I}_n a matriz de identidade de ordem n . Admite-se, ainda, que as observações medidas em um mesmo indivíduo são correlacionadas com distribuição normal multivariada enquanto que observações obtidas em indivíduos distintos não o são. A validade destas suposições será analisada posteriormente.

7.2. Seleção de variáveis independentes

Para a escolha das variáveis independentes utilizadas no ajuste dos modelos de regressão, foi proposta, inicialmente, a técnica de análise de componentes principais (Barroso e Artes, 2003), aplicada a cada grupo de variáveis explicativas: espessura das dobras cutâneas, perímetros e diâmetros ósseos. Tal técnica permitiu a redução da multicolinearidade entre as variáveis independentes, porém não era viável do ponto de vista da praticidade do modelo, uma vez que leva em conta todas as variáveis explicativas do estudo.

Optou-se, então, pela seleção das variáveis independentes através do procedimento *stepwise* (Montgomery et al., 2001), com o auxílio do aplicativo MINITAB. Tal procedimento consiste em remover e acrescentar variáveis no modelo de regressão com a finalidade de identificar um conjunto de variáveis preditoras úteis. Para tal propósito, foi executado um *stepwise* para cada variável resposta, com um nível de

significância de 5%, em que todas as variáveis explicativas poderiam ser inseridas no modelo. Os resultados obtidos estão apresentados na Tabela A3.

Com o auxílio do pesquisador, foram selecionadas as variáveis explicativas que tivessem sido apontadas pelo procedimento *stepwise* em pelo menos duas das variáveis resposta, resultando, assim, um total de doze variáveis explicativas, apresentadas na Tabela A4.

Para auxiliar na seleção das variáveis, foram obtidos os fatores de inflação da variância (FIV) para cada uma delas. Ao contrário da medida de multicolinearidade L (Apêndice D), que avalia a multicolinearidade global das variáveis explicativas; o FIV identifica quais variáveis oferecem maior contribuição para aumentar a multicolinearidade. Seu cálculo consiste em ajustar um modelo de regressão para cada variável independente, onde esta se torna a variável resposta das demais explicativas. É então calculado o coeficiente de explicação (R_x^2) desse modelo, a partir do qual se obtém o fator de inflação da variância da seguinte forma: $FIV = \frac{1}{1-R_x^2}$.

É conhecido na literatura que:

$$FIV = \begin{cases} 1, & \text{as variáveis explicativas são não correlacionadas;} \\ 1 \text{ a } 5, & \text{as variáveis explicativas são pouco correlacionadas;} \\ 5 \text{ a } 10, & \text{multicolinearidade moderada e} \\ >10, & \text{muticolinearidade prejudicial ao modelo.} \end{cases}$$

Há autores que consideram admissíveis fatores de inflação da variância até cinco e outros até dez. Nesse projeto, valores até dez foram considerados aceitáveis, o que significa que o coeficiente de explicação R_x^2 pode atingir até 90%. Na Tabela A5 encontram-se os valores dos FIV referentes às doze variáveis explicativas selecionadas preliminarmente, e estas apresentaram média multicolinearidade ($L \approx 330,47$).

Dentre as doze variáveis explicativas selecionadas até então, estão Dobra Cutânea Abdominal Vertical e Dobra Cutânea Abdominal Horizontal (Tabela A4). Como ambas medem a mesma região do corpo humano, optou-se, em conjunto com o

pesquisador, pela exclusão da variável Dobra Cutânea Abdominal Vertical, por possuir o maior FIV dentre as duas.

Ainda com o objetivo de reduzir a multicolinearidade, verificou-se que as variáveis PVC, Estatura e Peso também estão dentre as doze variáveis independentes pré-selecionadas. Porém, como o cálculo do PVC já inclui as variáveis Estatura e Peso, decidiu-se pela exclusão destas, mantendo apenas a variável PVC.

Apesar da variável Perímetro do Antebraço ser considerada significativa pelo procedimento *stepwise*; a pedido do pesquisador, ela foi eliminada do modelo, por não apresentar tanta importância prática. Em contrapartida, o pesquisador optou por inserir a variável Dobra Cutânea Tricipital em seu lugar, pois apesar de só ser significativa para a variável resposta massa de gordura, é útil do ponto de vista prático.

Os fatores de inflação da variância referentes às nove variáveis restantes encontram-se na Tabela A6. A medida de multicolinearidade L resultou em 95,9, indicando baixa multicolinearidade. Porém observou-se que a variável Diâmetro Bitrocantérico possui FIV acima de 10 (Tabela A6), justificando sua exclusão.

Após o término do procedimento, foram selecionadas, então, oito variáveis explicativas: PVC, Dobra Cutânea Coxa Medial, Dobra Cutânea Supra-Iliaca, Dobra Cutânea Peitoral, Dobra Cutânea Tricipital, Dobra Cutânea Abdominal Horizontal, Perímetro da Coxa e Diâmetro Bicrista-Iliaco.

Na Tabela A7 encontram-se os fatores de inflação da variância das oito variáveis independentes selecionadas e a medida de multicolinearidade para estas indicou baixa multicolinearidade ($L \approx 67,5$).

A partir desses resultados foram ajustados três modelos. A seguir, apresenta-se uma análise completa de cada modelo selecionado, obtida com o auxílio do aplicativo MINITAB.

7.3. Análise do Primeiro Modelo

O primeiro modelo se constituiu no modelo de regressão multivariada das três variáveis resposta em função das oito variáveis explicativas então selecionadas (vide Seção 7.2).

O modelo ajustado foi:

$$\widehat{MG} = -17,4 - 0,113 \text{ PVC} + 0,110 \text{ Dc_Tri} + 0,0197 \text{ Dc_Pt} + 0,206 \text{ Dc_Si} + 0,200 \text{ Dc_AbH} + \\ + 0,0978 \text{ Dc_Cx} + 0,213 \text{ Per_Cx} + 0,321 \text{ Dia_BiCri},$$

$$\widehat{MM} = -6,57 + 1,19 \text{ PVC} - 0,0633 \text{ Dc_Tri} - 0,107 \text{ Dc_Pt} + 0,0213 \text{ Dc_Si} - 0,0091 \text{ Dc_AbH} - \\ - 0,126 \text{ Dc_Cx} + 0,516 \text{ Per_Cx} + 0,370 \text{ Dia_BiCri} \text{ e}$$

$$\widehat{MO} = -0,271 + 0,163 \text{ PVC} - 0,00107 \text{ Dc_Tri} - 0,0114 \text{ Dc_Pt} + 0,0031 \text{ Dc_Si} - \\ - 0,00072 \text{ Dc_AbH} - 0,00543 \text{ Dc_Cx} + 0,0472 \text{ Per_Cx} + 0,029 \text{ Dia_BiCri},$$

sendo que o significado da abreviatura de cada variável explicativa encontra-se no Apêndice C.

Com o objetivo de verificar se cada variável que compõe o primeiro modelo era significativa, foi realizado o teste da razão de verossimilhança. Dessa forma, testava-se a hipótese de que as variáveis resposta não dependiam de cada variável explicativa:

$$H_0 : \beta_{(2)} = \mathbf{0}$$

contra

$$H_1 : \beta_{(2)} \neq \mathbf{0},$$

em que $\beta = \begin{bmatrix} \beta_{(1)} \\ ((q+1) \times m) \\ \dots \\ \beta_{(2)} \\ ((r-q) \times m) \end{bmatrix}.$

A estatística do teste $\lambda = -[n - r - 1 - 0,5(m - r + q + 1)] \ln \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_1|} \right)$, devido ao grande tamanho da amostra, é aproximada por uma distribuição qui-quadrado com $m(r-q)$ graus de liberdade. No caso, $\beta_{(2)}$ é o vetor contendo os parâmetros da variável explicativa a ser testada nas três regressões, de modo que $m=3$, $q=7$ e $r=8$. Além disso, $\hat{\Sigma}$ é a estimativa da matriz de covariância das medidas associadas ao mesmo indivíduo no modelo completo, enquanto que $\hat{\Sigma}_1$ é a mesma estimativa no modelo sem a variável explicativa que está sendo testada.

Verificou-se, então, que cada uma das variáveis é significativa para o modelo, dado que as demais já foram incluídas ($p < 0,020$), a um nível de significância de 5%. A única exceção é a variável Dobra Cutânea Tricipital ($p = 0,061$), porém, devido a sua importância prática, optou-se por mantê-la no primeiro modelo.

Os coeficientes de explicação univariados, obtidos no ajuste de cada uma das três equações de regressão, encontram-se na Tabela A8. Observou-se que, para todas as variáveis resposta, tal coeficiente ultrapassou 90%, indicando que as oito variáveis independentes que o compõe explicam uma parcela muito boa da variabilidade das variáveis resposta.

As suposições de homocedasticidade (homogeneidade de variâncias), bem como de normalidade dos erros foram verificadas com base em gráficos de resíduos (Gráficos E1, E4, E5 e E6). Observou-se que os resíduos padronizados distribuem-se numa faixa em torno do zero, sugerindo que não há grandes desvios da suposição de homocedasticidade dos erros do modelo ajustado. Quanto à normalidade, não há indícios de severa falha desta suposição, já que os pontos nos gráficos de “probabilidade normal” estão dispostos ao redor de uma reta, tanto quando analisados os resíduos univariadamente (Gráficos E4, E5 e E6), quanto na situação em que são analisados multivariadamente (Gráfico E1).

7.4. Análise do Segundo Modelo

Para compor um segundo modelo, partiu-se das oito variáveis explicativas já selecionadas, com a possibilidade de adição de outras, dentre cinco consideradas também de importância pelo pesquisador: Dobra Cutânea Subescapular, Perímetro Tórax Meso-Externa, Perímetro Perna Medial, Diâmetro Biepicondilar do Úmero e Diâmetro Biepicondilar do Fêmur.

Para decidir quais, dentre essas cinco variáveis, passariam a também compor o segundo modelo, foi ajustado um modelo com as oito variáveis iniciais juntamente com cada uma das cinco sugeridas e, em seguida, analisados os coeficientes de explicação e maior fator de inflação da variância associado (Tabela A9a, A9b, A9c, A9d e A9e). Esse procedimento foi realizado para cada uma das cinco variáveis propostas.

Analisando os resultados, observou-se que os valores dos FIV não aumentaram de forma significativa. Já o coeficiente de explicação, quando incluída uma das variáveis explicativas sugeridas, aumenta significativamente apenas para uma variável resposta e não para as três simultaneamente. Assim, optou-se pela inclusão de três variáveis explicativas, de forma que cada uma contribuísse para o aumento do coeficiente de explicação de uma determinada variável resposta. Tais variáveis foram: Dobra Cutânea Subescapular, Perímetro Perna Medial e Diâmetro Biepicondilar do Úmero, que contribuíam, respectivamente, para a explicação das variáveis resposta massa de gordura, massa óssea e massa muscular.

O segundo modelo ficou, inicialmente, composto por onze variáveis explicativas. Porém, após seu ajuste, o teste da razão de verossimilhança (vide Seção 7.2) detectou, a um nível de significância de 5%, que a variável Dobra Cutânea Tricipital não contribuía significativamente para a explicação das variáveis resposta ($p=0,123$), dado que as demais variáveis explicativas estavam no modelo.

Após a exclusão da variável Dobra Cutânea Tricipital, dez variáveis explicativas passam a formar o segundo modelo. O teste da razão de verossimilhança foi realizado novamente e todas as dez variáveis mostraram-se significativas ($p<0,001$). O ajuste desse modelo resultou em:

$$\widehat{MG} = -19,5 - 0,304 \text{ PVC} - 0,0433 \text{ Dc_Pt} + 0,175 \text{ Dc_Si} + 0,195 \text{ Dc_AbH} + 0,112 \text{ Dc_Cx} + \\ + 0,113 \text{ Per_Cx} + 0,335 \text{ Dia_BiCri} + 0,156 \text{ Dc_Se} + 0,279 \text{ Per_Pn} - \\ - 0,386 \text{ Dia_BiEpUm},$$

$$\widehat{MM} = -12,1 + 1,15 \text{ PVC} - 0,0958 \text{ Dc_Pt} + 0,0124 \text{ Dc_Si} - 0,0086 \text{ Dc_AbH} - 0,136 \text{ Dc_Cx} + \\ + 0,261 \text{ Per_Cx} + 0,119 \text{ Dia_BiCri} - 0,0335 \text{ Dc_Se} + 0,445 \text{ Per_Pn} + \\ + 1,42 \text{ Dia_BiEpUm} \text{ e}$$

$$\widehat{MO} = -0,658 + 0,154 \text{ PVC} - 0,0132 \text{ Dc_Pt} + 0,00166 \text{ Dc_Si} - 0,00118 \text{ Dc_AbH} - \\ - 0,00618 \text{ Dc_Cx} + 0,0258 \text{ Per_Cx} + 0,0160 \text{ Dia_BiCri} + 0,00305 \text{ Dc_Se} + \\ + 0,0458 \text{ Per_Pn} + 0,0348 \text{ Dia_BiEpUm}.$$

Na Tabela A10 são apresentados os fatores de inflação da variância relativos às dez variáveis que constituem o segundo modelo. Embora o maior valor alcançado tenha sido 11,7, decidiu-se pela manutenção desse modelo. Além disso, de acordo com o critério descrito no Apêndice D, a multicolinearidade detectada não é prejudicial ao modelo, já que $L \approx 116$.

Os coeficientes de explicação univariados (Tabela A11) resultaram pouco acima de 90%.

As suposições de homocedasticidade, bem como de normalidade dos erros foram verificadas com base em gráficos de resíduos (Gráficos E2, E7, E8 e E9), os quais sugerem que não há grandes desvios da suposição de homocedasticidade dos erros do modelo ajustado. Quanto à normalidade, também não há indícios de severa falha desta suposição, tanto na análise univariada (Gráficos E7, E8 e E9), quanto na análise multivariada (Gráfico E2).

7.5. Análise do Terceiro Modelo

Visando a praticidade do modelo, o pesquisador sugeriu que a variável independente PVC fosse excluída do conjunto das oito selecionadas inicialmente (primeiro modelo), uma vez que o cálculo do PVC envolve outras cinco variáveis. Ajustando-se o modelo com as sete variáveis explicativas restantes, foram obtidos coeficientes de explicação pouco abaixo de 90%. Numa etapa seguinte, optou-se, então, por substituir a variável PVC por uma das que a compõe, ou seja, Idade, Estatura ou Peso.

A variável explicativa Estatura foi a que apresentou resultados mais razoáveis dentre as três, no que diz respeito à contribuição significativa na explicação das variáveis resposta e fatores de inflação da variância. Os dados permitindo a comparação entre os modelos com as variáveis Idade, Estatura e Peso encontram-se nas Tabelas A12a, A12b e A12c.

O terceiro modelo passa, então, a ser constituído pelas variáveis explicativas: Estatura, Dobra Cutânea Coxa Medial, Dobra Cutânea Supra-Iliaca, Dobra Cutânea

Peitoral, Dobra Cutânea Tricipital, Dobra Cutânea Abdominal Horizontal, Perímetro da Coxa e Diâmetro Bicrista-Iliaco, com baixa multicolinearidade ($L \approx 66,26$).

O ajuste obtido foi:

$$\begin{aligned}\widehat{MG} &= -15,0 - 0,0137 \text{ Estatura} + 0,116 \text{ Dc_Tri} + 0,0206 \text{ Dc_Pt} + 0,204 \text{ Dc_Si} \\ &\quad + 0,200 \text{ Dc_AbH} + 0,0963 \text{ Dc_Cx} + 0,209 \text{ Per_Cx(m)} + 0,324 \text{ Dia_BiCri}, \\ \widehat{MM} &= -34,7 + 0,213 \text{ Estatura} - 0,0873 \text{ Dc_Tri} - 0,103 \text{ Dc_Pt} + 0,0445 \text{ Dc_Si} - \\ &\quad - 0,0134 \text{ Dc_AbH} - 0,0957 \text{ Dc_Cx} + 0,499 \text{ Per_Cx(m)} + 0,0890 \text{ Dia_BiCri} \text{ e} \\ \widehat{MO} &= -3,76 + 0,0208 \text{ Estatura} - 0,00792 \text{ Dc_Tri} - 0,0125 \text{ Dc_Pt} + 0,00586 \text{ Dc_Si} - \\ &\quad - 0,00189 \text{ Dc_AbH} - 0,00311 \text{ Dc_Cx} + 0,0528 \text{ Per_Cx(m)} + 0,0205 \text{ Dia_BiCri}.\end{aligned}$$

O teste da razão de verossimilhança não indicou, a um nível de significância de 5%, variáveis com contribuição não significativa (níveis descritivos $p < 0,009$).

Os fatores de inflação da variância relativos às variáveis que constituem o terceiro modelo são apresentados na Tabela A13 e os coeficientes de explicação univariados (Tabela A12c) resultaram pouco acima de 90%.

As suposições de homocedasticidade, bem como de normalidade dos erros foram verificadas com base em gráficos de resíduos (Gráficos E3, E10, E11 e E12), os quais sugerem que não há grandes desvios da suposição de homocedasticidade dos erros do modelo ajustado. Quanto à normalidade, também não há indícios de severa falha desta suposição, tanto na análise univariada (Gráficos E10, E11 e E12), quanto na análise multivariada (Gráfico E3).

7.6. Comparação entre os modelos

Quando se utiliza regressão como técnica de análise de dados, frequentemente determina-se vários modelos como candidatos a modelo final a ser adotado.

No caso, os três modelos ajustados se mostraram adequados com relação à validade das suposições necessárias à inferência.

A partir daí, foram utilizados métodos para comparação de modelos: critério de informação de Akaike (AIC) e PRESS.

O critério de Akaike baseia-se na função de verossimilhança do modelo e considera o melhor modelo aquele que apresentar menor valor da medida AIC que, para modelos de regressão linear multivariada, é definida como (Bedrick e Tsai, 1994):

$AIC = n(\ln|\hat{\Sigma}| + m) + 2\{m(r+1) + 0,5m(m+1)\}$, com $\hat{\Sigma} = \frac{(\mathbf{Y} - \mathbf{Z}\hat{\beta})'(\mathbf{Y} - \mathbf{Z}\hat{\beta})}{n}$ e $\hat{\beta}$ é o estimador da matriz de parâmetros β .

O critério PRESS consiste em eliminar uma observação e ajustar o modelo de regressão às observações restantes. Tal modelo ajustado é utilizado para prever o valor da observação retirada e calcular o resíduo correspondente. O procedimento é repetido para todas as observações. O melhor modelo, segundo esse critério, será aquele que apresentar a menor soma de quadrados deste resíduo e envolver o menor número de variáveis independentes (Draper e Smith, 1981). Verifica-se (Bedrick e Tsai, 1994) que, para modelos de regressão linear multivariada, essa medida é dada por:

$PRESS = \sum_{i=1}^n \sum_{j=1}^m \hat{e}_{ij}^2 / (1 - h_i)$, com \hat{e}_{ij} sendo o elemento de ordem i, j da matriz de resíduos $\mathbf{Y} - \mathbf{X}\hat{\beta}$ e h_i é o i -ésimo elemento da diagonal da matriz $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

Observa-se que, segundo esses critérios, o segundo modelo se mostra melhor que os outros dois, e o terceiro modelo se mostra melhor que o primeiro (Tabela A14).

7.7. Intervalo de predição para novas observações

Definido o modelo a ser utilizado, caso haja interesse, a predição por intervalo de uma nova observação (Johnson e Wichern, 2007), é da forma:

$$\mathbf{z}_0' \hat{\beta}_{(i)} \pm \sqrt{\left(\frac{m(n-r-1)}{n-r-m} \right) F_{m, n-r-m}(\alpha)} \sqrt{\left(1 + \mathbf{z}_0' (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{z}_0 \right) \left(\frac{n}{n-r-1} \hat{\sigma}_{ii} \right)},$$

em que:

i = índice da variável resposta, $i = 1, 2, \dots, m$;

$\hat{\beta}_{(i)}$ = i -ésima coluna da matriz de estimativas dos coeficientes de regressão;

$\hat{\sigma}_{ii}$ = i -ésimo elemento da diagonal da matriz $\hat{\Sigma}$ do modelo;

n = tamanho da amostra e

r = número de variáveis independentes.

$F_{m,n-r-m}(\alpha)$ é o quantil de ordem $1 - \alpha$ da distribuição F de Fisher e Snedecor com m graus de liberdade no numerador e $n - r - m$ graus de liberdade no denominador.

8. Considerações Parciais

Descritivamente não parece haver diferenças entre esportistas e não esportistas segundo as variáveis massa muscular e massa óssea, enquanto que para a massa de gordura há indícios de uma pequena diferença.

A massa de gordura não varia muito com a idade quando vista descritivamente, ao contrário da massa muscular e da massa óssea.

As variáveis resposta são altamente correlacionadas entre si, o que justifica o uso de uma regressão multivariada, e altamente correlacionadas com as variáveis explicativas, justificando a utilização de um modelo de regressão.

As variáveis explicativas possuem alta correlação entre si, gerando o problema da multicolinearidade. Uma sugestão para contorná-lo seria o uso das componentes principais. Porém essa técnica leva em conta todas as variáveis do estudo, o que seria inviável do ponto de vista da praticidade do modelo, uma vez que, a literatura da área tem mostrado que com apenas duas ou três medidas de dobras cutâneas é possível prever a massa de gordura, por exemplo.

Então, uma alternativa foi a seleção das variáveis, segundo o procedimento *stepwise* e, principalmente, levando em conta o conhecimento do pesquisador e o que a literatura dessa área tem observado sobre as variáveis mais significantes para o estudo.

Foram então obtidos três modelos, com oito, dez e oito variáveis explicativas, respectivamente. Em todos, as suposições básicas foram satisfeitas e nenhum deles

apresentava variáveis explicativas com contribuição não significativa. O segundo modelo se mostrou melhor de acordo com os critérios de Akaike e PRESS, mas o terceiro modelo apresentou medida de multicolinearidade ligeiramente inferior.

Além disso, este último modelo envolve variáveis explicativas mais simples e em menor quantidade que o segundo.

APÊNDICE A

TABELAS

Tabela A1: Medidas descritivas das variáveis resposta.

Massa	n	Média	Desvio Padrão	Coef. Var.	Mínimo	1º Quartil	Mediana	3º Quartil	Máximo
Gordura	408	9,34	7,48	80,14	1,28	4,30	6,62	11,86	41,79
Muscular	408	20,08	7,43	36,99	7,56	13,23	19,92	26,71	37,83
Óssea	408	2,14	0,79	37,14	0,70	1,39	2,09	2,85	4,09

Tabela A2: Medidas descritivas das variáveis resposta observadas por categoria.

Massa	Categ	n	Média	Desvio Padrão	Coef. Var.	Mínimo	1º Quartil	Mediana	3º Quartil	Máximo
Gordura	ES	181	7,56	5,69	75,25	1,52	4,00	5,98	8,94	41,79
	NE	227	10,75	8,39	78,10	1,28	4,48	7,76	14,67	40,66
Muscular	ES	181	20,62	7,39	35,86	7,73	14,20	20,88	27,00	37,83
	NE	227	19,66	7,45	37,88	7,56	12,79	19,41	26,63	36,15
Óssea	ES	181	2,20	0,80	36,40	0,92	1,49	2,12	2,90	3,84
	NE	227	2,09	0,79	37,69	0,70	1,37	2,05	2,76	4,09

Tabela A3: Variáveis explicativas selecionadas pelo procedimento *stepwise*.

Massa de Gordura	Massa Muscular	Massa Óssea
Dobra Cutânea Abdominal Vertical	Estatura	Altura Tronco-Cefálica
Perímetro Abdome	Perímetro Antebraço	Peso
Dobra Cutânea Coxa Medial	Dobra Cutânea Peitoral	Dobra Cutânea Abdominal Horizontal
Peso	Peso	PVC
Estatura	Dobra Cutânea Coxa Medial	Perímetro Punho
Dobra Cutânea Supra-Iliaca	Perímetro Coxa	Dobra Cutânea Supra-Iliaca
PVC	Dobra Cutânea Abdominal Vertical	Diâmetro Bitrocantérico
Perímetro Cintura	Diâmetro Bicrista-Iliaco	Perímetro Antebraço
Dobra Cutânea Subescapular	Diâmetro Bitrocantérico	Estatura
Perímetro Antebraço		Dobra Cutânea Peitoral
Dobra Cutânea Abdominal Horizontal		Perímetro Coxa
Idade		
Dobra Cutânea Tricipital		

Tabela A4: Variáveis explicativas selecionadas preliminarmente.

Dobra Cutânea Abdominal Vertical	Diâmetro Bicrista-Iliaco	Diâmetro Bitrocantérico
PVC	Perímetro Antebraço	Perímetro Coxa
Dobra Cutânea Coxa Medial	Dobra Cutânea Peitoral	Dobra Cutânea Abdominal Horizontal
Peso	Estatura	Dobra Cutânea Supra-Iliaca

Tabela A5: Fator de inflação da variância das variáveis selecionadas preliminarmente.

Variável	VIF	Variável	VIF
Dobra Cutânea Abdominal Vertical	27,104	Diâmetro Bitrocantérico	14,646
PVC	10,929	Perímetro Coxa	11,254
Dobra Cutânea Coxa Medial	4,689	Dobra Cutânea Abdominal Horizontal	22,687
Peso	26,740	Dobra Cutânea Supra-Iliaca	9,681
Diâmetro Bicrista-Iliaco	8,804	Dobra Cutânea Peitoral	8,651
Perímetro Antebraço	9,072	Estatutura	12,376

Tabela A6: Fator de inflação da variância das nove variáveis restantes.

Variável	VIF	Variável	VIF
Dobra Cutânea Tricipital	7,794	Diâmetro Bitrocantérico	12,912
PVC	8,158	Perímetro Coxa	7,208
Dobra Cutânea Coxa Medial	5,782	Dobra Cutânea Abdominal Horizontal	9,744
Dobra Cutânea Peitoral	8,140	Dobra Cutânea Supra-Iliaca	8,677
Diâmetro Bicrista-Iliaco	8,273		

Tabela A7: Fator de inflação da variância das oito variáveis explicativas selecionadas (primeiro modelo).

Variável	VIF	Variável	VIF
Dobra Cutânea Tricipital	7,794	Diâmetro Bicrista-Iliaco	5,824
PVC	7,065	Perímetro Coxa	6,204
Dobra Cutânea Coxa Medial	5,779	Dobra Cutânea Abdominal Horizontal	9,744
Dobra Cutânea Peitoral	8,125	Dobra Cutânea Supra-Iliaca	8,667

Tabela A8: Coeficientes de explicação univariados do primeiro modelo.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,7\%$
Massa Muscular	$R^2 = 91,9\%$
Massa Óssea	$R^2 = 90,7\%$

Tabela A9a: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas acrescidas da variável Dobra Cutânea Subescapular.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 93,1\%$
Massa Muscular	$R^2 = 91,9\%$
Massa Óssea	$R^2 = 90,7\%$
Maior VIF	9,8

Tabela A9b: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas acrescidas da variável Perímetro Tórax Meso External.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,9\%$
Massa Muscular	$R^2 = 92,3\%$
Massa Óssea	$R^2 = 91,0\%$
Maior VIF	9,9

Tabela A9c: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas acrescidas da variável Perímetro Perna Medial.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 93,0\%$
Massa Muscular	$R^2 = 92,8\%$
Massa Óssea	$R^2 = 91,4\%$
Maior VIF	11,4

Tabela A9d: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas acrescidas da variável Diâmetro Biepicondilar do Úmero.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,7\%$
Massa Muscular	$R^2 = 92,6\%$
Massa Óssea	$R^2 = 90,8\%$
Maior VIF	9,8

Tabela A9e: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas acrescidas da variável Diâmetro Biepicondilar do Femur.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,7\%$
Massa Muscular	$R^2 = 92,2\%$
Massa Óssea	$R^2 = 90,9\%$
Maior VIF	10,0

Tabela A10: Fator de inflação da variância das dez variáveis explicativas que compõem o segundo modelo.

Variável	VIF	Variável	VIF
Perímetro Coxa	11,7	Dobra Cutânea Supra-Iliaca	9,5
Dobra Cutânea Coxa Medial	4,6	Dobra Cutânea Peitoral	9,4
PVC	7,1	Diâmetro Biepicondilar do Úmero	4,9
Diâmetro Bicrista-Iliaco	6,8	Dobra Cutânea Abdominal Horizontal	9,6
Dobra Cutânea Subescapular	7,7	Perímetro Perna-Medial	10,5

Tabela A11: Coeficientes de explicação univariados do segundo modelo.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 93,3\%$
Massa Muscular	$R^2 = 93,2\%$
Massa Óssea	$R^2 = 91,4\%$

Tabela A12a: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas inicialmente, substituindo-se a variável PVC pela variável Idade.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,7\%$
Massa Muscular	$R^2 = 90,9\%$
Massa Óssea	$R^2 = 89,1\%$
Maior VIF	9,8

Tabela A12b: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas inicialmente, substituindo-se a variável PVC pela variável Peso.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 94,2\%$
Massa Muscular	$R^2 = 95,0\%$
Massa Óssea	$R^2 = 93,2\%$
Maior VIF	16,0

Tabela A12c: Coeficientes de explicação univariados e maior fator de inflação da variância do modelo ajustado com as oito variáveis selecionadas inicialmente, substituindo-se a variável PVC pela variável Estatura.

Variável resposta	Coeficiente de explicação
Massa de Gordura	$R^2 = 92,7\%$
Massa Muscular	$R^2 = 93,5\%$
Massa Óssea	$R^2 = 90,4\%$
Maior VIF	9,7

Tabela A13: Fator de inflação da variância das oito variáveis explicativas que compõem o terceiro modelo.

Variável	VIF	Variável	VIF
Dobra Cutânea Tricipital	7,5	Diâmetro Bicrista-Iliaco	6,7
Estatura	7,2	Perímetro Coxa	5,7
Dobra Cutânea Coxa Medial	5,9	Dobra Cutânea Abdominal Horizontal	9,7
Dobra Cutânea Peitoral	8,1	Dobra Cutânea Supra-Iliaca	8,7

Tabela A14: Comparação entre os modelos segundo os critérios de Akaike e PRESS.

critério	Primeiro Modelo	Segundo Modelo	Terceiro Modelo
Akaike	1170,324	1073,678	1115,738
PRESS	3526,460	3106,891	3166,671

APÊNDICE B

GRÁFICOS

Gráfico B1: Gráfico de setores da distribuição média do peso corporal nos indivíduos da amostra.

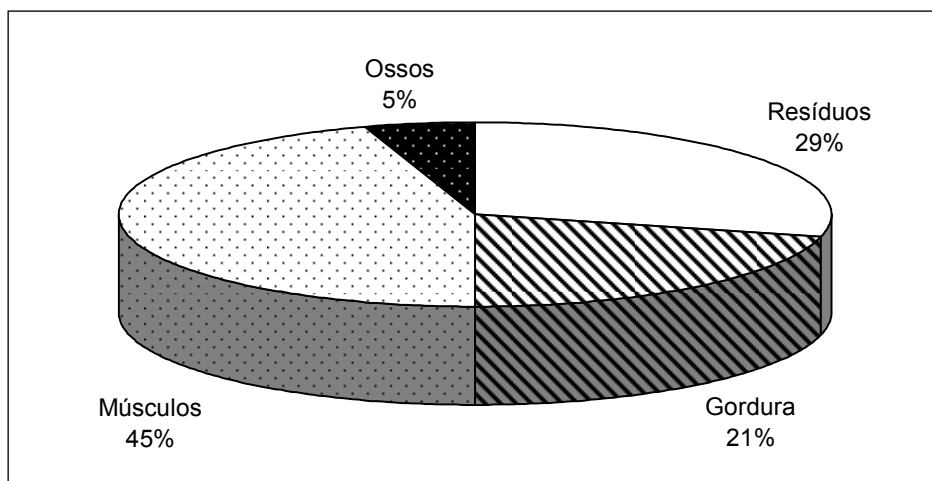


Gráfico B2: *Box-plot* da massa de gordura (MG).

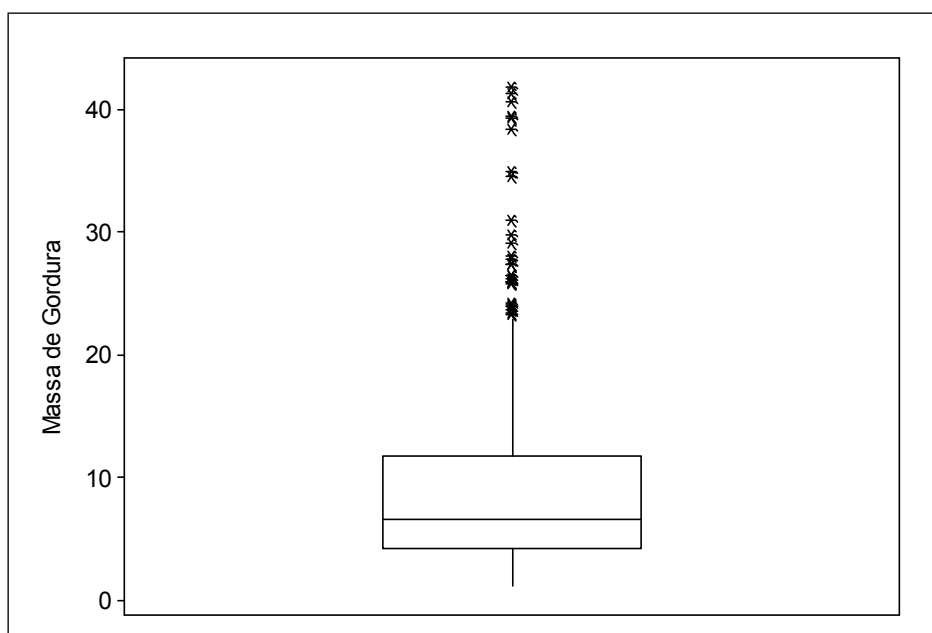


Gráfico B3: *Box-plot* da massa muscular (MM).

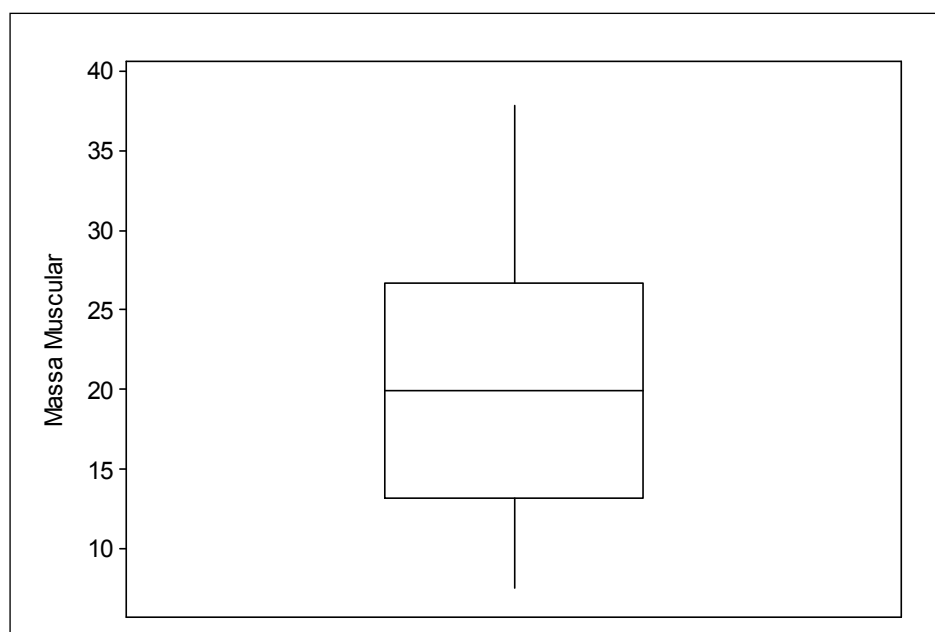


Gráfico B4: *Box-plot* da massa óssea (MO).

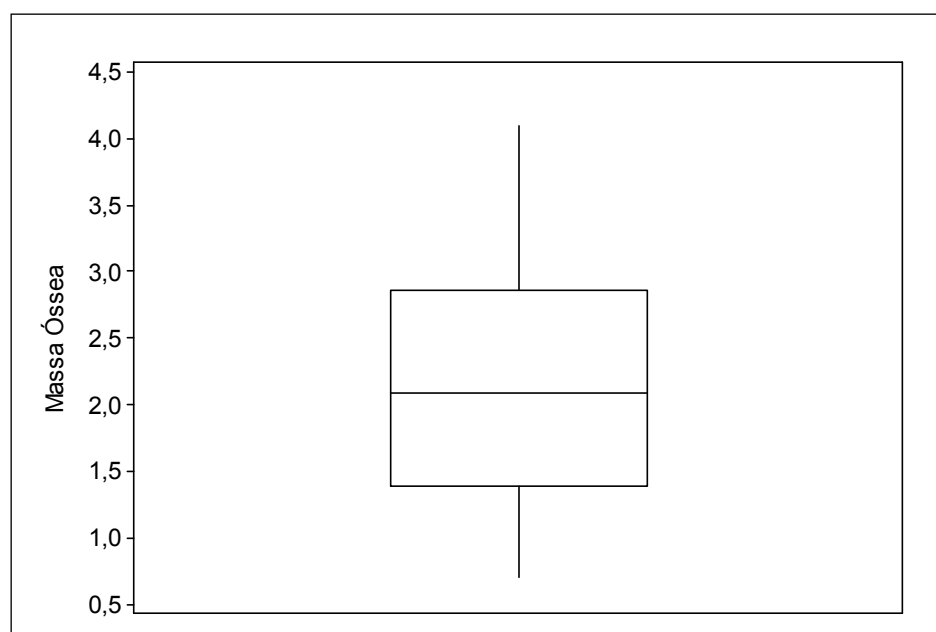


Gráfico B5: *Box-plot* da massa de gordura (MG) observada por categoria.

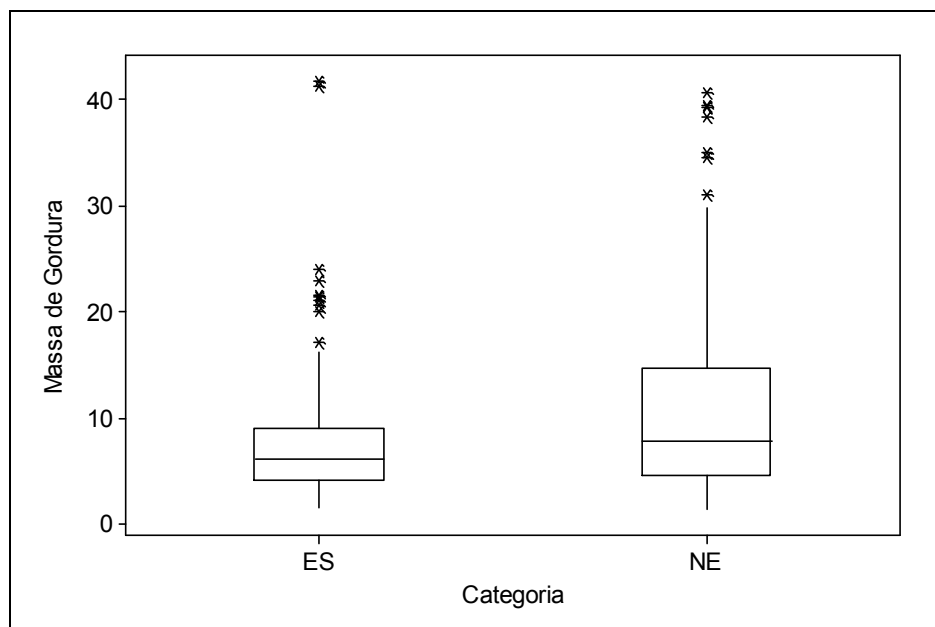


Gráfico B6: *Box-plot* da massa muscular (MM) observada por categoria.

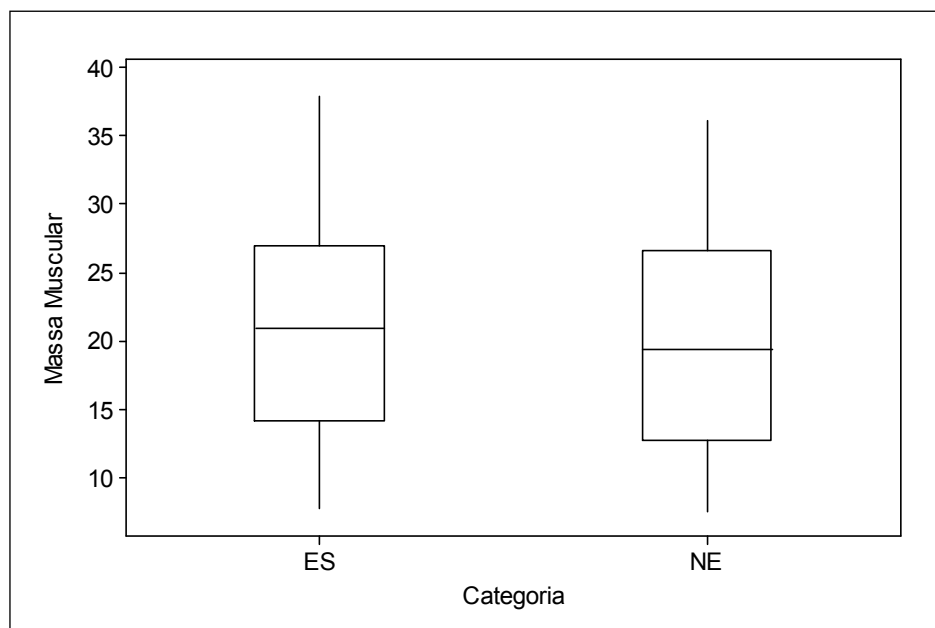


Gráfico B7: *Box-plot* da massa óssea (MO) observada por categoria.

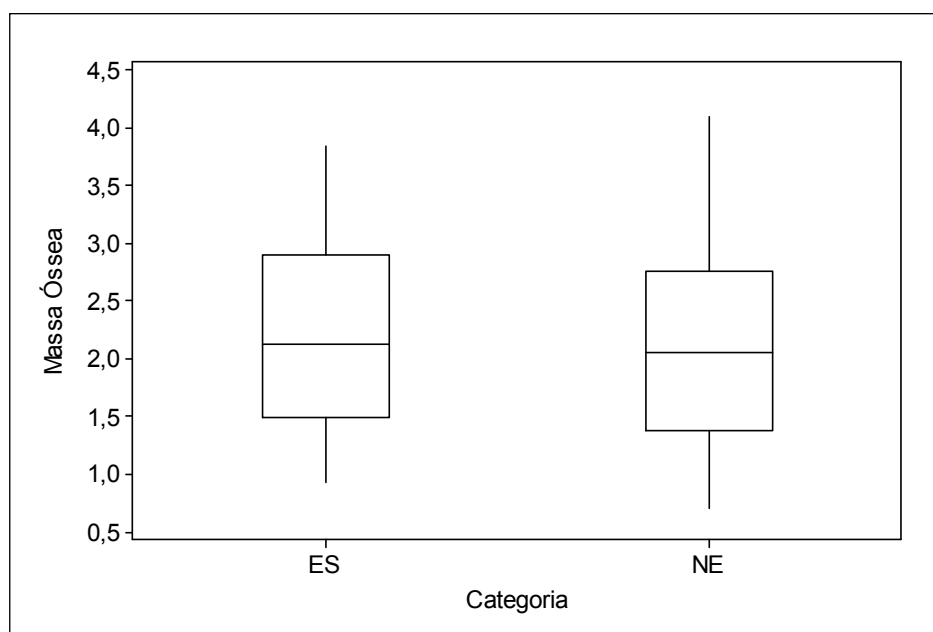


Gráfico B8: Massa de gordura média em função da idade.

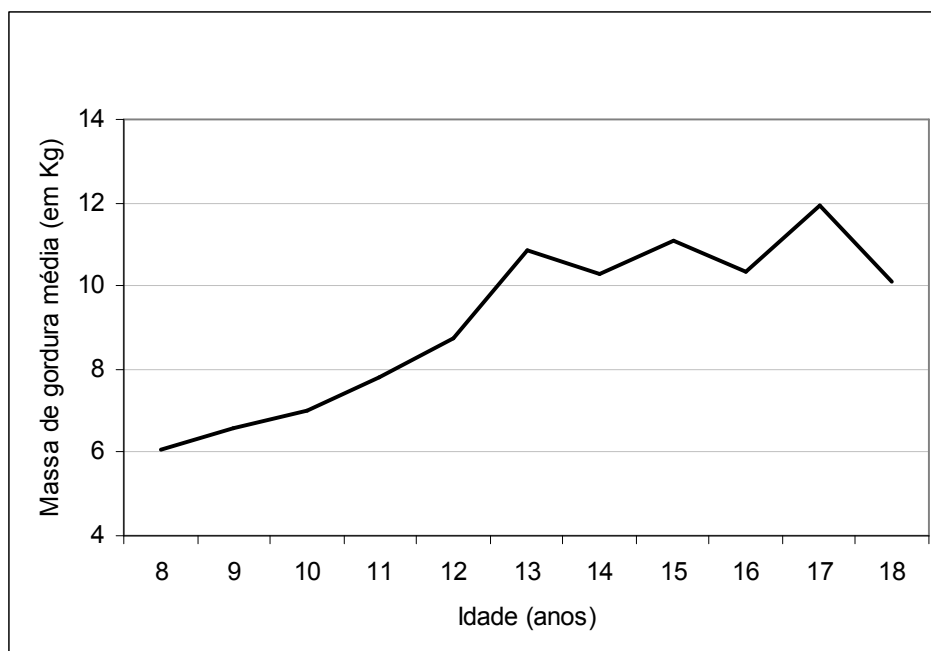


Gráfico B9: Massa de gordura média em função do nível de maturação.

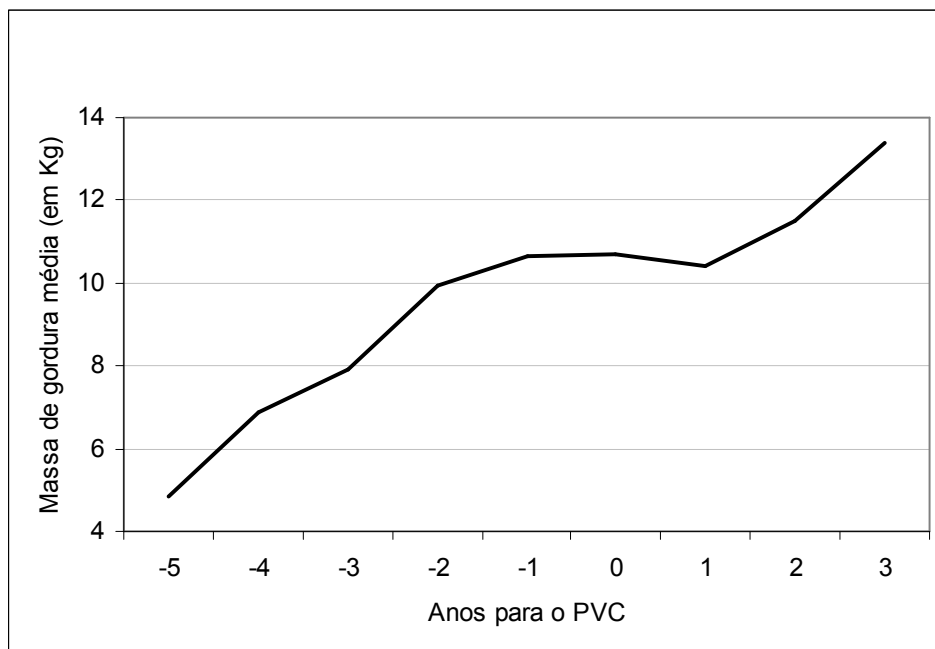


Gráfico B10: Massa muscular média em função da idade.

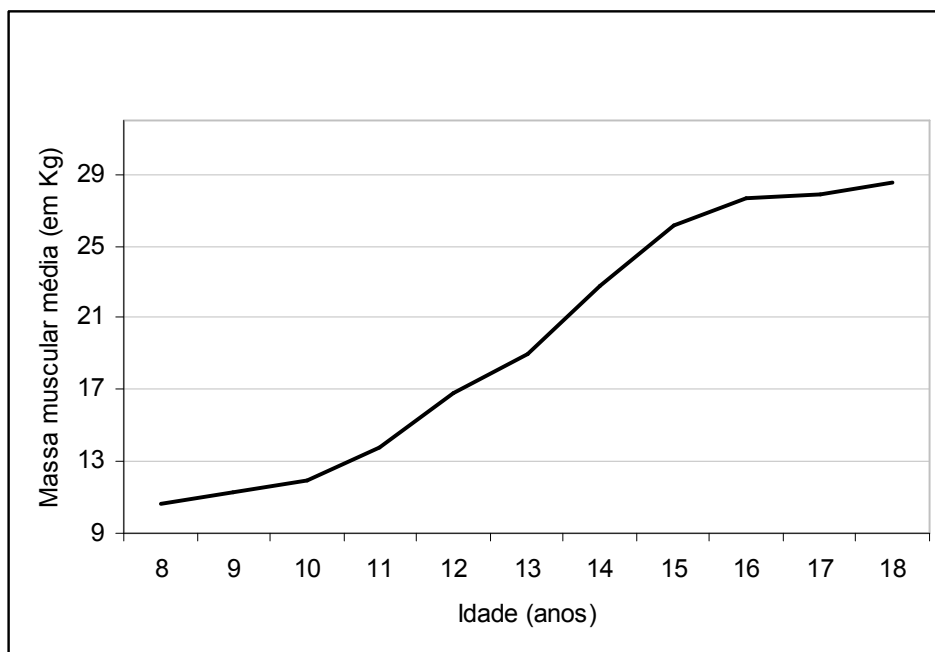


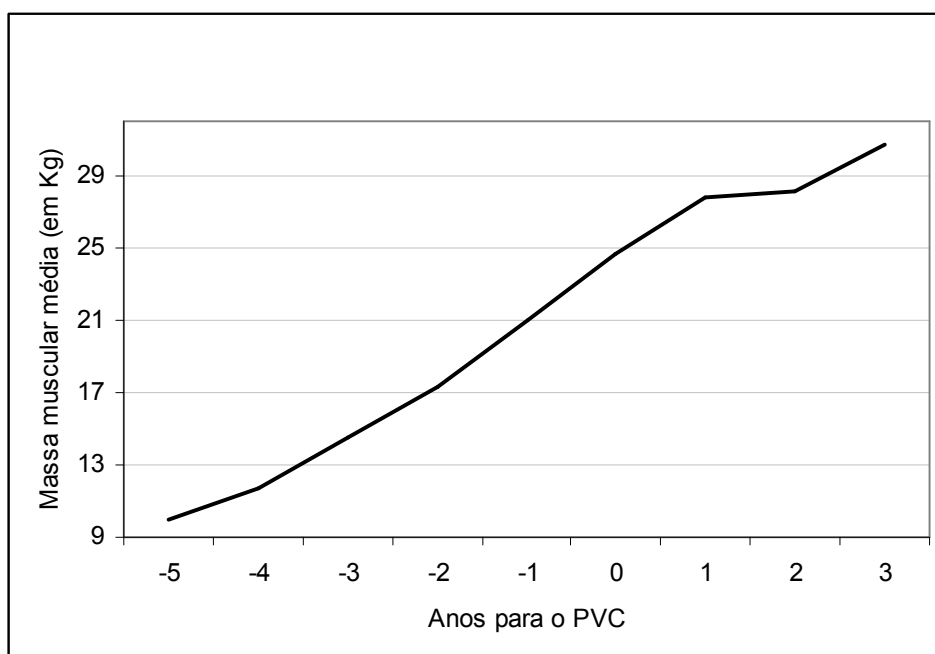
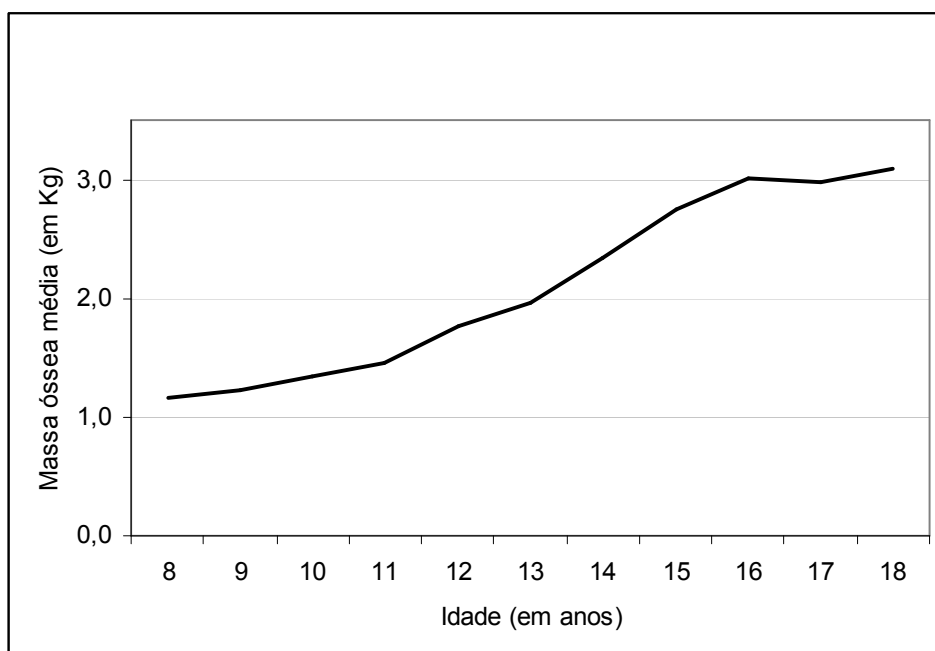
Gráfico B11: Massa muscular média em função do nível de maturação.**Gráfico B12:** Massa óssea média em função da idade.

Gráfico B13: Massa óssea média em função do nível de maturação.

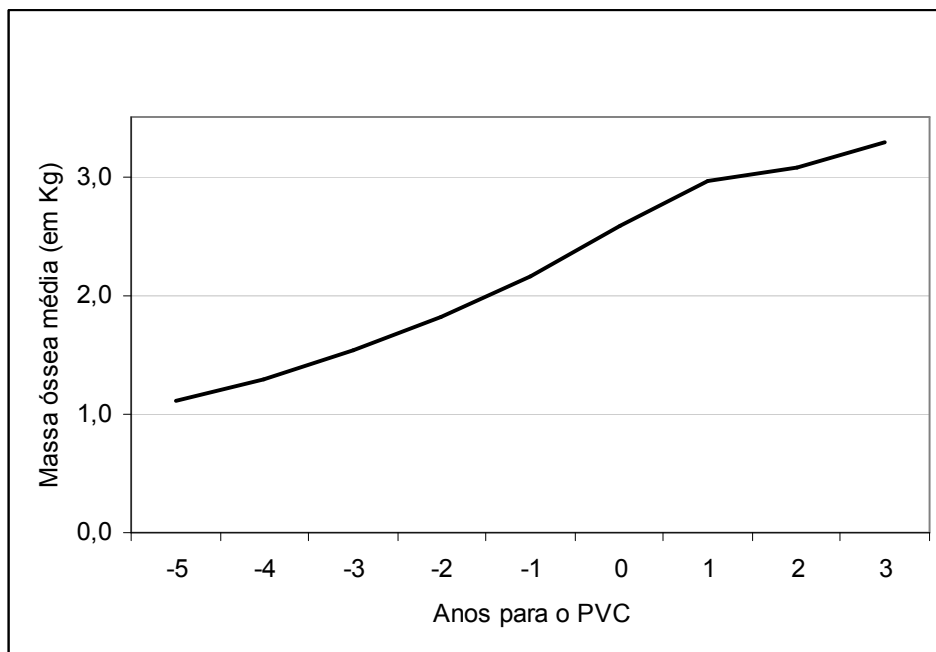


Gráfico B14: Gráfico de dispersão da massa de gordura por idade.

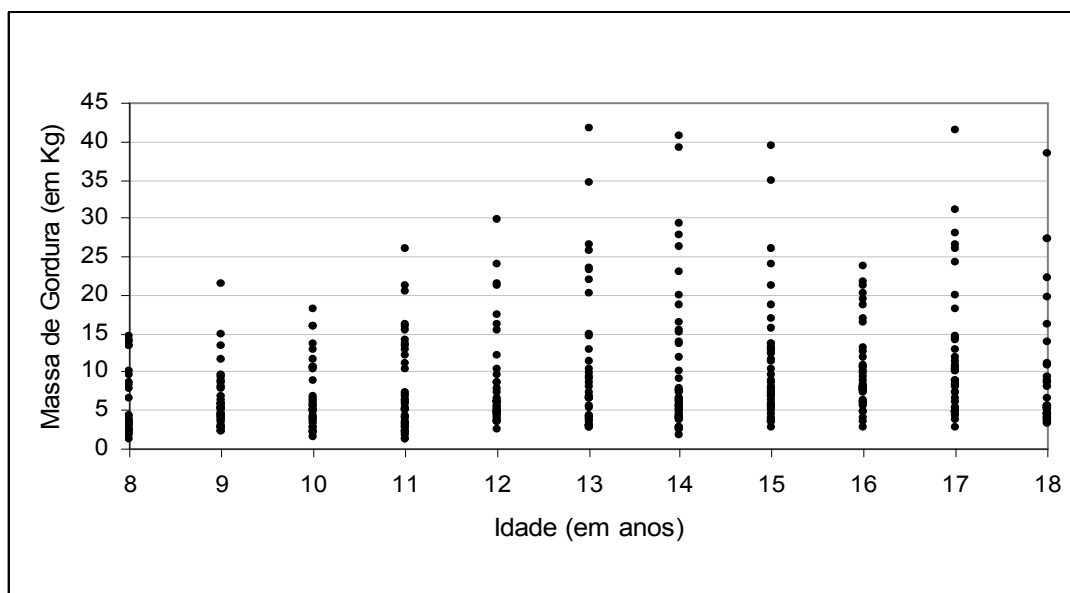


Gráfico B15: Gráfico de dispersão da massa de gordura por nível de maturação.

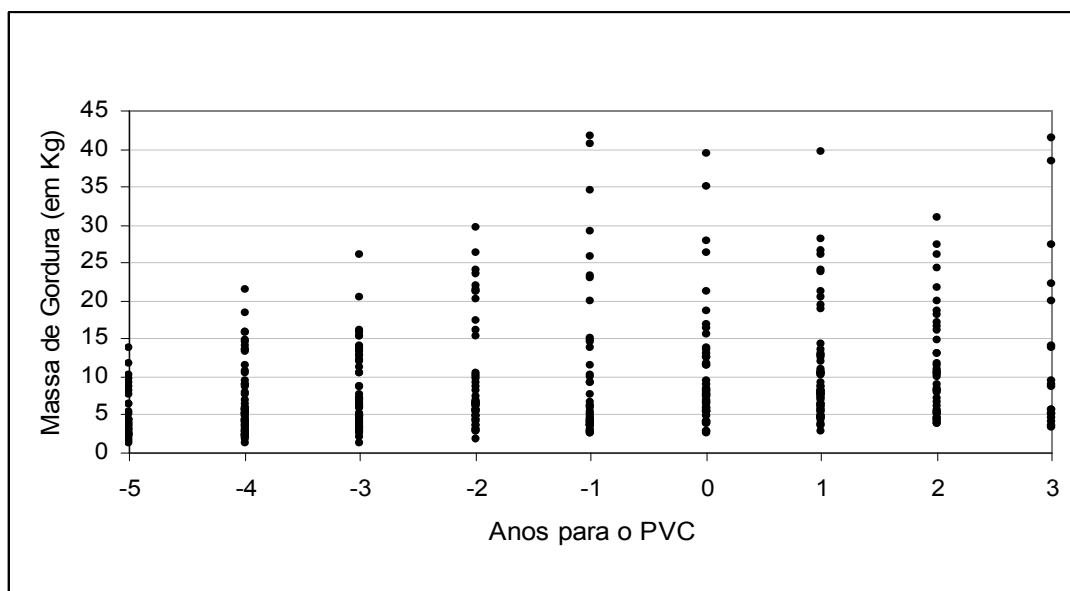


Gráfico B16: Gráfico de dispersão da massa muscular por idade.

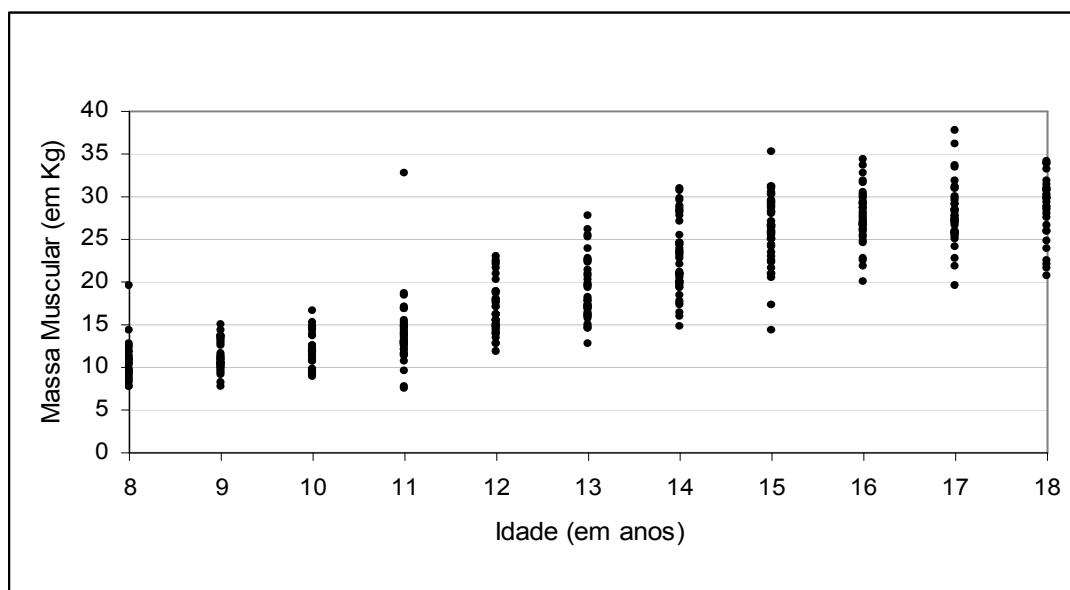


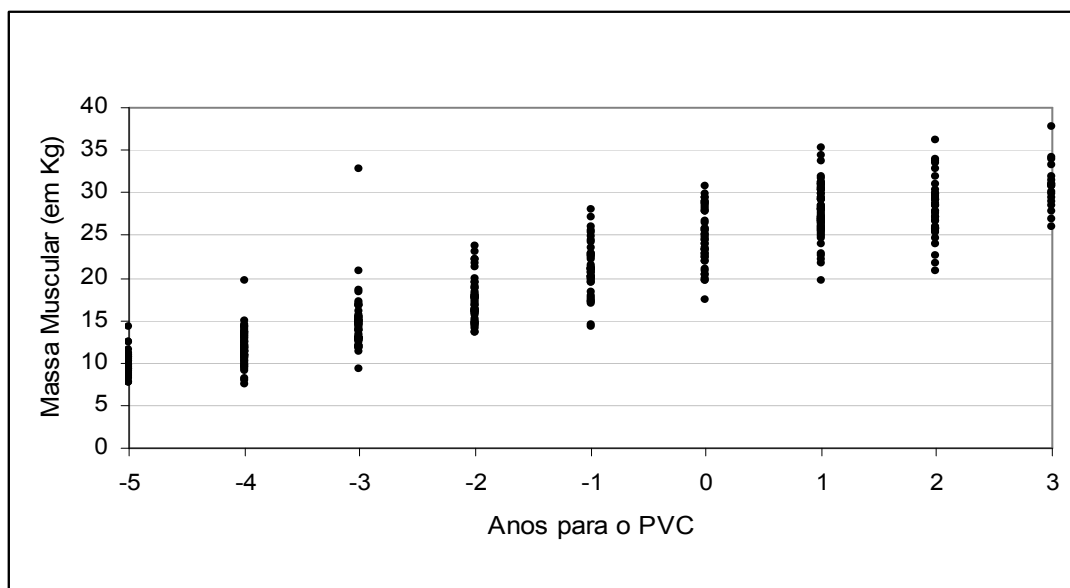
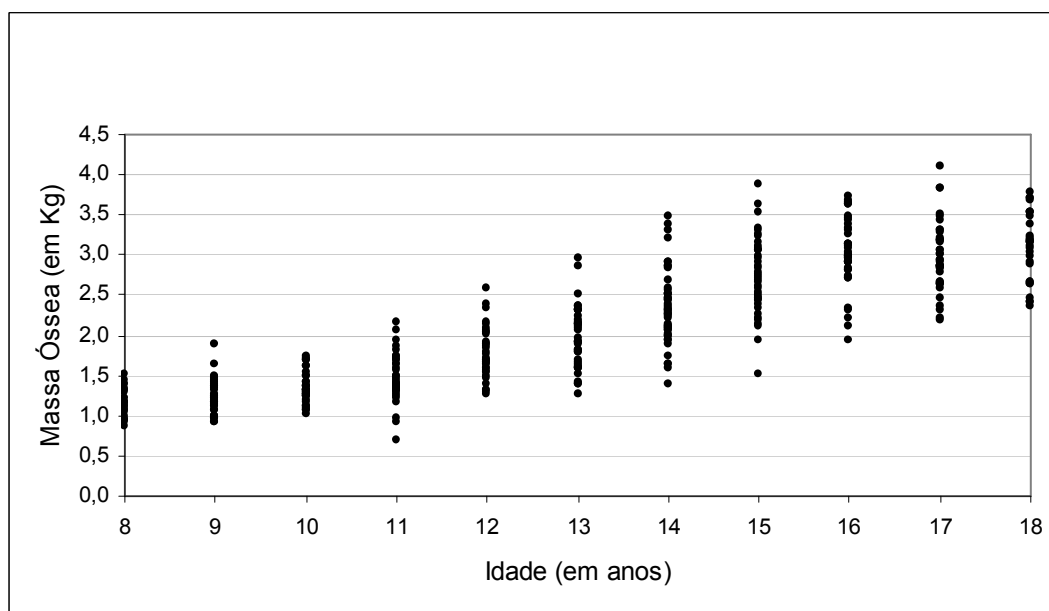
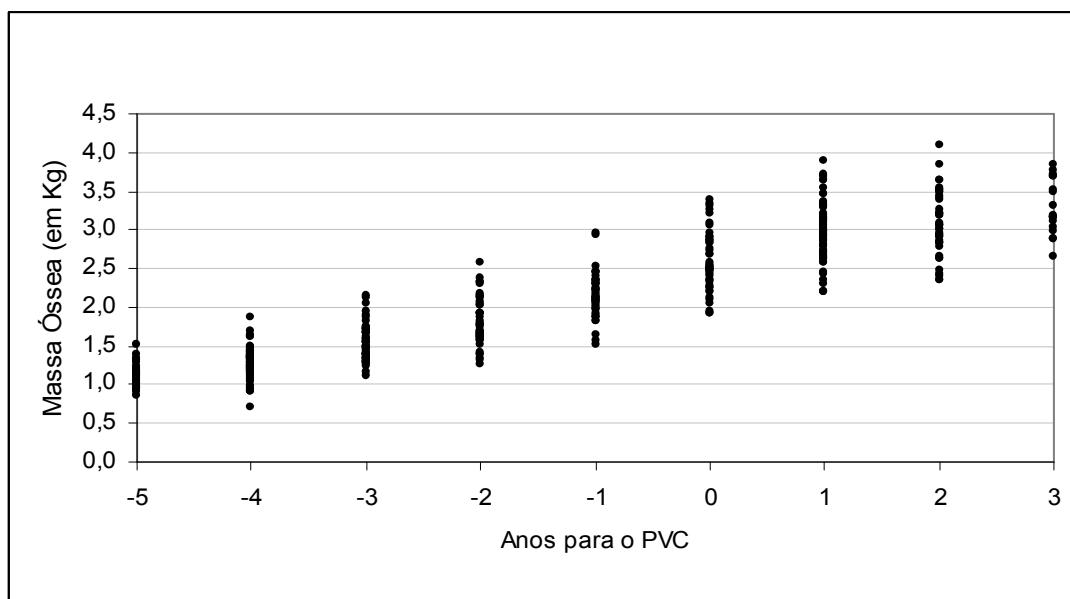
Gráfico B17: Gráfico de dispersão da massa muscular por nível de maturação.**Gráfico B18:** Gráfico de dispersão da massa óssea por idade.

Gráfico B19: Gráfico de dispersão da massa óssea por nível de maturação.



APÊNDICE C

MATRIZ DE CORRELAÇÕES LINEARES DE PEARSON

Sigla	Nome da Variável
Idade	Idade
PVC	Pico de Velocidade de Crescimento
Estatura	Estatura
Tr_cefal	Altura do Tronco-Cefálico
Peso	Peso
Dc_Bi	Espessura da Dobra Cutânea Bicipital
Dc_Tri	Espessura da Dobra Cutânea Tricipital
Dc_Se	Espessura da Dobra Cutânea Subescapular
Dc_Pt	Espessura da Dobra Cutânea Peitoral
Dc_Am	Espessura da Dobra Cutânea Axilar Média
Dc_Si	Espessura da Dobra Cutânea Supra-Iliaca
Dc_AbV	Espessura da Dobra Cutânea Abdominal Vertical
Dc_AbH	Espessura da Dobra Cutânea Abdominal Horizontal
Dc_Cx	Espessura da Dobra Cutânea Coxa Medial
Dc_Pm	Espessura da Dobra Cutânea Panturrilha
Per_Br	Perímetro Braço-Relaxado
Per_Bc	Perímetro Braço-Contraído
Per_Tx(MeEx)	Perímetro Tórax Meso-Exernal
Per_AnB	Perímetro Antebraço
Per_Pun	Perímetro Punho
Per_Cin(md)	Perímetro Cintura
Per_Ab(um)	Perímetro Abdome
Per_Qd	Perímetro Quadril
Per_Cx(m)	Perímetro Coxa
Per_Pn(m)	Perímetro Perna-Medial
Dia_BiAc	Diâmetro Biacromial
Dia_TxTr	Diâmetro Tórax Transverso
Dia_BiCri	Diâmetro Bicrista-Iliaco
Dia_BiTr	Diâmetro Bitrocantérico
Dia_BiEpUm	Diâmetro Biepicondilar do Úmero
Dia_BiEsR_U	Diâmetro Biestiloidal de Rádio e Ulna
Dia_BiEpFem	Diâmetro Biepicondilar do Fêmur
Dia_BiMa	Diâmetro Bimaleolar
MG	Massa de Gordura
MM	Massa Muscular
MO	Massa Óssea

Matriz dos Coeficientes de Correlação linear de Pearson

	Idade	PVC	Estatura	Tr_cefal	Peso
Idade	1	0,973	0,885	0,883	0,778
PVC	0,973	1	0,936	0,960	0,845
Estatura	0,885	0,936	1	0,950	0,843
Tr_cefal	0,883	0,960	0,950	1	0,878
Peso	0,778	0,845	0,843	0,878	1
Dc_Bi	-0,223	-0,180	-0,166	-0,096	0,222
Dc_Tri	-0,174	-0,134	-0,129	-0,054	0,284
Dc_Se	0,145	0,179	0,144	0,229	0,542
Dc_Pt	-0,080	-0,040	-0,040	0,036	0,360
Dc_Am	0,088	0,131	0,122	0,192	0,522
Dc_Si	0,080	0,129	0,119	0,201	0,531
Dc_AbV	0,087	0,136	0,137	0,209	0,542
Dc_AbH	0,060	0,113	0,111	0,187	0,521
Dc_Cx	-0,159	-0,121	-0,134	-0,044	0,277
Dc_Pm	-0,106	-0,064	-0,046	0,022	0,327
Per_Br	0,658	0,713	0,674	0,740	0,909
Per_Bc	0,744	0,798	0,755	0,812	0,933
Per_Tx(MeEx)	0,760	0,813	0,790	0,830	0,940
Per_AnB	0,755	0,816	0,801	0,844	0,936
Per_Pun	0,704	0,755	0,775	0,792	0,829
Per_Cin(md)	0,632	0,683	0,664	0,713	0,921
Per_Ab(um)	0,570	0,628	0,623	0,675	0,899
Per_Qd	0,688	0,748	0,743	0,792	0,924
Per_Cx(m)	0,698	0,760	0,745	0,801	0,941
Per_Pn(m)	0,691	0,756	0,763	0,812	0,942
Dia_BiAc	0,741	0,786	0,798	0,802	0,774
Dia_TxTr	0,758	0,808	0,801	0,821	0,899
Dia_BiCri	0,795	0,849	0,863	0,882	0,906
Dia_BiTr	0,820	0,876	0,884	0,904	0,938
Dia_BiEpUm	0,752	0,806	0,845	0,848	0,865
Dia_BiEsR_U	0,698	0,742	0,779	0,769	0,728
Dia_BiEpFem	0,598	0,653	0,699	0,715	0,773
Dia_BiMa	0,674	0,728	0,788	0,782	0,785
MG	0,222	0,279	0,274	0,350	0,689
MM	0,876	0,926	0,939	0,926	0,888
MO	0,869	0,924	0,915	0,926	0,914

	Dc_Bi	Dc_Tri	Dc_Se	Dc_Pt	Dc_Am	Dc_Si	Dc_AbV	Dc_AbH	Dc_Cx	Dc_Pm
Idade	-0,223	-0,174	0,145	-0,080	0,088	0,080	0,087	0,060	-0,159	-0,106
PVC	-0,180	-0,134	0,179	-0,040	0,131	0,129	0,136	0,113	-0,121	-0,064
Estatura	-0,166	-0,129	0,144	-0,040	0,122	0,119	0,137	0,111	-0,134	-0,046
Tr_cefal	-0,096	-0,054	0,229	0,036	0,192	0,201	0,209	0,187	-0,044	0,022
Peso	0,222	0,284	0,542	0,360	0,522	0,531	0,542	0,521	0,277	0,327
Dc_Bi	1	0,872	0,771	0,856	0,732	0,804	0,808	0,819	0,829	0,796
Dc_Tri	0,872	1	0,818	0,874	0,788	0,856	0,856	0,857	0,894	0,873
Dc_Se	0,771	0,818	1	0,886	0,862	0,905	0,898	0,884	0,783	0,770
Dc_Pt	0,856	0,874	0,886	1	0,854	0,892	0,911	0,906	0,841	0,807
Dc_Am	0,732	0,788	0,862	0,854	1	0,876	0,881	0,856	0,726	0,714
Dc_Si	0,804	0,856	0,905	0,892	0,876	1	0,930	0,916	0,825	0,795
Dc_AbV	0,808	0,856	0,898	0,911	0,881	0,930	1	0,975	0,812	0,804
Dc_AbH	0,819	0,857	0,884	0,906	0,856	0,916	0,975	1	0,826	0,802
Dc_Cx	0,829	0,894	0,783	0,841	0,726	0,825	0,812	0,826	1	0,897
Dc_Pm	0,796	0,873	0,770	0,807	0,714	0,795	0,804	0,802	0,897	1
Per_Br	0,363	0,439	0,656	0,479	0,615	0,649	0,643	0,629	0,413	0,435
Per_Bc	0,254	0,317	0,565	0,376	0,525	0,550	0,552	0,535	0,293	0,330
Per_Tx(MeEx)	0,212	0,277	0,542	0,352	0,511	0,520	0,525	0,505	0,262	0,305
Per_AnB	0,184	0,238	0,495	0,316	0,441	0,477	0,479	0,465	0,233	0,270
Per_Pun	0,099	0,161	0,361	0,198	0,322	0,355	0,352	0,332	0,131	0,210
Per_Cin(md)	0,421	0,478	0,722	0,564	0,692	0,699	0,703	0,683	0,470	0,501
Per_Ab(um)	0,503	0,562	0,766	0,644	0,739	0,764	0,771	0,753	0,548	0,573
Per_Qd	0,312	0,379	0,600	0,458	0,569	0,603	0,605	0,582	0,373	0,410
Per_Cx(m)	0,288	0,364	0,574	0,425	0,555	0,573	0,591	0,574	0,367	0,391
Per_Pn(m)	0,276	0,346	0,561	0,406	0,529	0,548	0,568	0,547	0,342	0,397
Dia_BiAc	-0,029	0,001	0,250	0,084	0,222	0,221	0,242	0,217	0,004	0,069
Dia_TxTr	0,122	0,163	0,428	0,256	0,425	0,418	0,433	0,411	0,159	0,209
Dia_BiCri	0,123	0,172	0,419	0,259	0,399	0,406	0,421	0,397	0,151	0,227
Dia_BiTr	0,110	0,171	0,429	0,259	0,398	0,413	0,429	0,402	0,163	0,227
Dia_BiEpUm	0,052	0,106	0,341	0,176	0,317	0,335	0,335	0,315	0,089	0,168
Dia_BiEsR_U	-0,067	-0,026	0,209	0,035	0,161	0,169	0,166	0,146	-0,020	0,047
Dia_BiEpFem	0,172	0,226	0,391	0,258	0,365	0,379	0,385	0,355	0,197	0,261
Dia_BiMa	0,001	0,045	0,253	0,115	0,253	0,232	0,253	0,234	0,025	0,104
MG	0,745	0,814	0,894	0,845	0,861	0,917	0,925	0,919	0,797	0,788
MM	-0,155	-0,110	0,178	-0,028	0,169	0,151	0,162	0,138	-0,115	-0,042
MO	-0,075	-0,018	0,267	0,056	0,241	0,234	0,246	0,220	-0,016	0,048

	Per_Br	Per_Bc	Per_Tx(MeEx)	Per_AnB	Per_Pun	Per_Cin(md)	Per_Ab(um)	Per_Qd	Per_Cx(m)	Per_Pn(m)
Idade	0,658	0,744	0,760	0,755	0,704	0,632	0,570	0,688	0,698	0,691
PVC	0,713	0,798	0,813	0,816	0,755	0,683	0,628	0,748	0,760	0,756
Estatutura	0,674	0,755	0,790	0,801	0,775	0,664	0,623	0,743	0,745	0,763
Tr_cefal	0,740	0,812	0,830	0,844	0,792	0,713	0,675	0,792	0,801	0,812
Peso	0,909	0,933	0,940	0,936	0,829	0,921	0,899	0,924	0,941	0,942
Dc_Bi	0,363	0,254	0,212	0,184	0,099	0,421	0,503	0,312	0,288	0,276
Dc_Tri	0,439	0,317	0,277	0,238	0,161	0,478	0,562	0,379	0,364	0,346
Dc_Se	0,656	0,565	0,542	0,495	0,361	0,722	0,766	0,600	0,574	0,561
Dc_Pt	0,479	0,376	0,352	0,316	0,198	0,564	0,644	0,458	0,425	0,406
Dc_Am	0,615	0,525	0,511	0,441	0,322	0,692	0,739	0,569	0,555	0,529
Dc_Si	0,649	0,550	0,520	0,477	0,355	0,699	0,764	0,603	0,573	0,548
Dc_AbV	0,643	0,552	0,525	0,479	0,352	0,703	0,771	0,605	0,591	0,568
Dc_AbH	0,629	0,535	0,505	0,465	0,332	0,683	0,753	0,582	0,574	0,547
Dc_Cx	0,413	0,293	0,262	0,233	0,131	0,470	0,548	0,373	0,367	0,342
Dc_Pm	0,435	0,330	0,305	0,270	0,210	0,501	0,573	0,410	0,391	0,397
Per_Br	1	0,978	0,917	0,915	0,794	0,915	0,900	0,873	0,906	0,884
Per_Bc	0,978	1	0,930	0,940	0,815	0,896	0,867	0,879	0,916	0,897
Per_Tx(MeEx)	0,917	0,930	1	0,915	0,823	0,908	0,877	0,880	0,904	0,893
Per_AnB	0,915	0,940	0,915	1	0,848	0,867	0,840	0,880	0,910	0,907
Per_Pun	0,794	0,815	0,823	0,848	1	0,770	0,736	0,773	0,793	0,798
Per_Cin(md)	0,915	0,896	0,908	0,867	0,770	1	0,972	0,880	0,909	0,887
Per_Ab(um)	0,900	0,867	0,877	0,840	0,736	0,972	1	0,885	0,896	0,873
Per_Qd	0,873	0,879	0,880	0,880	0,773	0,880	0,885	1	0,913	0,899
Per_Cx(m)	0,906	0,916	0,904	0,910	0,793	0,909	0,896	0,913	1	0,943
Per_Pn(m)	0,884	0,897	0,893	0,907	0,798	0,887	0,873	0,899	0,943	1
Dia_BiAc	0,659	0,721	0,745	0,737	0,671	0,655	0,624	0,704	0,697	0,712
Dia_TxTr	0,821	0,861	0,884	0,854	0,771	0,839	0,801	0,833	0,841	0,840
Dia_BiCri	0,795	0,835	0,859	0,856	0,795	0,818	0,811	0,862	0,837	0,843
Dia_BiTr	0,816	0,866	0,880	0,886	0,802	0,828	0,811	0,893	0,885	0,892
Dia_BiEpUm	0,774	0,823	0,822	0,867	0,842	0,768	0,744	0,802	0,817	0,828
Dia_BiEsR_U	0,609	0,674	0,679	0,722	0,718	0,598	0,563	0,662	0,681	0,700
Dia_BiEpFem	0,675	0,711	0,720	0,738	0,696	0,703	0,690	0,741	0,760	0,778
Dia_BiMa	0,662	0,715	0,747	0,749	0,771	0,683	0,653	0,717	0,743	0,777
MG	0,751	0,665	0,650	0,610	0,482	0,803	0,861	0,721	0,711	0,696
MM	0,740	0,822	0,837	0,862	0,797	0,717	0,660	0,776	0,811	0,818
MO	0,771	0,842	0,859	0,871	0,770	0,765	0,716	0,818	0,840	0,844

	Dia_BiAc	Dia_TxTr	Dia_BiCri	Dia_BiTr	Dia_BiEpUm	Dia_BiEsR_U	Dia_BiEpFem	Dia_BiMa	MG	MM	MO
Idade	0,741	0,758	0,795	0,820	0,752	0,698	0,598	0,674	0,222	0,876	0,869
PVC	0,786	0,808	0,849	0,876	0,806	0,742	0,653	0,728	0,279	0,926	0,924
Estatutura	0,798	0,801	0,863	0,884	0,845	0,779	0,699	0,788	0,274	0,939	0,915
Tr_cefal	0,802	0,821	0,882	0,904	0,848	0,769	0,715	0,782	0,350	0,926	0,926
Peso	0,774	0,899	0,906	0,938	0,865	0,728	0,773	0,785	0,689	0,888	0,914
Dc_Bi	-0,029	0,122	0,123	0,110	0,052	-0,067	0,172	0,001	0,745	-0,155	-0,075
Dc_Tri	0,001	0,163	0,172	0,171	0,106	-0,026	0,226	0,045	0,814	-0,110	-0,018
Dc_Se	0,250	0,428	0,419	0,429	0,341	0,209	0,391	0,253	0,894	0,178	0,267
Dc_Pt	0,084	0,256	0,259	0,259	0,176	0,035	0,258	0,115	0,845	-0,028	0,056
Dc_Am	0,222	0,425	0,399	0,398	0,317	0,161	0,365	0,253	0,861	0,169	0,241
Dc_Si	0,221	0,418	0,406	0,413	0,335	0,169	0,379	0,232	0,917	0,151	0,234
Dc_AbV	0,242	0,433	0,421	0,429	0,335	0,166	0,385	0,253	0,925	0,162	0,246
Dc_AbH	0,217	0,411	0,397	0,402	0,315	0,146	0,355	0,234	0,919	0,138	0,220
Dc_Cx	0,004	0,159	0,151	0,163	0,089	-0,020	0,197	0,025	0,797	-0,115	-0,016
Dc_Pm	0,069	0,209	0,227	0,227	0,168	0,047	0,261	0,104	0,788	-0,042	0,048
Per_Br	0,659	0,821	0,795	0,816	0,774	0,609	0,675	0,662	0,751	0,740	0,771
Per_Bc	0,721	0,861	0,835	0,866	0,823	0,674	0,711	0,715	0,665	0,822	0,842
Per_Tx(MeEx)	0,745	0,884	0,859	0,880	0,822	0,679	0,720	0,747	0,650	0,837	0,859
Per_AnB	0,737	0,854	0,856	0,886	0,867	0,722	0,738	0,749	0,610	0,862	0,871
Per_Pun	0,671	0,771	0,795	0,802	0,842	0,718	0,696	0,771	0,482	0,797	0,770
Per_Cin(md)	0,655	0,839	0,818	0,828	0,768	0,598	0,703	0,683	0,803	0,717	0,765
Per_Ab(um)	0,624	0,801	0,811	0,811	0,744	0,563	0,690	0,653	0,861	0,660	0,716
Per_Qd	0,704	0,833	0,862	0,893	0,802	0,662	0,741	0,717	0,721	0,776	0,818
Per_Cx(m)	0,697	0,841	0,837	0,885	0,817	0,681	0,760	0,743	0,711	0,811	0,840
Per_Pn(m)	0,712	0,840	0,843	0,892	0,828	0,700	0,778	0,777	0,696	0,818	0,844
Dia_BiAc	1	0,743	0,776	0,799	0,748	0,665	0,643	0,697	0,359	0,793	0,803
Dia_TxTr	0,743	1	0,861	0,886	0,818	0,688	0,742	0,756	0,558	0,834	0,855
Dia_BiCri	0,776	0,861	1	0,931	0,855	0,747	0,770	0,789	0,549	0,846	0,859
Dia_BiTr	0,799	0,886	0,931	1	0,873	0,782	0,802	0,804	0,559	0,893	0,909
Dia_BiEpUm	0,748	0,818	0,855	0,873	1	0,752	0,789	0,820	0,464	0,855	0,838
Dia_BiEsR_U	0,665	0,688	0,747	0,782	0,752	1	0,700	0,748	0,295	0,777	0,765
Dia_BiEpFem	0,643	0,742	0,770	0,802	0,789	0,700	1	0,765	0,500	0,717	0,727
Dia_BiMa	0,697	0,756	0,789	0,804	0,820	0,748	0,765	1	0,388	0,800	0,777
MG	0,359	0,558	0,549	0,559	0,464	0,295	0,500	0,388	1	0,314	0,406
MM	0,793	0,834	0,846	0,893	0,855	0,777	0,717	0,800	0,314	1	0,947
MO	0,803	0,855	0,859	0,909	0,838	0,765	0,727	0,777	0,406	0,947	1

APÊNDICE D

MULTICOLINEARIDADE

MULTICOLINEARIDADE

Multicolinearidade é a existência de relações lineares entre variáveis independentes que, no ajuste do modelo de regressão linear, traz consequências indesejáveis para a análise inferencial, tais como:

- no caso de perfeita multicolinearidade, não é possível estimar os parâmetros do modelo pelo método de mínimos quadrados.
- na maioria das situações práticas em que há multicolinearidade, as variáveis independentes são altamente correlacionadas, mas não perfeitamente colineares.

Neste caso:

- as estimativas dos coeficientes de regressão tornam-se muito sensíveis a pequenas alterações na matriz de planejamento.
- as variâncias dos estimadores são altas, tornando os testes do tipo $\beta_j = 0$ contra $\beta_j \neq 0$ (além de outros) pouco sensíveis e, assim, variáveis independentes importantes podem ser eliminadas.

A principal fonte de multicolinearidade neste projeto vem das relações naturais existentes entre variáveis independentes.

Existem vários métodos para verificação da existência de multicolinearidade. A seguir, é descrito o método utilizado neste projeto.

Uso dos auto-valores da matriz de correlações lineares de Pearson das variáveis independentes.

Sejam $\lambda_1, \lambda_2, \dots, \lambda_p$ os auto-valores da matriz de correlações lineares de Pearson das variáveis independentes (\mathbf{R}_X).

Verifica-se que $\prod_{i=1}^p \lambda_i = \det(\mathbf{R}_X)$, em que $\det(\mathbf{R}_X)$ é o determinante de \mathbf{R}_X .

Se houver um ou mais auto-valores próximos de zero, então, $\prod_{i=1}^p \lambda_i = \det(\mathbf{R}_X)$ é aproximadamente zero, indicando existência de multicolinearidade.

Nesta análise, define-se $L = \frac{\lambda_{\max}}{\lambda_{\min}}$,

em que λ_{\max} é o maior auto-valor da matriz \mathbf{R}_X e

λ_{\min} é o menor auto-valor da matriz \mathbf{R}_X .

A literatura sugere o critério:

$L < 100$	→	baixa multicolinearidade
$100 \leq L < 1000$	→	média multicolinearidade
$L \geq 1000$	→	alta multicolinearidade

Este método de verificação da existência de multicolinearidade é um dos mais utilizados.

Para maiores detalhes sobre multicolinearidade, vide Montgomery et al. (2001).

APÊNDICE E

ANÁLISE DE RESÍDUOS

ANÁLISE DE RESÍDUOS – NORMALIDADE TRIVARIADA

Para verificar a suposição de normalidade trivariada dos erros do modelo de regressão foi utilizado um método baseado nas distâncias generalizadas ao quadrado, para amostras grandes (Johnson e Wichern, 2007, p.186). Primeiramente, é calculada a distância para cada observação, definida como: $d_j^2 = (\mathbf{x}_j - \bar{\mathbf{x}})' \hat{\Sigma}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})$, em que \mathbf{x}_j é o vetor dos resíduos da j -ésima observação, com $j = 1, \dots, 408$, $\bar{\mathbf{x}}$ é o vetor das médias dos resíduos, $\hat{\Sigma} = \frac{(\mathbf{Y} - \mathbf{Z}\hat{\beta})'(\mathbf{Y} - \mathbf{Z}\hat{\beta})}{n}$ e $n = 408$ indivíduos.

Para amostras grandes, essas quantidades devem se comportar como uma variável aleatória com distribuição qui-quadrado.

Constrói-se então um gráfico do tipo “q-q plot”. Nessa construção, deve-se primeiramente ordenar as distâncias em ordem crescente, e cada ponto do gráfico será dado por $(\chi_m^2((j-0,5)/n), d_{(j)}^2)$, sendo que $\chi_m^2((j-0,5)/n)$ é o percentil de ordem $100((j-0,5)/n)$ de uma distribuição qui-quadrado com m graus de liberdade, e $m = 3$, o número de variáveis resposta. Havendo normalidade, tal gráfico deve resultar em uma linha reta; e um ou dois pontos distantes e acima da reta indicariam observações discrepantes.

Os Gráficos E1, E2 e E3 sugerem a normalidade trivariada para os erros dos três modelos ajustados.

ANÁLISE DE RESÍDUOS UNIVARIADA

Para modelos de regressão linear multivariada, como os ajustados neste projeto, a análise de resíduos pode ser feita univariadamente (Johnson e Wichern, 2007).

Os gráficos dos resíduos para os três modelos são os Gráficos E4 a E12, para as três variáveis resposta.

Gráfico E1: Análise de normalidade trivariada — primeiro modelo.

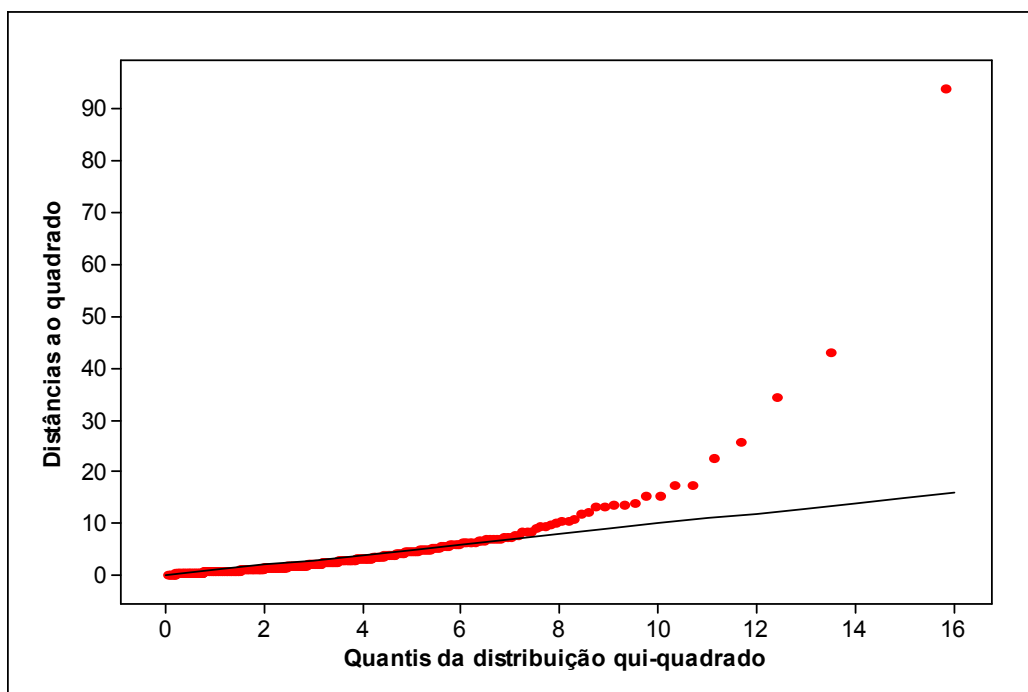


Gráfico E2: Análise de normalidade trivariada — segundo modelo.

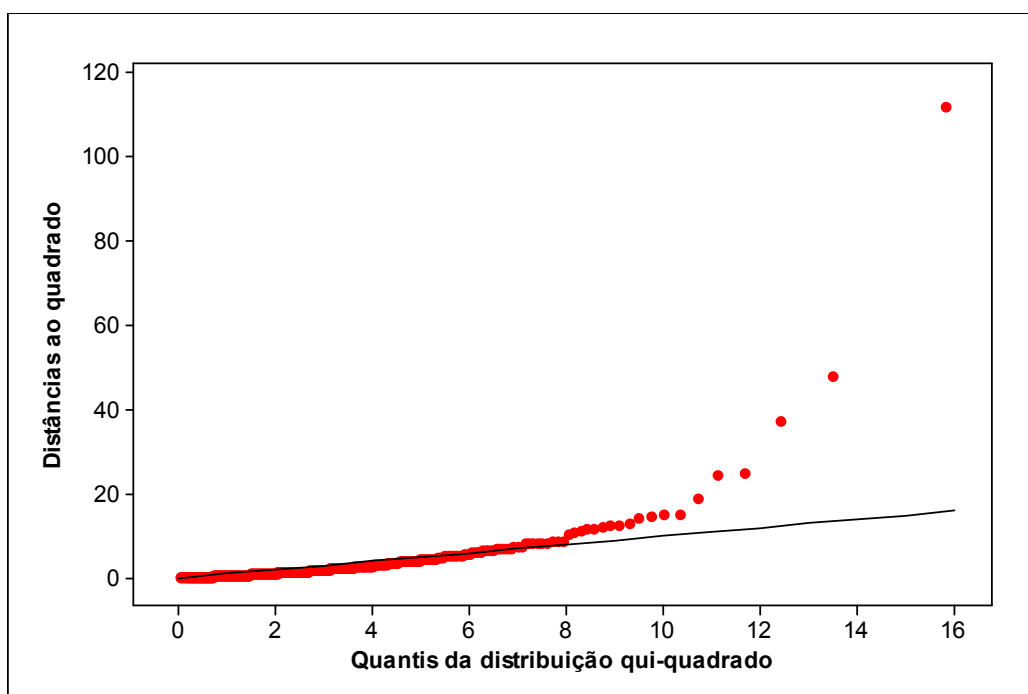


Gráfico E3: Análise de normalidade trivariada — terceiro modelo.

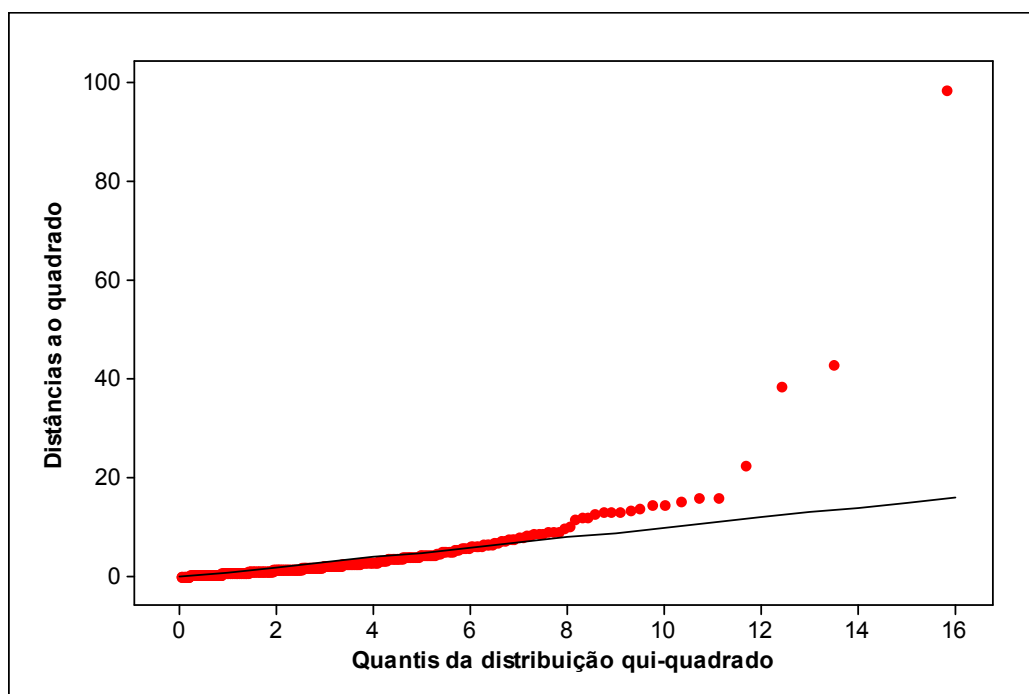


Gráfico E4: Gráfico de resíduos para massa de gordura – primeiro modelo.

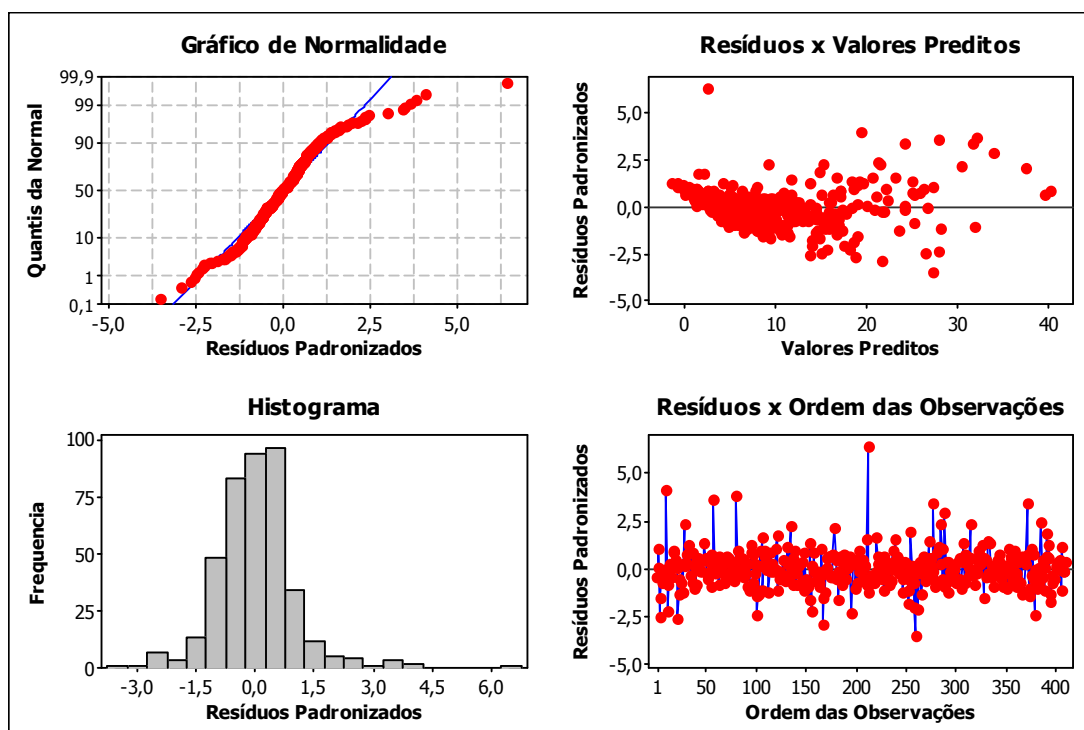


Gráfico E5: Gráfico de resíduos para massa muscular – primeiro modelo.

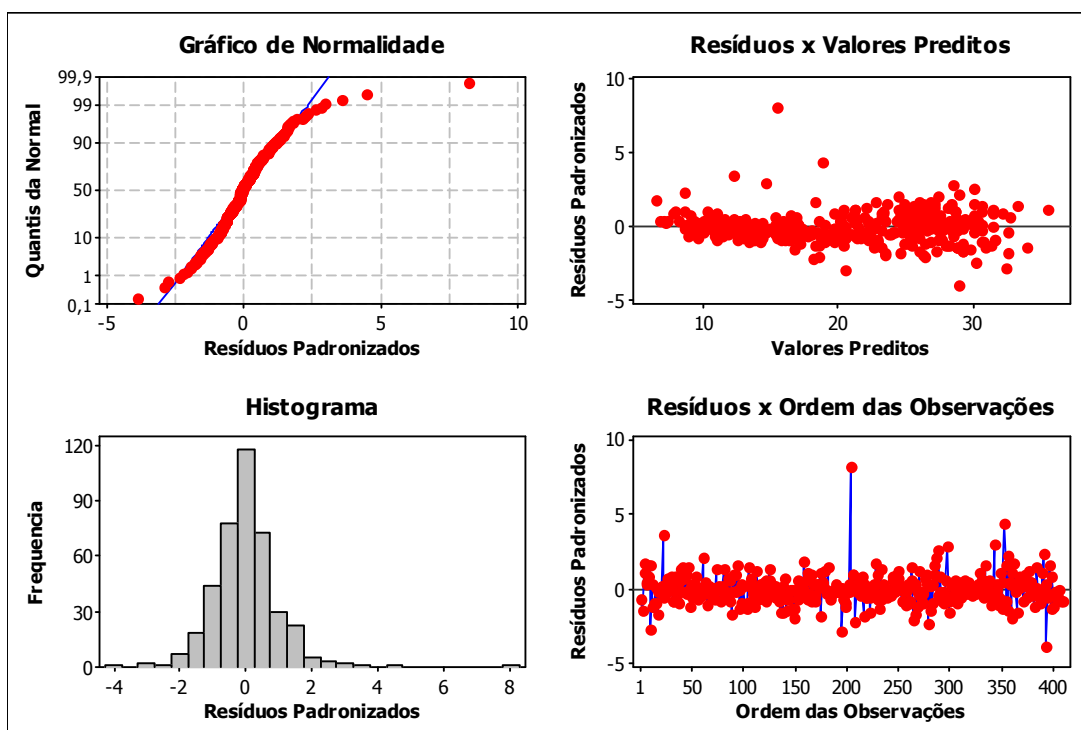


Gráfico E6: Gráfico de resíduos para massa óssea – primeiro modelo.

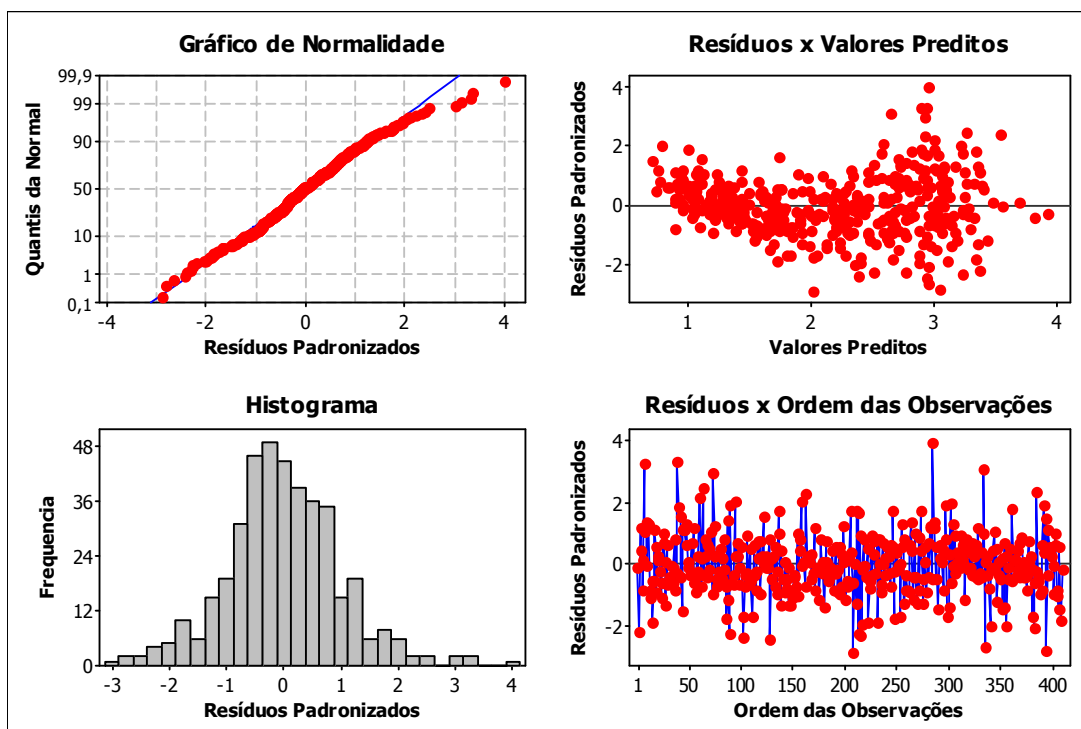


Gráfico E7: Gráfico de resíduos para massa de gordura – segundo modelo.

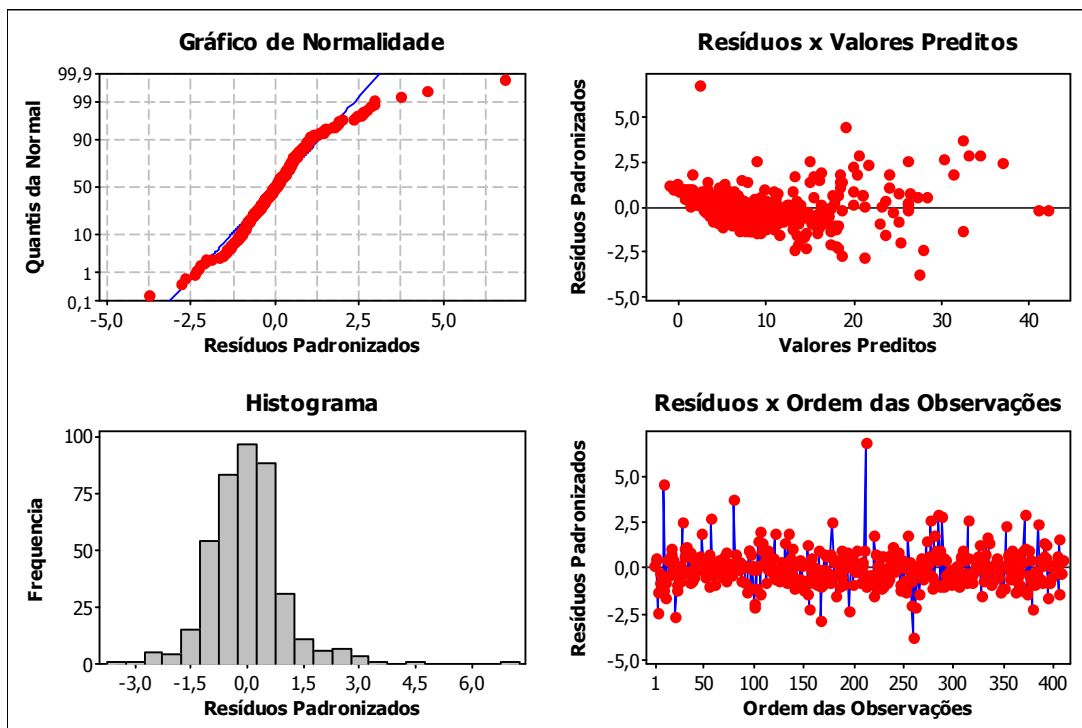


Gráfico E8: Gráfico de resíduos para massa muscular – segundo modelo.

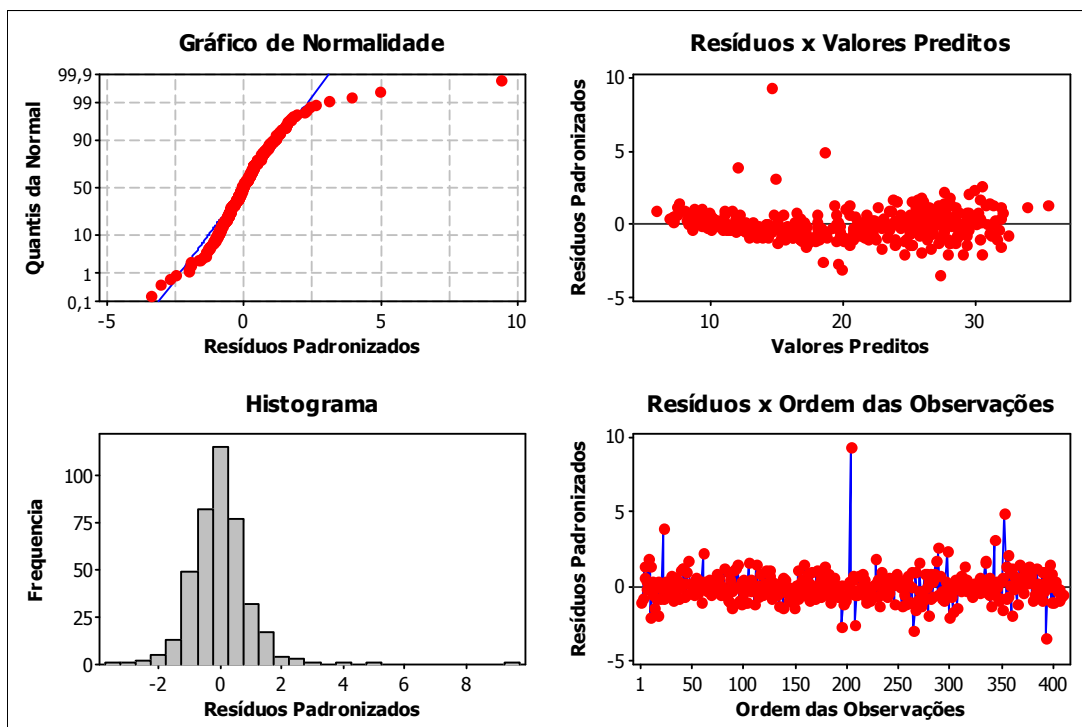


Gráfico E9: Gráfico de resíduos para massa óssea – segundo modelo.

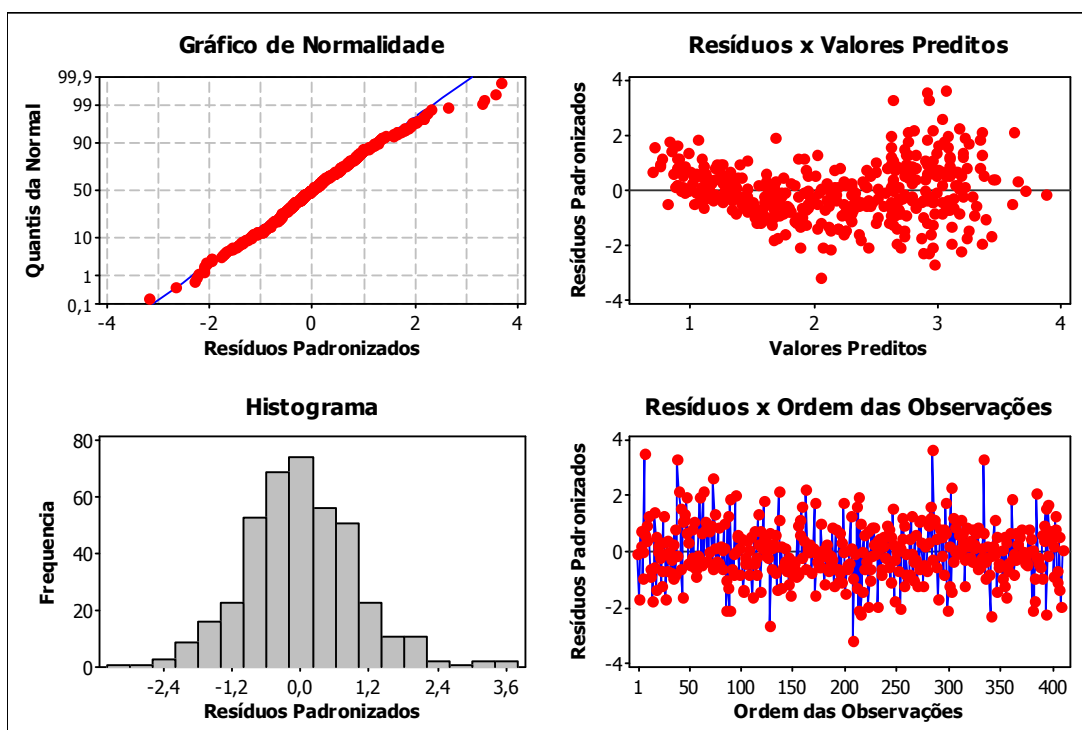


Gráfico E10: Gráfico de resíduos para massa de gordura – terceiro modelo.

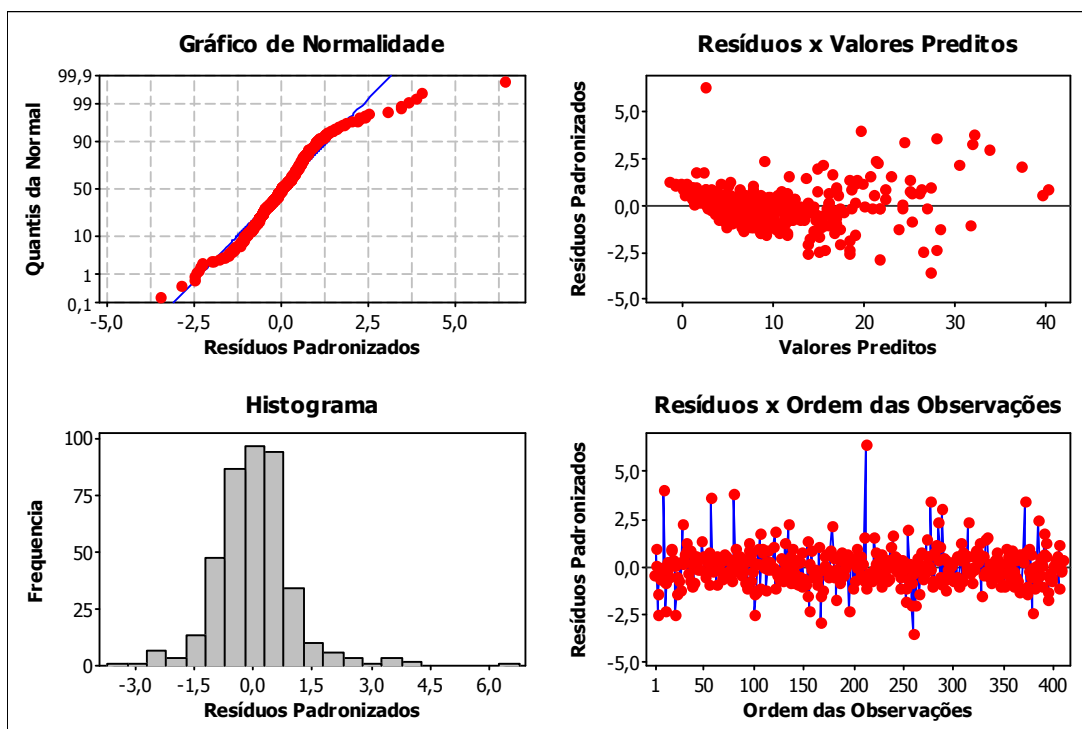


Gráfico E11: Gráfico de resíduos para massa muscular – terceiro modelo.

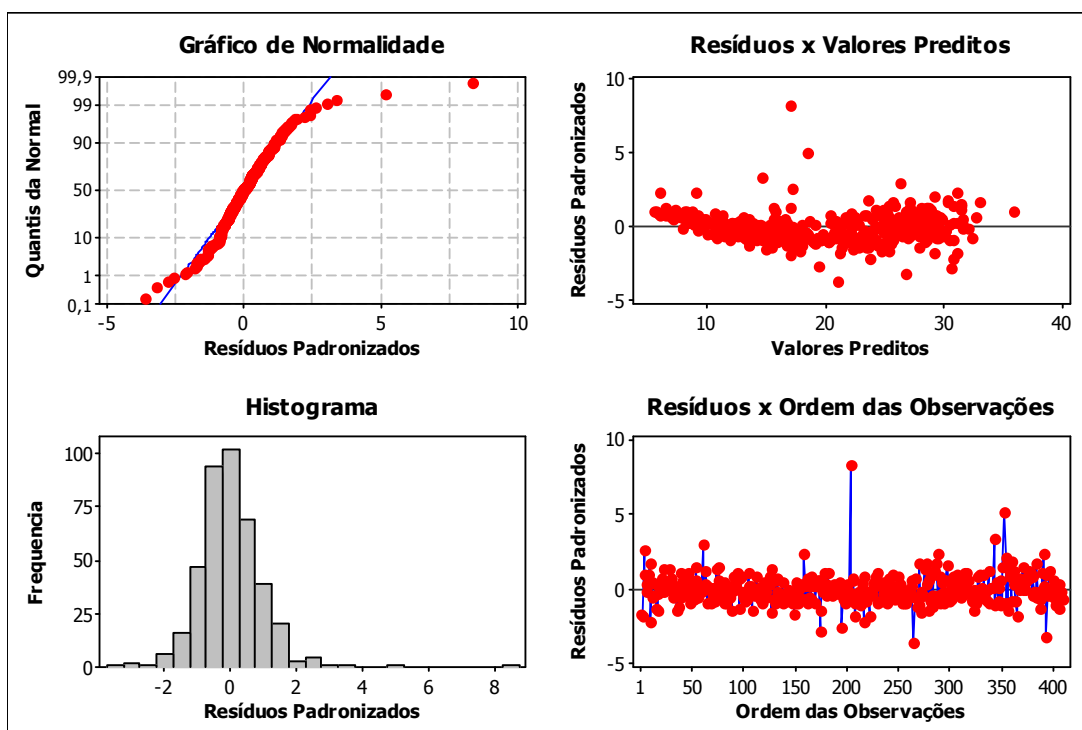


Gráfico E12: Gráfico de resíduos para massa óssea – terceiro modelo.

