Relatório Análise de dados DL-2015

Felipe Figueiredo

January 21, 2016

# Metodologia utilizada

A análise descritiva foi apresentada na forma de tabelas os dados observados, expressos pela média, desvio padrão, mediana, amplitude interquartílica (AIQ), mínimo e máximo para dados numéricos (quantitativos) e pela frequência (n) e percentual (%) para dados categóricos (qualitativos). São apresentados gráficos descritivos das variáveis basais e finais. Ao longo do texto, as variáveis normais foram sumarizadas como média +/ desvio padrão (DP), e as variáveis não normais como mediana +- AIQ.

Foi aplicado o teste de normalidade de Shapiro-Wilks nas variáveis numéricas. A comparação das variáveis basais e finais foi feita com o teste t pareado para variáveis normalmente distribuídas, e com o teste de Mann-Whitney pareado para variáveis que tiveram a hipótese de normalidade rejeitada. Foram calculadas as correlações entre o nível final de Vitamina D e o nível final de cada marcador. Para variáveis normais, foi aplicada a correlação de Pearson, e para variáveis não normais, a correlação não paramétrica de Spearman. Em todas as análises, foi adotada a significância de 5%. As análises estatística foram feitas utilizando o software estatístico R, versão 3.2.2 (<https://www.R-project.org>).

# Resultados

## Estatísticas descritivas

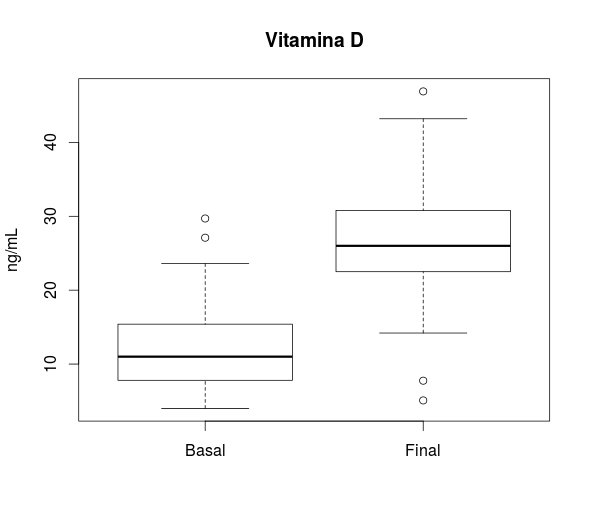
### Estatísticas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Média | DP | Mediana | AIQ |
| **VITAMINAD.BASAL** | 13.41 | 7.27 | 11 | 7.6 |
| **VITAMINAD.FINAL** | 26.78 | 10.87 | 26 | 8.3 |
| **ADIPONECTINA.BASAL** | 0.16 | 0.05 | 0.15 | 0.07 |
| **ADIPONECTINA.FINAL** | 0.27 | 0.09 | 0.26 | 0.12 |
| **CTX.BASAL** | 0.31 | 0.13 | 0.31 | 0.19 |
| **CTX.FINAL** | 0.16 | 0.06 | 0.14 | 0.08 |
| **FAO.BASAL** | 30.23 | 13.8 | 33.5 | 11.9 |
| **FAO.FINAL** | 21.99 | 8.95 | 22.9 | 12 |
| **LEPTINA.BASAL** | 0.28 | 0.04 | 0.27 | 0.07 |
| **LEPTINA.FINAL** | 0.2 | 0.08 | 0.2 | 0.15 |
| **OPG.BASAL** | 2.68 | 0.66 | 2.7 | 0.8 |
| **OPG.FINAL** | 3.12 | 0.56 | 2.9 | 0.7 |
| **TNFALFA.BASAL** | 4.83 | 1.06 | 4.8 | 1.2 |
| **TNFALFA.FINAL** | 3.87 | 0.87 | 3.6 | 1.3 |

Tabela 1: Estatísticas descritivas da vitamina D e marcadores (siglas), em média (DP), mediana (AIQ).

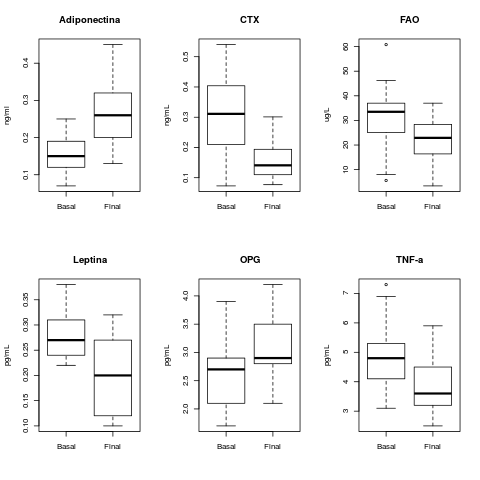
As estatísticas descritivas das variáveis analisadas estão sumarizadas na tabela

### Gráficos



Boxplots da Vitamina D

Figura 1: Níveis basal e final da Vitamina D (boxplot: mediana, quartis, outlier)

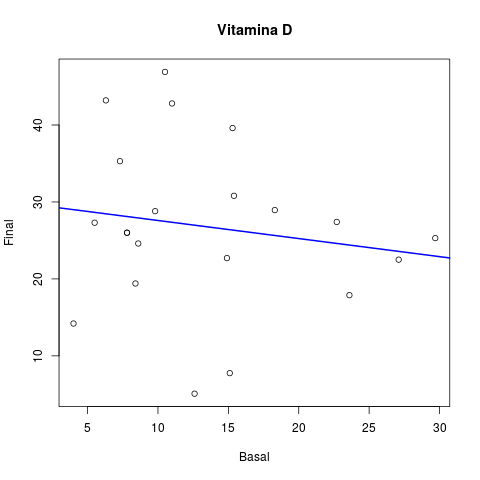


Boxplots dos marcadores

Figura 2: Níveis basal e final dos marcadores ADP, CTX, FAO, LEP, OPG e TNF (siglas e boxplot: mediana, quartis, outliers)

São apresentadas nas figuras 3 e 4 as retas de melhor ajuste aos ...

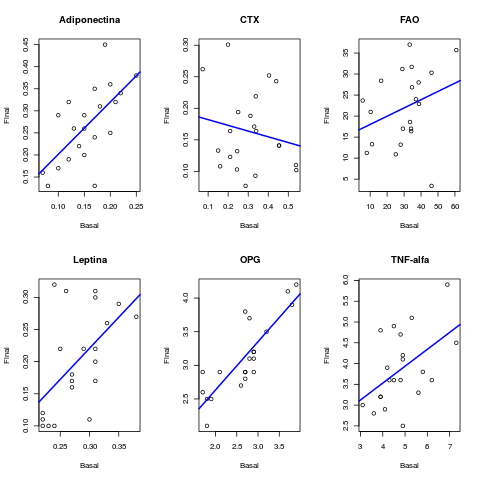
Os níveis basal e final foram significativamente diferentes, conforme seção X, e resumidos na tabela X.



Scatterplot Vitamina D

Figura 3: Comparação entre a Vitamina D basal e final (completar legenda). Índice de correlação de Spearman: -0.1241066 (p-valor=0.592)

Sugestão: **não incluir a figura acima, da reta de melhor ajuste da vitamina D**. Incluir no trabalho apenas as figuras 1 e 2 que ilustram a diferença entre os níveis basal e final, e a figura abaixo, com os dados e a reta de melhor ajuste dos marcadores.



Scatterplot marcadores

Figura 4: Comparação entre os níveis basal e final dos marcadores ADP, CTX, FAO, LEP, OPG e TNF (completar, siglas, etc). A linha azul é a reta de melhor ajuste aos dados. Índices de correlação de Spearman: ADP (rho=0.6564454, p-valor=0.001229), CTX (rho=-0.1102081, p-valor=0.6344), FAO (rho=0.3164392, p-valor=0.1623), LEP (rho=0.547378, p-valor=0.01022), OPG (rho=0.8328939, p-valor=2.791e-06) e TNF (rho=0.464204, p-valor=0.03402).

## Normalidade

Considerando conjuntamente os valores basal e final, nenhuma das variáveis parece ter distribuição normal. Deve-se rejeitar a hipótese de normalidade de **todas** as variáveis, e utilizar apenas métodos e testes não paramétricos. Tabela 2.

|  |  |
| --- | --- |
| Variável | p.valor |
| VITAMINAD.BASAL | 0.05342 |
| VITAMINAD.FINAL | 0.6507 |
| ADIPONECTINA.BASAL | 0.9775 |
| ADIPONECTINA.FINAL | 0.8827 |
| CTX.BASAL | 0.7354 |
| CTX.FINAL | 0.16 |
| FAO.BASAL | 0.2291 |
| FAO.FINAL | 0.8838 |
| LEPTINA.BASAL | 0.2637 |
| LEPTINA.FINAL | 0.05006 |
| OPG.BASAL | 0.0624 |
| OPG.FINAL | 0.3222 |
| TNFALFA.BASAL | 0.2568 |
| TNFALFA.FINAL | 0.6046 |

Tabela 2: Testes de normalidade dos níveis basal e final da Vitamina D e dos marcadores ADP, CTX, FAO, LEP, OPG e TNF (completar, siglas, etc). P-valores apresentados para o teste de Shapiro-Wilks.

## Testes pareados

Todos os testes pareados indicaram aumento/diminuição altamente significativa entre o valor basal e final de cada variável. Tabela 3.

|  |  |  |
| --- | --- | --- |
| Variável | Desfecho | p.valor |
| VITD | aumentou | 0.0003735 |
| ADP | aumentou | 3.683e-05 |
| CTX | diminuiu | 0.0004796 |
| FAO | diminuiu | 0.008773 |
| LEP | diminuiu | 0.0002087 |
| OPG | aumentou | 4.704e-05 |
| TNF | diminuiu | 0.0003489 |

Tabela 3: Diferenças entre os valores basal e final dos marcadores ADP, CTX, FAO, LEP, OPG e TNF (siglas). P-valores apresentados de acordo com o teste de Mann-Whitney pareado.

## Correlações

Nenhuma das correlações entre Vitamina D e marcadores foi significativa com os dados analisados. Tabela 4

|  |  |  |
| --- | --- | --- |
| Variável | rho | p.valor |
| ADP | 0.1167 | 0.6143 |
| CTX | 0.0156 | 0.9465 |
| FAO | 0.2671 | 0.2419 |
| LEP | 0.1255 | 0.5878 |
| OPG | 0.03174 | 0.8914 |
| TNF | -0.003259 | 0.9888 |

Tabela 4: Coeficientes de correlação de Spearman entre o nível final de Vitamina D e o nível final dos marcadores ADP, CTX, FAO, LEP, OPG e TNF (siglas).

# Conclusões sintéticas

* Nenhuma das variáveis parece ser normalmente distribuída (Tabela 2). Foram utilizados testes não-paramétricos.
* Todas as diferenças entre o nível basal e final foram altamente significativas (Tabela 3)
* Nenhuma correlação entre a Vitamina D e os marcadores foi significativa (Tabela 4)
* Algumas correlações entre o nível basal e final foram significativas (Figura 4)

A maior parte das correlações não foram significativas, no entanto os testes de diferença foram altamente significativos. Isso pode ser observado quando a correlação discorda da diferença. Por exemplo, FAO tem correlação positiva (rho=0.3164392), embora seu nível final tenha sido significativamente **menor** que o nível basal (p-valor=0.008773), o que pode ser observado na Figura 2.

Minha sugestão é evitar discutir os valores de correlação mas apenas apresentar os resultados, indicando que não são significativos.

Possíveis justificativas para estas observações incluem:

* poucos dados
* outliers que possivelmente deveriam ser removidos

Observe que com tão poucos casos, remover outliers possivelmente tornaria a maior parte das análises inconcludentes. Lembre-se que este trabalho é um estudo de caso, e portanto tudo que foi observado pode e deve ser reportado como tal, sem garantias de generalidade.