

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



Efeito Antioxidante das Framboesas — Um Estudo Longitudinal

Adilson Gracindo Cardoso Delgado

Mestrado em Bioestatística

Trabalho de Projeto orientado por:
Professora Doutora Maria Isabel Calisto Frade Barão

Agradecimentos

Agradeço à Professora Doutora Maria Isabel Barão toda a ajuda e aconselhamento que me permitiram melhorar este trabalho, e à Doutora Maria Eduardo Figueira, Professora de Bromatologia da Faculdade de Farmácia da Universidade de Lisboa, pela disponibilização dos dados. Quero ainda agradecer à Professora Doutora Lisete de Sousa, que foi uma coordenadora exemplar do Mestrado em Bioestatística, sempre disponível para ajudar todos os alunos a ultrapassarem as dificuldades que surgiam.

Estou grato aos meus professores de Estatística no DEIO/FCUL, cujos ensinamentos mudaram a minha vida; em particular, agradeço à Professora Doutora Marília Antunes a grande ajuda no uso eficiente do *R*. O Doutor Paulo Semblano ajudou-me, também, a fazer um *script* reproduzindo os resultados do exemplo de modelação longitudinal dos dados *bodyweight* do capítulo 6 de Cabral e Gonçalves. Agradeço ainda ao Doutor Carlos Ventura, do Departamento de Línguas, Literaturas e Culturas, pelo muito que me ensinou sobre como escrever com correção este trabalho, quer no texto português quer no sumário inglês, e espero não ter estragado a revisão que ele teve a generosidade de ir fazendo capítulo a capítulo quando preparei o texto final usando o acordo ortográfico, a que ele chama desacordo ortográfico, não querendo por isso ver a versão final.

A empresa em que trabalho — Grupo Jerónimo Martins, nos supermercados Pingo Doce — deram-me facilidades e ajudaram a pagar as propinas de Mestrado, pelo que estou muito grato. Agradeço à Dr^a Inês Martins, dos Recursos Humanos, que se interessou pelo meu destino e cuja intervenção esteve na origem da generosidade do Grupo Jerónimo Martins.

Agradeço à minha Mulher e Filho por me aturarem, e aturarem as minhas ausências (a minha filha ainda não deu por isso). E a todos os meus familiares, em particular os meus avôs, que me educaram desde que fiquei órfão de pai, e à minha ex-cunhada Maria José Cardoso Delgado, que me tem tratado como um filho predileto.

Sumário

Procedeu-se à medição de oito variáveis — glicose, triglicéridos, colesterol, HDL, LDL, LDLox, IL-1 e IL-6 — em dez voluntários, em jejum e a intervalos diversos após tomarem um pequeno almoço calórico, em Maio; medições essas que serão consideradas *controles* de medições de *casos*, medições análogas feitas em Junho, depois de esses voluntários terem uma dieta rica em framboesas, de reconhecido poder antioxidante, durante um mês. O desafio foi proceder à análise estatística dos dados, no que se refere ao efeito protetor das framboesas.

No Capítulo 1 apresentamos os dados disponibilizados, referimos o padrão de evolução das variáveis e referimos, de forma muito genérica o conceito de dados longitudinais. Abordamos ainda uma questão prévia importante, a possibilidade de um ajustamento normal, por ser relevante para a análise de dados longitudinais.

No Capítulo 2 fazemos uma análise inicial dos dados, centrada na investigação, explicitamente pedida, da eventual semelhança de casos e auto-controles *em jejum*. Usamos métodos gráficos reveladores, em particular caras de Chernoff, e procedemos também a testes formais, quer admitindo normalidade sempre que tal não foi rejeitado, quer numa perspetiva não paramétrica.

No Capítulo 3 são analisados os perfis de evolução das variáveis, e no Capítulo 4 a evolução de cada voluntário, fazendo-se sempre comentários críticos sobre a coleção de dados que foi disponibilizada.

No Capítulo 5 procede-se a uma análise das médias relevantes, que evidenciam que há aparentemente resultados benéficos daquela dieta rica em antioxidantes: subida do “bom” colesterol HDL, descida do “mau” colesterol LDL (e muito particularmente do LDLox), e descida e regularização das interleucinas, que são indicadores de inflamação.

No Capítulo 6 inspiramo-nos nos exemplos tratados por Cabral e Gonçalves (2012, cap. 6) e Ramsey and Schaffer(2002, cap. 16), sobre análise de dados longitudinais e de medições repetidas, constatando porém que os dados disponibilizados não permitem qualquer das análises referidas naquelas fontes. Foi feito um estudo sistemático de modelos mistos, usando-se o critério de Akaike para selecionar os que proporcionam melhor ajustamento.

O Capítulo 7 regista os comentários finais que nos parecem relevantes.

Palavras chave: Dados longitudinais, Modelos mistos, Antioxidantes, Framboesas

Abstract

We have been asked to peer into data collected in May and in June — glucose, triglicerides, cholesterol, HDL, LDL, LDLox, IL-1 and IL-6, in 10 volunteers, taken at relevant periods of time before and after breakfast. The 10 volunteers have a massive intake of cranberries, known to be powerful antioxidants, during the three weeks before the June measurements (cases); the May measurements are considered as self-controls.

Chapter 1 presents the data and their evolution pattern with comments on the metabolic normal values, pointing out that some of the volunteers used seem quite abnormal. Aside from an analysis on changes that occur during the daily measurement times, there is a very elementary presentation on longitudinal data analysis. Since normality is a simplifying assumption in longitudinal data analysis, we also investigate normal fit to the data.

An initial data analysis is carried out in Chapter 2, both using graphical revealing tools such as profile plots and Chernoff faces, and some formal tests are performed. An important goal is to comment on the possibility of similitude of May and June data before breakfast intake.

In Chapter 3 we analyse the variables profile evolution, and in Chapter 4 the evolution of each volunteer is studied. Critical comments on the poor quality of the data that have been gathered are recurrently presented.

In Chapter 5 we calculate averages and confidence intervals, that indicate that the cranberries diet can contribute to reinforce the good HDL cholesterol and to lower the bad LDL cholesterol and LDLox, and to reduce as well the inflammation indicators IL-1 and IL-6.

Chapter 6 refers the analysis of longitudinal data made in Cabral and Gonçalves (2012, chap. 6) and of repeated measurements in Ramsey and Schaffer (2001, chap. 16), concluding that the data at hand have been gathered in a way that precludes similar analyses. A systematic investigation of mixed models is carried out, and the Akaike criteriom is used to decide which of them provide a better fit.

In Chapter 7 we make some concluding comments.

Keywords: Longitudinal data, Mixed models, Antioxidants, Raspberries

Nota Prévia

Os dados disponibilizados são poucos (foram obtidos usando uma amostra com apenas 10 indivíduos); aquela amostra não é aleatória (são “voluntários”), não parece ter havido critérios de inclusão/exclusão dos voluntários no estudo, e como aparentemente alguns deles estão bastante doentes há a possibilidade de tomarem medicamentos podendo assim haver confundimento incontrolado. Variantes destes comentários serão repetidamente feitos, para que nunca haja a tentação de confundir os resultados de uma análise de dados académica com resultados inferenciais sobre uma população.

Por outro lado, o mesmo tratamento (uso de antioxidantes — framboesas — entre Maio e Junho, originando respetivamente dados de controle e dados experimentais) é aplicado aos 10 voluntários, não se podendo pois concluir se os resultados se devem à dieta ou à simples passagem do tempo (ou outros fatores de confundimento). Também não se esperou um período de *washout* para proceder a um terceiro conjunto de medições depois de a dieta antioxidante ser interrompida, o que seria uma outra forma de analisar objetivamente se a dieta de framboesas tem de facto efeito no controle das lipoproteínas e das inflamações vasculares que lhe são genericamente atribuídas.

Nestas condições procuramos mostrar a capacidade de abordar criticamente a análise dos dados, usando uma variedade de métodos gráficos e de abordagens mais formais, e de usar *software* diversificado (EXCEL, SPSS, R e RStudio), alertando sempre para o carácter de estudo piloto que este estudo não pode ultrapassar.

Procurei na *www* indicações sobre as variáveis do estudo (glicose, triglicéridos, colesterol, HDL, LDL, LDLox, IL-1, IL-6), nomeadamente os níveis séricos considerados normais e os perigos associados a valores fora dos intervalos de confiança indicados. Mas a informação que se encontra na *net* nem sempre é confiável, pelo que alguns dos comentários que faço podem ser, involuntariamente, sem sentido.

Índice

1	Introdução	1
1.1	Os dados	1
1.1.1	Glucose	2
1.1.2	Triglicéridos	2
1.1.3	Colesterol	3
1.1.4	HDL	3
1.1.5	LDL	4
1.1.6	LDLox	4
1.1.7	Interleucina-1	5
1.1.8	Interleucina-6	5
1.1.9	Comentário	5
1.2	Padrões de Evolução das Variáveis	6
1.3	Dados longitudinais	8
1.4	Questão prévias — normalidade?	9
1.5	Conclusões provisórias	10
2	Análise Inicial dos Dados	15
2.1	Os perfis em jejum dos 10 voluntários são idênticos?	16
2.2	Houve alteração dos perfis?	19
2.3	Testes formais	21
2.3.1	Amostra	22
2.3.2	Comparações emparelhadas	22
3	Perfis	31
4	Evolução de cada voluntário	37
4.1	Dados individuais	37
4.2	Correlação	43
4.3	Retorno às Caras de Chernoff	45
5	Gráficos das médias	49
5.1	Gráficos: Médias com Barras de Confiança	49
5.2	Gráficos: Evolução das Variáveis nos Voluntários Cotejada com Médias e Limites de Confiança	51

6	Análise de Dados Longitudinais	53
6.1	IL-1	54
6.2	Conclusões provisórias	61
6.3	Comentário crítico	61
7	Considerações Finais	77

Lista de Tabelas

1.1	Glicose	6
1.2	Triglicéridos	6
1.3	Colesterol	7
1.4	LDL	7
1.5	HDL	7
1.6	LDLox	8
1.7	IL-1	8
1.8	IL-6	8
1.9	Dados	10
2.1	Perfis em jejum em Maio	16
2.2	Perfis em jejum em Junho	16
2.3	Parametrização das faces de Chernoff	18
2.4	Perfis estandardizados em jejum em Maio	20
2.5	Perfis estandardizados em jejum em Junho	20
2.6	<i>Ranks</i> em jejum em Maio	22
4.1	Voluntário 1	37
4.2	Voluntário 2	37
4.3	Voluntário 3	38
4.4	Voluntário 4	38
4.5	Voluntário 5	38
4.6	Voluntário 6	39
4.7	Voluntário 7	39
4.8	Voluntário 8	39
4.9	Voluntário 9	40
4.10	Voluntário 10	40
4.11	Registos às 6 h (após imputação do valor da glicose).	41
4.12	Valores padronizados às 6 h (após imputação do valor da glicose).	42
4.13	Correlações de Pearson	45

Lista de Figuras

2.1	Voluntários 1 a 5, Maio, em jejum	18
2.2	Voluntários 6 a 10, Maio, em jejum	18
2.3	Voluntários 1 a 5, Junho, em jejum	19
2.4	Voluntários 6 a 10, Junho, em jejum	19
2.5	Comparação da Glucose em Maio e Junho (1)	26
2.6	Comparação da Glucose em Maio e Junho (2)	26
2.7	Comparação dos Triglicéridos em Maio e Junho (1)	26
2.8	Comparação dos Triglicéridos em Maio e Junho (2)	26
2.9	Comparação do Colesterol em Maio e Junho (1)	27
2.10	Comparação do Colesterol em Maio e Junho (2)	27
2.11	Comparação das LDL em Maio e Junho (1)	27
2.12	Comparação das LDL em Maio e Junho (2)	27
2.13	Comparação das HDL em Maio e Junho (1)	27
2.14	Comparação das HDL em Maio e Junho (2)	28
2.15	Comparação das LDLox em Maio e Junho (1)	28
2.16	Comparação das LDLox em Maio e Junho (2)	28
2.17	Comparação da IL-1 em Maio e Junho (1)	28
2.18	Comparação da IL-1 em Maio e Junho (2)	29
2.19	Comparação da IL-6 em Maio e Junho (1)	29
2.20	Comparação da IL-6 em Maio e Junho (2)	29
4.1	Voluntário 1 após 6h, Maio e Junho	41
4.2	Voluntário 2 após 6h, Maio e Junho	41
4.3	Voluntário 3 após 6h, Maio e Junho	42
4.4	Voluntário 4 após 6h, Maio e Junho	42
4.5	Voluntário 5 após 6h, Maio e Junho	42
4.6	Voluntário 6 após 6h, Maio e Junho	43
4.7	Voluntário 7 após 6h, Maio e Junho	43
4.8	Voluntário 8 após 6h, Maio e Junho	43
4.9	Voluntário 9 após 6h, Maio e Junho	43
4.10	Voluntário 10 após 6h, Maio e Junho	43
5.1	Glicose	51
5.2	Triglicéridos	51
5.3	Colesterol	51
5.4	LDL	51

5.5	HDL	52
5.6	LDL _{ox}	52
5.7	IL-1	52
5.8	IL-6	52
6.1	Citação de Ramsey and Schaffer	54
6.2	Varição de IL-1 ao longo do tempo. Além da representação do que se passa com cada voluntário, representa-se a negro grosso as linhas unindo os valores médios e limites dos intervalos de confiança para a média (95%).	55

Capítulo 1

Introdução

Colesterol e triglicéridos são substâncias lipídicas essenciais mas que, combinadas no sangue com proteínas (L — Lipoproteins), em excesso podem depositar-se em placas nas artérias, comprometendo o sistema circulatório e possivelmente causando aterosclerose, aumentando por isso o risco de acidentes vasculares, cardíacos ou cerebrais. Admite-se que as lipoproteínas de elevada densidade (HDL ou o chamado “bom colesterol”) diminuem o risco de acidentes vasculares, enquanto as de baixa densidade (LDL ou “mau colesterol”) aumentam esse risco. Julga-se que as lipoproteínas de baixa densidade oxidadas — LDLox — aumentam o risco de aterosclerose e conseqüentemente de doença cardíaca e de acidentes vasculares.

A hereditariedade é provavelmente a causa mais importante de colesterol ou triglicéridos elevados, e doenças metabólicas como a diabetes contribuem também para agravamento do nível sérico de colesterol. Uma dieta rica em gorduras piora a situação, e uma dieta equilibrada com poucas gorduras ajuda a reduzir os valores séricos do mau colesterol, dos triglicéridos, do colesterol total e da glucose.

Juntamente com os triglicéridos, o colesterol é uma das principais e essenciais substâncias gordas presentes no sangue, e que tem duas origens: uma parte produzida pelo próprio organismo, em particular pelo fígado, e outra parte obtida através da alimentação, em particular pela ingestão de produtos animais, como a carne, os ovos, e os produtos lácteos.

1.1 Os dados

Procedeu-se à determinação da glucose, triglicérideos, colesterol total, HDL, LDL, LDLox, IL-1 e IL-6 de 10 voluntários, em jejum (tomado com origem, 0 horas, dos tempos de sucessivas medições) e após ingestão de um pequeno almoço abundante, com intervalos de tempo convenientes.

As primeiras medições foram feitas em Maio, e seguidamente durante 3 semanas os voluntários ingeriram diariamente uma dose de 150g de framboesas, consideradas um antioxidante potente — *“Raspberry has an ORAC value (oxygen radical absorbance capacity) of about 4900 $\mu\text{mol TE per 100 grams, crediting it among the top-ranked ORAC fruits.$ ”*(ver <https://www.nutrition-and-you.com/raspberry.html>). Além da ação antioxidante acredita-se que tem efeito benéfico no controle da glucose e de mediadores inflamatórios, como interleucinas. Após essas três semanas de tratamento, procedeu-se em Junho a nova determinação das variáveis acima referidas, em condições idênticas.

Assim, cada paciente foi o seu próprio controle, e relativamente a cada variável dispomos de duas séries de medições repetidas. No entanto, como o mesmo tratamento foi administrado aos 10 voluntários, não vai ser possível concluir se aquelas doses de framboesas causam as alterações que se observam nas variáveis investigadas.

1.1.1 Glucose

No que se refere a glicemia, deve ter valores entre 90 e 110 mg/dL em jejum, e inferiores a 145 mg/dL após alimentação. Em nenhum caso deve descer abaixo de 80 mg/dL, e abaixo de 70 mg/dL há hipoglicemia, perigosa. Doentes diabéticos controlados devem ter valores inferiores a 126 mg/dL em jejum (0h), e hemoglobina glicada inferior a 6%. A glucose elevada no sangue contribui para um aumento do c-LDL e diminuição do c-HDL, para além de prejudicar a parede das artérias.

Glucose em Maio				Glucose em Junho			
Tempo Voluntário	0h	1h	2h	Tempo Voluntário	0h	1h	2h
1	69.0	60.5	67.6	1	65.7	44.1	64.8
2	48.3	33.4	58.8	2	71.9	68.4	93.8
3	56.6	59.1	70.8	3	69.0	69.5	89.2
4	66.2	57.5	64.3	4	67.3	70.3	84.4
5	69.5	62.3	81.0	5	65.1	77.1	69.9
6	78.8	77.7	112.5	6	80.4	103.6	74.9
7	85.9	81.5	88.5	7	89.4	110.5	95.9
8	65.5	71.0	72.4	8	80.5	73.4	66.8
9	73.2	59.9	68.1	9	81.0	92.1	99.7
10	63.8	43.7	71.3	10	80.0	49.9	68.7

Se as informações acima são corretas, observa-se que no início do estudo, em Maio, grande parte destes voluntários têm problemas de hipoglicemia.

Observa-se também que em alguns dos voluntários o valor da glicemia desce após ingestão de um pequeno-almoço rico, o que parece contrariar o que seria de esperar.

Em Junho o valor da glucose em jejum parece ter subido em vários voluntários para níveis mais seguros, mas observa-se de novo uma inesperada baixa de valores séricos uma hora após a refeição em alguns dos voluntários, em particular nos voluntários 1 e 10.

1.1.2 Triglicéridos

Após 12 horas de jejum, o nível sérico de triglicéridos deve ser inferior a 150 mg/dL. No caso de historial de doença cardiocerebrovascular deve ser inferior a 100 mg/dL.

Tempo	Triglicéridos em Maio					Triglicéridos em Junho				
	0h	1h	2h	4h	6h	0h	1h	2h	4h	6h
Vol. 1	125.7	193.0	196.9	180.7	141.7	126.3	186.1	249.4	241.9	198.2
Vol. 2	83.7	86.4	119.6	148.8	87.4	142.5	156.0	239.0	172.4	165.8
Vol. 3	177.4	220.6	293.3	363.6	231.0	219.3	263.5	328.3	322.9	299.8
Vol. 4	85.8	155.6	189.7	272.8	112.0	88.0	112.5	166.2	143.4	113.0
Vol. 5	94.6	124.0	163.3	210.5	111.6	89.7	103.6	154.0	134.8	127.3
Vol. 6	52.4	65.7	90.5	85.7	122.2	75.5	92.7	166.7	172.2	168.8
Vol. 7	52.9	62.8	64.4	63.8	135.9	54.8	66.7	62.9	75.4	111.6
Vol. 8	85.4	101.5	141	187.7	107.4	77.3	101.0	126.7	108.9	82.8
Vol. 9	69.3	89.5	143.7	152.4	111.0	59.7	106.6	117.9	73.4	70.6
Vol. 10	112.1	147.8	169.9	213.5	142.1	104.0	165.3	156.8	126.2	206.7

Assim, o voluntário 3 ao ser recrutado para a amostra de conveniência apresenta valores excessivamente altos nesta variável, tendo portanto um risco acrescido de problemas cardiovasculares. Essa situação mantém-se em Junho após as três semanas de tratamento com framboesas.

1.1.3 Colesterol

O colesterol total deve ser inferior a 190 mg/dL, mas se estiver entre 190 mg/dL e 220 mg/dL deve poder ser corrigido com dieta e exercício, desde que não haja doenças metabólicas. Acima de 220 mg/dL inspira maiores cuidados, sobretudo se exceder 240 mg/dL. Abaixo de 150 mg/dL também se deve consultar o médico.

Colesterol em Maio			Colesterol em Junho		
Tempo Voluntário	0h	6h	Tempo Voluntário	0h	6h
1	232	196.1	1	230.2	227.0
2	184.5	152.8	2	216.4	213.9
3	219.0	188.6	3	182.2	201.0
4	171.7	188.3	4	174.2	165.8
5	147.9	154.7	5	132.0	143.8
6	158.8	156.9	6	164.5	178.9
7	193.8	163.9	7	171.0	187.7
8	194.3	199.2	8	215.1	213.4
9	158.4	136.3	9	145.6	164.0
10	168.8	159.7	10	174.2	170.8

Entre Maio e Junho observa-se em alguns doentes uma subida dos níveis séricos do colesterol em jejum, quando decerto o tratamento com framboesa tinha como objetivo baixar esse valor.

1.1.4 HDL

O HDL (High Density Lipoproteins) deve ser superior a 40 mg/dL em homens e superior a 45 mg/dL em mulheres. Observa-se assim que alguns dos voluntários recrutados têm valores demasiado baixos de HDL (e não foi indicado o sexo dos voluntários).

HDL em Maio			HDL em Junho		
Tempo Voluntário	0h	6h	Tempo Voluntário	0h	6h
1	63.4	50.3	1	72.3	67.8
2	69.1	57.3	2	76.3	67.3
3	18.8	10.5	3	23.0	21.5
4	55.6	46.3	4	61.4	60.8
5	78.8	54.6	5	79.6	78.0
6	63.7	45.7	6	59.4	58.5
7	54.8	42.4	7	46.2	49.0
8	71.4	64.8	8	83.7	78.3
9	38.5	26.7	9	35.3	36.7
10	35.0	25.7	10	37.9	33.7

De novo se observa que o voluntário 3 apresenta valores muito anormais.

No que se refere às medições 6 horas após pequeno almoço, os valores em Junho são mais elevados do que em Maio, parecendo portanto que o tratamento teve um efeito benéfico. No entanto, no que se refere aos valores em jejum houve pelo contrário um abaixamento no que se refere aos voluntários 6, 7 e 9.

1.1.5 LDL

O LDL (Low Density Lipoproteins) deve ser inferior a 115mg/dL. Esta recomendação é mais estrita — 100 mg/dL (ou mesmo 70 mg/dL) — em doentes cardio-vasculares.

LDL em Maio			LDL em Junho		
Tempo Voluntário	0h	6h	Tempo Voluntário	0h	6h
1	102.7	115.3	1	106.7	103.4
2	97.1	76.2	2	99.8	98.2
3	176.3	107.4	3	156.7	134.9
4	103.2	99.7	4	108.4	102.5
5	51.2	77.1	5	61.2	69.2
6	85.4	71.6	6	79.2	77.3
7	132.4	102.7	7	113.6	122.9
8	106.3	101.3	8	106.3	111.3
9	101.5	79.4	9	99.8	85.7
10	112.8	121.3	10	123.4	116.0

Não há informação sobre se alguns dos voluntários são doentes cardio-vasculares. De qualquer forma, mais uma vez o voluntário 3 se evidencia por ter valores indesejavelmente altos de LDL.

Observe-se que nas tabelas fornecidas o valor do voluntário 2 em Junho, 6h, era 982, um valor discrepante que nos pareceu dever ser corrigido para 98.2.

No que se refere aos níveis em jejum o tratamento parece ter tido o efeito secundário perverso de aumentar o “mau” colesterol nos voluntários 1, 2, 4, 5 e 10. Já no que se refere aos valores 6 horas após ingestão do pequeno almoço não se nota um padrão definido.

1.1.6 LDLox

O LDLox elevado é ainda mais perigoso do que o LDL (mas não encontrei valores de referência); é aceite que pode causar inflamação nas artérias, contribuindo para a formação de placas e aterosclerose.

LDLox em Maio				LDLox em Junho			
Tempo Voluntário	0h	4h	6h	Tempo Voluntário	0h	4h	6h
1	51.6	85.10	49.7	1	42.2	50.9	37.0
2	42.8	61.7	43.7	2	39.8	44.0	39.2
3	47.5	71.1	50.0	3	43.2	49.1	42.3
4	44.2	68.6	46.3	4	38.6	43.3	38.0
5	50.4	79.8	54.2	5	30.0	40.7	31.0
6	41.1	71.2	43.4	6	30.3	52.40	29.7
7	55.3	75.3	55.0	7	37.4	52.6	35.0
8	36.6	51.3	32.1	8	32.0	45.30	33.8
9	40.0	56.0	35.4	9	29.4	42.1	27.7
10	42.0	59.8	54.0	10	28.70	37.70	26.2

Observa-se que na generalidade das situações as medições em Junho têm valores mais baixos do que em Maio.

1.1.7 Interleucina-1

A família de IL-1 é um grupo de 11 citosinas que têm um papel importante na regulação de reações de imunidade ou inflamatórias no caso de exposição a inflamações. Não consegui ter informação sobre os níveis séricos adequados, nem se era desejável que o tratamento os aumentasse ou diminuísse. Na falta dessa informação, vou assumir que seria desejável baixar.

O facto de as determinações em Maio 6h após a refeição serem feitas com dois decimais para alguns dos voluntários parece indicar uma alteração no protocolo da análise.

IL-1 em Maio						IL-1 em Junho					
Tempo Voluntário	0h	1h	2h	4h	6h	Tempo Voluntário	0h	1h	2h	4h	6h
1	4.6	4.8	4.9	4.0	4.65	1	2.1	3.0	3.7	3.5	3.0
2	1.5	3.9	3.0	2.5	2.21	2	2.2	2.5	3.5	3.2	3.9
3	1.7	2.7	2.0	2.2	2.9	3	2.9	3.2	2.3	2.3	2.0
4	2.7	2.8	3.0	3.1	3.3	4	2.5	2.8	2.8	2.7	2.5
5	1.9	2.3	2.1	2.5	3.1	5	1.8	2.2	3.1	3.2	2.7
6	3.3	2.9	1.3	2.5	1.48	6	2.0	2.3	2.3	2.9	1.7
7	2.9	3.7	2.1	2.4	2.97	7	3.5	2.9	2.4	2.0	1.9
8	3.4	2.2	2.5	1.6	2.44	8	2.1	2.2	2.0	2.7	2.1
9	3.7	3.8	3.9	6.3	3.86	9	1.6	2.3	1.8	2.1	1.8
10	3.4	2.6	4.3	4.6	4.0	10	2.5	2.1	2.7	4.1	2.8

1.1.8 Interleucina-6

A IL-6 actua como uma citosina pro-inflamatória e como uma miocina anti-inflamatória. Não consegui encontrar níveis de referência para os valores séricos desejáveis, e o duplo papel que tem em processos inflamatórios contribui para não me esclarecer se seria desejável que o tratamento com framboesas diminuísse ou aumentasse os valores séricos. Na falta dessa informação, vou assumir que seria desejável baixar.

IL-6 em Maio					IL-6 em Junho				
Tempo Voluntário	0h	2h	4h	6h	Tempo Voluntário	0h	2h	4h	6h
1	1.71	2.32	2.57	2.73	1	2.13	2.45	2.46	2.99
2	2.02	3.40	3.62	4.31	2	1.86	2.26	2.84	2.00
3	1.35	2.46	3.31	3.70	3	1.04	1.47	1.53	1.29
4	1.66	2.8	3.65	4.08	4	1.56	2.02	2.12	1.67
5	1.87	2.34	3.43	4.30	5	1.43	1.93	2.10	1.77
6	2.00	2.09	3.02	5.20	6	2.12	2.52	3.46	2.44
7	1.59	1.60	2.00	2.34	7	0.85	1.13	2.74	1.96
8	1.50	2.02	2.83	3.24	8	0.54	0.73	1.69	1.34
9	1.58	3.10	4.10	3.80	9	1.62	1.91	2.74	2.8
10	0.55	1.10	2.34	3.52	10	0.32	0.41	1.10	0.24

1.1.9 Comentário

Observamos assim que se tem uma amostra de tamanho muito pequeno e de conveniência, que parece não ter havido critérios de inclusão/exclusão — o que poderá trazer confundimento — e que os tempos

de medição das diversas variáveis não foram sempre os mesmos, o que limita as possibilidades da análise por causa dos dados omissos, como se verá no capítulo sobre análise de dados longitudinais.

Por isso este estudo deve ser encarado como estudo piloto inspirador de investigação usando amostras apropriadas.

1.2 Padrões de Evolução das Variáveis

As tabelas 1.1 a 1.8 registam o sentido da variação de cada uma das variáveis nos períodos em que foram medidas.

Tabela 1.1: Glicose

	Maio		1h		2h		Junho		1h		2h
	0h						0h				
Vol. 1		↘		↗				↘		↗	
Vol. 2		↘		↗				↘		↗	
Vol. 3		↗		↗				↗		↗	
Vol. 4		↘		↗				↗		↗	
Vol. 5		↘		↗				↗		↘	
Vol. 6		↘		↗				↗		↘	
Vol. 7		↘		↗				↗		↘	
Vol. 8		↗		↗				↘		↘	
Vol. 9		↘		↗				↗		↗	
Vol. 10		↘		↗				↘		↗	

Tabela 1.2: Triglicéridos

	Maio		1h		2h		4h		6h		Junho		1h		2h		4h		6h
	0h										0h								
Vol. 1		↗		↗		↘		↘				↗		↗		↘		↘	
Vol. 2		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 3		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 4		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 5		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 6		↗		↗		↘		↗				↗		↗		↗		↘	
Vol. 7		↗		↗		↘		↗				↗		↘		↗		↗	
Vol. 8		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 9		↗		↗		↗		↘				↗		↗		↘		↘	
Vol. 10		↗		↗		↗		↘				↗		↘		↘		↗	

Tabela 1.3: Colesterol

	Maio 0h		6h			Junho 0h		6h
Vol. 1		↘					↘	
Vol. 2		↘					↘	
Vol. 3		↘					↗	
Vol. 4		↗					↘	
Vol. 5		↗					↗	
Vol. 6		↘					↗	
Vol. 7		↘					↗	
Vol. 8		↗					↘	
Vol. 9		↘					↗	
Vol. 10		↘					↘	

Tabela 1.4: LDL

	Maio 0h		6h			Junho 0h		6h
Vol. 1		↗					↘	
Vol. 2		↘					↘	
Vol. 3		↘					↘	
Vol. 4		↘					↘	
Vol. 5		↗					↗	
Vol. 6		↘					↘	
Vol. 7		↘					↗	
Vol. 8		↘					↗	
Vol. 9		↘					↘	
Vol. 10		↗					↘	

Tabela 1.5: HDL

	Maio 0h		6h			Junho 0h		6h
Vol. 1		↘					↘	
Vol. 2		↘					↘	
Vol. 3		↘					↘	
Vol. 4		↘					↘	
Vol. 5		↘					↘	
Vol. 6		↘					↘	
Vol. 7		↘					↗	
Vol. 8		↘					↘	
Vol. 9		↘					↗	
Vol. 10		↘					↘	

*

Observa-se que há um padrão bem definido no que se refere a HDL, LDLox e IL-6 — com variações semelhantes em Maio e em Junho — mas mais difícil de interpretar no que se refere às outras variáveis.

Obviamente os eventuais padrões são mais facilmente observáveis no caso de a variável ter poucos registros. É por isso mais interessante observar padrões na IL-6, e a ausência de padrões no colesterol e no LDL.

Tabela 1.6: LDLox

	Maio		4h		6h		Junho		4h		6h
	0h						0h				
Vol. 1		↗		↘				↗		↘	
Vol. 2		↗		↘				↗		↘	
Vol. 3		↗		↘				↗		↘	
Vol. 4		↗		↘				↗		↘	
Vol. 5		↗		↘				↗		↘	
Vol. 6		↗		↘				↗		↘	
Vol. 7		↗		↘				↗		↘	
Vol. 8		↗		↘				↗		↘	
Vol. 9		↗		↘				↗		↘	
Vol. 10		↗		↘				↗		↘	

Tabela 1.7: IL-1

	Maio		1h		2h		4h		6h		Junho		1h		2h		4h		6h
	0h										0h								
Vol. 1		↗		↗		↘		↗				↗		↗		↘		↘	
Vol. 2		↗		↘		↘		↘				↗		↗		↘		↗	
Vol. 3		↗		↘		↗		↗				↗		↘				↘	
Vol. 4		↗		↗		↗		↗				↗				↘		↗	
Vol. 5		↗		↘		↗		↗				↗		↗		↗		↘	
Vol. 6		↘		↘		↗		↘				↗				↗		↘	
Vol. 7		↗		↘		↗		↗				↘		↘		↘		↘	
Vol. 8		↘		↗		↘		↗				↗		↘		↗		↘	
Vol. 9		↗		↗		↗		↘				↗		↘		↗		↘	
Vol. 10		↘		↗		↗		↘				↘		↗		↗		↘	

Tabela 1.8: IL-6

	Maio		2h		4h		6h		Junho		2h		4h		6h
	0h								0h						
Vol. 1		↗		↗		↗				↗		↗		↗	
Vol. 2		↗		↗		↗				↗		↗		↘	
Vol. 3		↗		↗		↗				↗		↗		↘	
Vol. 4		↗		↗		↗				↗		↗		↘	
Vol. 5		↗		↗		↗				↗		↗		↘	
Vol. 6		↗		↗		↗				↗		↗		↘	
Vol. 7		↗		↗		↗				↗		↗		↘	
Vol. 8		↗		↗		↗				↗		↗		↘	
Vol. 9		↗		↗		↘				↗		↗		↗	
Vol. 10		↗		↗		↗				↗		↗		↘	

1.3 Dados longitudinais

Os dados disponibilizados são portanto medições repetidas sobre os mesmos 10 indivíduos, em Maio e em Junho, estes últimos após um “tratamento” com framboesas, consideradas anti-oxidantes e anti-inflamatórias. Assim, cada indivíduo é o seu próprio controle (em Maio), e em cada um dos períodos há uma série de medições intervaladas, decerto correspondentes a diversas situações metabólicas específicas, pelo que o tempo é um fator de interesse. Trata-se portanto do que é em geral designado como

“dados longitudinais”.

Para a sua análise usaremos o formalismo exposto em Cabral e Gonçalves (2011, pp. 14–16): y_{it} , $t = 1, 2, \dots, T_i$ é o valor da resposta na ocasião t para a unidade amostral i , $i = 1, 2, \dots, n$ da *variável resposta* Y_{it} ; $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{iT_i})^T$ é o vetor das variáveis respostas para o sujeito amostral i , e a sequência das medições efetuadas nesse indivíduo, $y_i = (y_{i1}, y_{i2}, \dots, y_{iT_i})^T$ é o *perfil do indivíduo*.

Além disso, considera-se quando apropriado um vetor de covariáveis associado a cada variável aleatória resposta Y_i (ver formalismo em Cabral e Gonçalves, 2011, p. 15). No caso de variáveis contínuas, os modelos paramétricos usados na sua análise são em geral baseados no pressuposto de distribuição normal (Cabral e Gonçalves, 2011, p. 16).

No Capítulo 6, apoiando-nos em citações de Cabral e Gonçalves (2011), consideramos que faria eventualmente sentido uma análise longitudinal das variáveis mais diretamente influenciadas pelo tratamento com doses abundantes de framboesa (interleucinas e lipoproteínas) admitindo efeitos aleatórios.

A análise foi feita usando o *software* R, seguindo as indicações de Cabral e Gonçalves, e a informação contida nos manuais do software usado; é portanto um trabalho de índole prática, não havendo razões para detalhar mais do que o que acima ficou sobre o formalismo da modelação de dados longitudinais, pois Cabral e Gonçalves (2011) o fazem com detalhe. Pareceu-nos por isso mais interessante, no referido capítulo 6, começar por um arrolamento sistemático dos modelos para IL-1 usando Mês, Hora e a intereacção entre Mês e Hora, escolhendo depois alguns cuja análise prévia foi mais interessante, seguindo depois para a modelação de IL-1 usando LDLox como covariável influente (o que nos obriga a descartar dados, visto que o LDLox apenas foi registado em jejum, 4h e 6h depois).

O ajustamento preferido é escolhido tendo usando em particular o critério de Akaike, e os valores- p correspondentes aos parâmetros estimados.

1.4 Questão prévias — normalidade?

Há uma questão prévia que vamos abordar: com vista a análises anteriores mais formais, podemos assumir um ajustamento normal? Na análise de dados longitudinais, os modelos paramétricos usados na análise de dados contínuos assumem, como já referimos, normalidade¹.

Decidimos por isso testar a normalidade dos dados, usando o teste de Kolmogorov-Smirnov (usando a modificação de Lilliefors, uma vez que há estimação do valor médio e da variância). O facto de as séries de dados corresponderem aos mesmo 10 indivíduos, e a correlação serial óbvia nas diversas medições de cada uma das variáveis em tempos diferentes deve ser levada em linha de conta, usando por exemplo a correção de Bonferroni, cf. https://en.wikipedia.org/wiki/Bonferroni_correction, onde de qualquer modo há o aviso de que “*The Bonferroni correction can be somewhat conservative if there are a large number of tests and/or the test statistics are positively correlated. The correction also comes at the cost of increasing the probability of producing false negatives, and consequently reducing statistical power.*”.

Note-se no entanto que o uso da correção de Bonferroni não tem um padrão consensual, na mesma https://en.wikipedia.org/wiki/Bonferroni_correction lê-se “*Another criticism concerns the concept of a family of hypotheses. There is not a definitive consensus on how to define a family in all cases. As there is no standard definition, test results may change dramatically, only by modifying the way we consider the hypotheses families.*”. Por isso, nesta fase exploratória inicial, decidimos não as usar

¹ Veja-se a discussão em Cabral e Gonçalves, (2011, 1.4.1), em que se recomenda considerar que na análise de dados longitudinais com respostas contínuas há vantagem em considerar modelos normais.

Os resultados devem ser tomados como apenas indicativos. Usando o *SPSS 23*, a rejeição do ajustamento normal apenas ocorre em três casos.

1. GlucoseMaio2h com $p - value$ 0.025;
2. TriglMaio6h com $p - value$ 0.024;
3. IL-1Maio4h, com $p - value$ 0.023.

1.5 Conclusões provisórias

Um estudo feito com *voluntários* não será totalmente inútil, mas não permite inferir para a população de que foi retirada a coleção de dados. Não parece ter havido qualquer preocupação com critérios de inclusão/exclusão, e o tamanho da coleção de dados parece muito pequeno. Além disso, não há qualquer garantia de ter sido evitado confundimento — o voluntário 3, por exemplo, tem um grau de hipoglicemia e valores de colesterol, LDL, HDL e LDLox tão preocupantes que é razoável pensar que não lhe mandarem tomar medicação específica seria pouco ético, e se assim for isso vai quase certamente ter influência nos valores de Junho, possivelmente mesmo influência muito maior do que a influência do tratamento com framboesas.

Por outro lado, não é claro quais as questões que se pretende investigar, nomeadamente que variável ou variáveis deve(m) ser considerada(s) resposta(s), e que variáveis devem ser consideradas covariáveis. Por outro lado, a apresentação global dos dados na tabela abaixo (codificamos 0 o mês Maio e 1 o mês Junho) mostra que há muitos dados omissos (codificados como *NA – Not Available*), não tendo sido fornecidas medições das variáveis nos mesmos períodos), o que limita a exploração longitudinal.

Tabela 1.9: Dados

VOL	MES	HOR	IL1	IL6	COL	HDL	LDL	LDLox	TRI	GLU
1	0	0	4.6	1.7	232.0	63.4	102.7	51.6	125.7	69.0
2	0	0	1.5	2.0	184.5	69.1	97.1	42.8	83.7	48.3
3	0	0	1.7	1.4	219.0	18.8	176.3	47.5	177.4	56.6
4	0	0	2.7	1.7	171.7	55.6	103.2	44.2	85.8	66.2
5	0	0	1.9	1.9	147.9	78.8	51.2	50.4	94.6	69.5
6	0	0	3.3	2.0	158.8	63.7	85.4	41.1	52.4	78.8
7	0	0	2.9	1.6	193.8	54.8	132.4	55.3	52.9	85.9
8	0	0	3.4	1.5	194.3	71.4	106.3	36.6	85.4	65.5
9	0	0	3.7	1.6	158.4	38.5	101.5	40.0	69.3	73.2
10	0	0	3.4	0.6	168.8	35.0	112.8	42.0	112.1	63.8
1	0	1	4.8	NA	NA	NA	NA	NA	193.0	60.5
2	0	1	3.9	NA	NA	NA	NA	NA	86.4	33.4
3	0	1	2.7	NA	NA	NA	NA	NA	220.6	59.1
4	0	1	2.8	NA	NA	NA	NA	NA	155.6	57.5
5	0	1	2.3	NA	NA	NA	NA	NA	124.0	62.3
6	0	1	2.9	NA	NA	NA	NA	NA	65.7	77.7
7	0	1	3.7	NA	NA	NA	NA	NA	62.8	81.5

8	0	1	2.2	NA	NA	NA	NA	NA	101.5	71.0
9	0	1	3.8	NA	NA	NA	NA	NA	89.5	59.9
10	0	1	2.6	NA	NA	NA	NA	NA	147.8	43.7
1	0	2	4.9	2.3	NA	NA	NA	NA	196.9	67.6
2	0	2	3.0	3.4	NA	NA	NA	NA	119.6	58.8
3	0	2	2.0	2.5	NA	NA	NA	NA	293.3	70.8
4	0	2	3.0	2.8	NA	NA	NA	NA	189.7	64.3
5	0	2	2.1	2.3	NA	NA	NA	NA	163.3	81.0
6	0	2	1.3	2.1	NA	NA	NA	NA	90.5	112.5
7	0	2	2.1	1.6	NA	NA	NA	NA	64.4	88.5
8	0	2	2.5	2.0	NA	NA	NA	NA	141.0	72.4
9	0	2	3.9	3.1	NA	NA	NA	NA	143.7	68.1
10	0	2	4.3	1.1	NA	NA	NA	NA	169.9	71.3
1	0	4	4.0	2.6	NA	NA	NA	85.1	180.7	NA
2	0	4	2.5	3.6	NA	NA	NA	61.7	148.8	NA
3	0	4	2.2	3.3	NA	NA	NA	71.1	363.6	NA
4	0	4	3.1	3.7	NA	NA	NA	68.6	272.8	NA
5	0	4	2.5	3.4	NA	NA	NA	79.8	210.5	NA
6	0	4	2.5	3.0	NA	NA	NA	71.2	85.7	NA
7	0	4	2.4	2.0	NA	NA	NA	75.3	63.8	NA
8	0	4	1.6	2.8	NA	NA	NA	51.3	187.7	NA
9	0	4	6.3	4.1	NA	NA	NA	56.0	152.4	NA
10	0	4	4.6	2.3	NA	NA	NA	59.8	213.5	NA
1	0	6	4.65	2.7	196.1	50.3	115.3	49.7	141.7	NA
2	0	6	2.21	4.3	152.8	57.3	76.2	43.7	87.4	NA
3	0	6	2.9	3.7	188.6	10.5	107.4	50.0	231.0	NA
4	0	6	3.3	4.1	188.3	46.3	99.7	46.3	112.0	NA
5	0	6	3.1	4.3	154.7	54.6	77.1	54.2	111.6	NA
6	0	6	1.48	5.2	156.9	45.7	71.6	43.4	122.2	NA
7	0	6	2.97	2.3	163.9	42.4	102.7	55.0	135.9	NA
8	0	6	2.44	3.2	199.2	64.8	101.3	32.1	107.4	NA
9	0	6	3.86	3.8	136.3	26.7	79.4	35.4	111.0	NA
10	0	6	4	3.5	159.7	25.7	121.3	54.0	142.1	NA
1	1	0	2.1	2.1	230.2	72.3	106.7	42.2	126.3	65.7
2	1	0	2.2	1.9	216.4	76.3	99.8	39.8	142.5	71.9
3	1	0	2.9	1.0	182.2	23.0	156.7	43.2	219.3	69.0
4	1	0	2.5	1.6	174.2	61.4	108.4	38.6	88.0	67.3
5	1	0	1.8	1.4	132.0	79.6	61.2	30.0	89.7	65.1
6	1	0	2.0	2.1	164.5	59.4	79.2	30.3	75.5	80.4
7	1	0	3.5	0.9	171.0	46.2	113.6	37.4	54.8	89.4
8	1	0	2.1	0.5	215.1	83.7	106.3	32.0	77.3	80.5
9	1	0	1.6	1.6	145.6	35.3	99.8	29.4	59.7	81.0
10	1	0	2.5	0.3	174.2	37.9	123.4	28.7	104.0	80.0

1	1	1	3.0	NA	NA	NA	NA	NA	186.1	44.1
2	1	1	2.5	NA	NA	NA	NA	NA	156.0	68.4
3	1	1	3.2	NA	NA	NA	NA	NA	263.5	69.5
4	1	1	2.8	NA	NA	NA	NA	NA	112.5	70.3
5	1	1	2.2	NA	NA	NA	NA	NA	103.6	77.1
6	1	1	2.3	NA	NA	NA	NA	NA	92.7	103.6
7	1	1	2.9	NA	NA	NA	NA	NA	66.7	110.5
8	1	1	2.2	NA	NA	NA	NA	NA	101.0	73.4
9	1	1	2.3	NA	NA	NA	NA	NA	106.6	92.1
10	1	1	2.1	NA	NA	NA	NA	NA	165.3	49.9
1	1	2	3.7	2.5	NA	NA	NA	NA	249.4	64.8
2	1	2	3.5	2.3	NA	NA	NA	NA	239.0	93.8
3	1	2	2.3	1.5	NA	NA	NA	NA	328.3	89.2
4	1	2	2.8	2.0	NA	NA	NA	NA	166.2	84.4
5	1	2	3.1	1.9	NA	NA	NA	NA	154.0	69.9
6	1	2	2.3	2.5	NA	NA	NA	NA	166.7	74.9
7	1	2	2.4	1.1	NA	NA	NA	NA	62.9	95.9
8	1	2	2.0	0.7	NA	NA	NA	NA	126.7	66.8
9	1	2	1.8	1.9	NA	NA	NA	NA	117.9	99.7
10	1	2	2.7	0.4	NA	NA	NA	NA	156.8	68.7
1	1	4	3.5	2.5	NA	NA	NA	50.9	241.9	NA
2	1	4	3.2	2.8	NA	NA	NA	44.0	172.4	NA
3	1	4	2.3	1.5	NA	NA	NA	49.1	322.9	NA
4	1	4	2.7	2.1	NA	NA	NA	43.3	143.4	NA
5	1	4	3.2	2.1	NA	NA	NA	40.7	134.8	NA
6	1	4	2.9	3.5	NA	NA	NA	52.4	172.2	NA
7	1	4	2.0	2.7	NA	NA	NA	52.6	75.4	NA
8	1	4	2.7	1.7	NA	NA	NA	45.3	108.9	NA
9	1	4	2.1	2.7	NA	NA	NA	42.1	73.4	NA
10	1	4	4.1	1.1	NA	NA	NA	37.7	126.2	NA
1	1	6	3.0	3.0	227.0	67.8	103.4	37.0	198.2	NA
2	1	6	3.9	2.0	213.9	67.3	98.2	39.2	165.8	NA
3	1	6	2.0	1.3	201.0	21.5	134.9	42.3	299.8	NA
4	1	6	2.5	1.7	165.8	60.8	102.5	38.0	113.0	NA
5	1	6	2.7	1.8	143.8	78.0	69.2	31.0	127.3	NA
6	1	6	1.7	2.4	178.9	58.5	77.3	29.7	168.8	NA
7	1	6	1.9	2.0	187.7	49.0	122.9	35.0	111.6	NA
8	1	6	2.1	1.3	213.4	78.3	111.3	33.8	82.8	NA
9	1	6	1.8	2.8	164.0	36.7	85.7	27.7	70.6	NA
10	1	6	2.8	0.2	170.8	33.7	116.0	26.2	206.7	NA

Nas leituras que fiz na internet encontrei várias vezes a informação de que às interleucinas se ajusta-

riam modelos lognormais. Vimos atrás que o ajustamento por modelos normais não foi rejeitado senão em três casos (que não eram da classe das interleucinas), mas tal deve-se com certeza ao facto de as amostras terem dimensão 10, muito baixa.

Capítulo 2

Análise Inicial dos Dados

Pretende o presente capítulo fazer uma análise inicial dos dados que nos indique quais as análises mais formais que podem esclarecer o efeito do tratamento com framboesas. Por estas serem antioxidantes fortes, espera-se que reduzam o LDLox, o que pode ter como consequência redução da formação de placas de colesterol nas artérias, que podem fissurar e provocar inflamação e, ao soltar-se, formar coágulos e levar a acidentes vasculares cerebrais ou cardíacos. Se este raciocínio estiver certo, é de esperar redução dos valores das interleucinas 1 e 6, por serem indicadoras da gravidade da inflamação.

Sublinhamos de novo que o facto de se ter usado uma coleção de dados de voluntários não permite fazer inferências sobre a população alvo; é portanto um estudo piloto, que pode inspirar hipóteses a investigar com amostras apropriadas representativas de alguma população alvo bem definida.

Acresce que o número de voluntários é muito baixo, e como apontado na descrição dos dados, aparentemente alguns deles (em particular o voluntário 3) são bastante “anormais”, não tendo a questão de critérios de inclusão/exclusão sido considerada. Como já referido, é provável que alguns dos voluntários estejam em tratamento, de que pode resultar confundimento.

De qualquer modo, vamos abordar as seguintes questões:

1. Os perfis iniciais (0h em Maio) dos 10 voluntários são semelhantes?
2. O perfil de cada um dos candidatos entre Maio e Junho alterou-se, isto é: a ingestão de framboesas durante 3 semanas teve algum efeito na alteração dos perfis, em jejum, dos voluntários?
3. O tratamento com framboesas teve um efeito detetável sobre as variáveis em análise?

O tratamento da última questão será abordado apenas no capítulo seguinte, onde se analisará a evolução dos perfis dos voluntários.

Tabela 2.1: Perfis em jejum em Maio

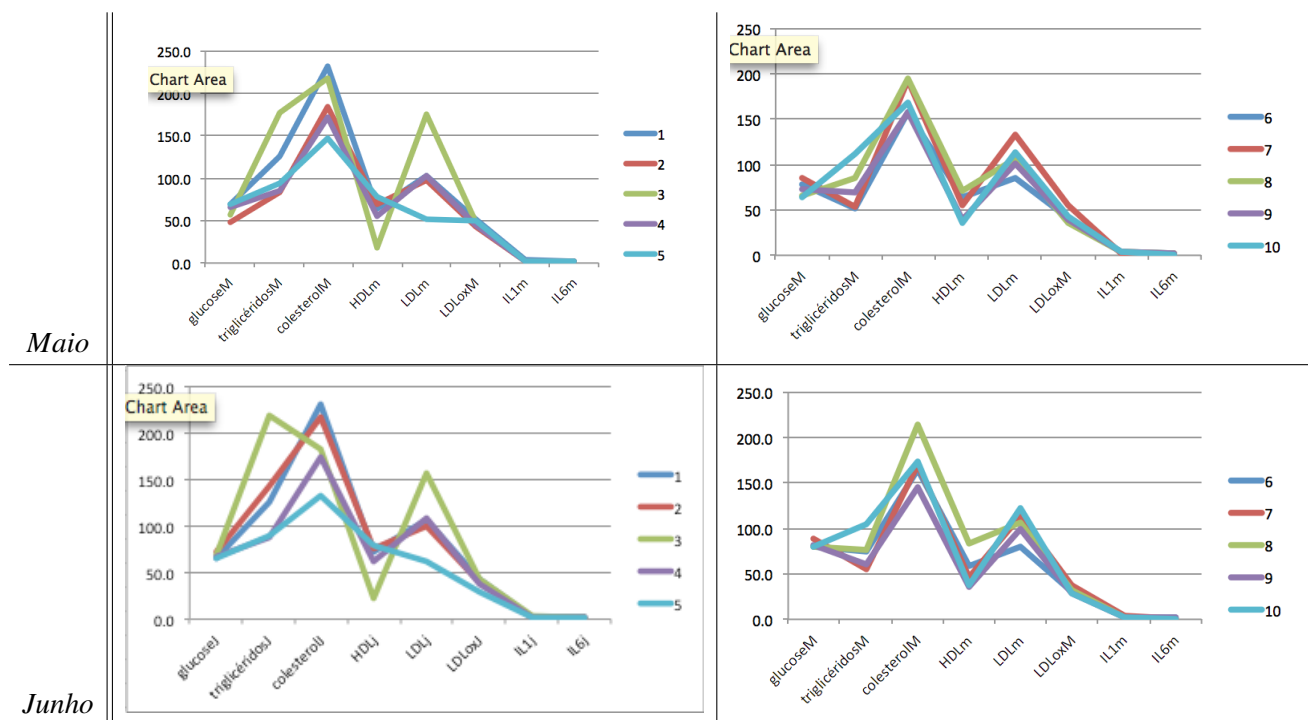
Voluntário	glucose	triglicéridos	colesterol	HDL	LDL	LDL _{ox}	IL1	IL6
1	69.0	125.7	232.0	63.4	102.7	51.6	4.6	1.71
2	48.3	83.7	184.5	69.1	97.1	42.8	1.5	2.02
3	56.6	177.4	219.0	18.8	176.3	47.5	1.7	1.35
4	66.2	85.8	171.7	55.6	103.2	44.2	2.7	1.66
5	69.5	94.6	147.9	78.8	51.2	50.4	1.9	1.87
6	78.8	52.4	158.8	63.7	85.4	41.1	3.3	2.00
7	85.9	52.9	193.8	54.8	132.4	55.3	2.9	1.59
8	65.5	85.4	194.3	71.4	106.3	36.6	3.4	1.5
9	73.2	69.3	158.4	38.5	101.5	40.0	3.7	1.58
10	63.8	112.1	168.8	35.0	112.8	42.0	3.4	0.55
\bar{x}	67.7	93.9	182.9	54.9	106.9	45.2	2.9	1.58
s	10.60	37.37	27.28	18.75	32.01	5.87	0.96	0.421

Tabela 2.2: Perfis em jejum em Junho

Voluntário	glucose	triglicéridos	colesterol	HDL	LDL	LDL _{ox}	IL1	IL6
1	65.7	126.3	230.2	72.3	106.7	42.2	2.1	2.13
2	71.9	142.5	216.4	76.3	99.8	39.8	2.2	1.86
3	69.0	219.3	182.2	23.0	156.7	43.2	2.9	1.04
4	67.3	88.0	174.2	61.4	108.4	38.6	2.5	1.56
5	65.1	89.7	132.0	79.6	61.2	30.0	1.8	1.43
6	80.4	75.5	164.5	59.4	79.2	30.3	2.0	2.12
7	89.4	54.8	171.0	46.2	113.6	37.4	3.5	0.85
8	80.5	77.3	215.1	83.7	106.3	32.0	2.1	0.54
9	81.0	59.7	145.6	35.3	99.8	29.4	1.6	1.62
10	80.0	104.0	174.2	37.9	123.4	28.70	2.5	0.32
\bar{x}	75.0	105.3	180.5	57.5	105.5	35.2	2.3	1.35
s	8.28	51.75	31.54	20.98	25.25	5.65	0.56	0.637

2.1 Os perfis em jejum dos 10 voluntários são idênticos?

A observação dos gráficos de perfis (na matriz de gráficos que apresentamos separamos, para mais simples interpretação visual, na coluna da esquerda estão os correspondentes aos voluntários 1 a 5 e na da direita os dos voluntários 6 a 10) sugere imediatamente que os voluntários 1 e 3 são bastante distintos dos outros.



No entanto, as escalas muito diversas das variáveis não permitem uma análise mais detalhada, pois por exemplo os valores das interleucinas são tão baixos quando comparados com os do colesterol que não se distinguem diferenças no que se refere às interleucinas, que afinal à escala podem ser muito grandes. Portanto esta representação não é conveniente.

Recorremos por isso a uma representação mais expressiva, as faces de Chernoff.



Para tal usamos o *package aplpack* do R, que usa as variáveis para parametrizar os diversos aspetos do gráfico¹.

Uma vez que o *software* considera 15 parâmetros (altura da face, largura da face, estrutura da face, altura da boca, largura da boca, expressão da boca, altura dos olhos, largura dos olhos, altura do cabelo, largura do cabelo, estilo do cabelo, altura do nariz, largura do nariz, altura das orelhas) há uma atribuição sequencial que leva a que a parametrização seja a descrita na tabela 2.3.

A expressividade das faces de Chernoff pode ser aumentada usando cores.

¹ Não podemos deixar de anotar de novo que não nos foi dada qualquer informação sobre variáveis e covariáveis, e sobre a importância relativa que têm nesta investigação, e por isso usámos as variáveis na sequência glucose – triglicéridos – colesterol – HDL – LDL – LDLox – IL-1 – IL-6, de acordo com a numeração das variáveis na folha EXCEL que foi fornecida. Obviamente o aspeto das faces de Chernoff seria muito diferente se usássemos uma outra ordenação das variáveis, o que nos leva a sublinhar de novo que este estudo está a ser feito por mim sem um objetivo definido.

Tabela 2.3: Parametrização das faces de Chernoff

Variável	Características da face	Número
glucose	altura da face, altura do cabelo	1, 9
triglicéridos	largura da face, largura do cabelo	2, 10
colesterol	estrutura da face, estilo do cabelo	3, 11
HDL	altura da boca, altura do nariz	4, 12
LDL	largura da boca, largura do nariz	5, 13
LDLox	expressão da boca, largura das orelhas	6, 14
IL-1	altura dos olhos, altura das orelhas	7, 15
IL-6	largura dos olhos	8



Nos gráficos que adiante apresentamos as cores são parametrizadas por: olhos e íris (7, 8), lábios (média de 1, 2, 3), orelhas (média de 14,15) nariz (média de 12, 13), cabelo (média de 9, 10, 11), face (média de 1, 2).

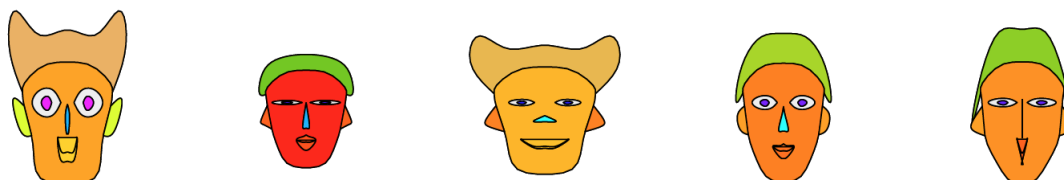


Figura 2.1: Voluntários 1 a 5, Maio, em jejum



Figura 2.2: Voluntários 6 a 10, Maio, em jejum

Observando as figuras 2.1 e 2.2 a grande variabilidade dos voluntários é bem evidente, sobressaindo a “anormalidade” dos voluntários 1 e 3, como anteriormente, mas notando-se agora também diversidade marcada no que se refere às interleucinas, pelo aspeto e cor dos olhos e das orelhas.

As faces de Chernoff das figuras 2.3 e 2.4, correspondentes ao período em jejum em Junho, após tratamento com doses diárias de framboesas, mostram que a diversidade da coleção de dados persiste:

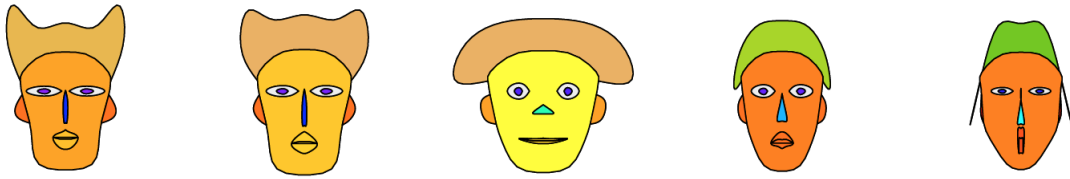


Figura 2.3: Voluntários 1 a 5, Junho, em jejum

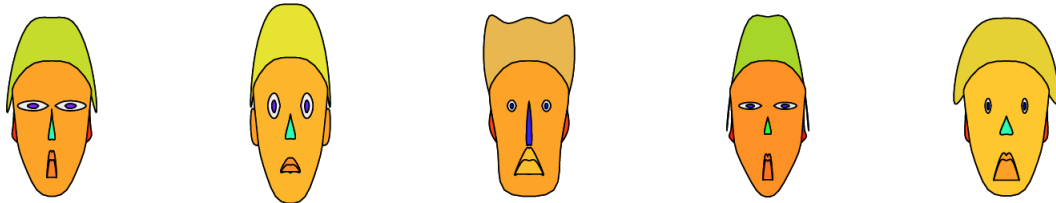
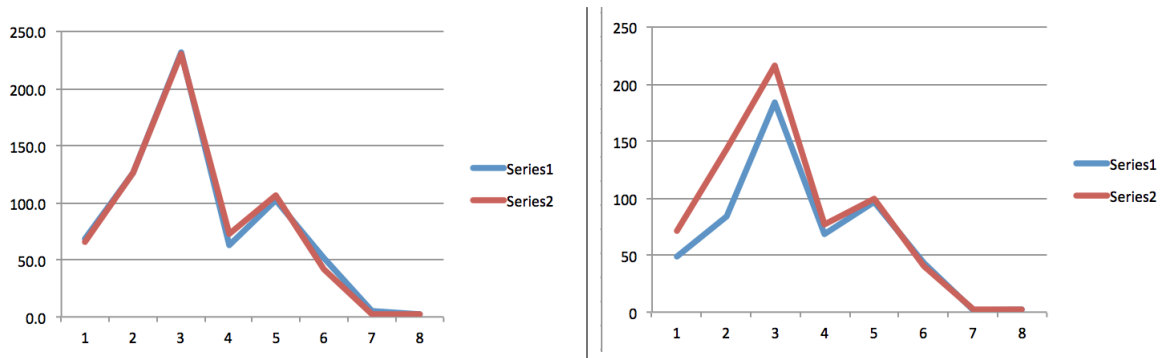


Figura 2.4: Voluntários 6 a 10, Junho, em jejum

O voluntário 3 continua a distinguir-se de forma mais evidente, mas a heterogeneidade muito evidente de todos os dados leva-nos a duvidar da representatividade desta “amostra”, e portanto da validade de conclusões.

2.2 Houve alteração dos perfis?

Observe-se a figura abaixo, com os perfis dos voluntários 1 e 2, em jejum, em Maio e em Junho:



Os gráficos sugerem que no caso do voluntário 1 não houve alteração, enquanto no que respeita o voluntário 2 houve um aumento dos níveis séricos de glucose, triglicéridos e colesterol total.

Observe-se no entanto que a escala muito diversa das variáveis em estudo mascara as possíveis alterações, sobretudo no que se refere a IL-1 e IL-6.

Vamos por isso considerar em vez dos dados brutos os dados standardizados, porque a harmonização das escala vai evidenciar as alterações de perfis:

A representação gráfica dos perfis, apesar de ser uma técnica elementar, mostra que não é possível perceber um padrão de alterações:

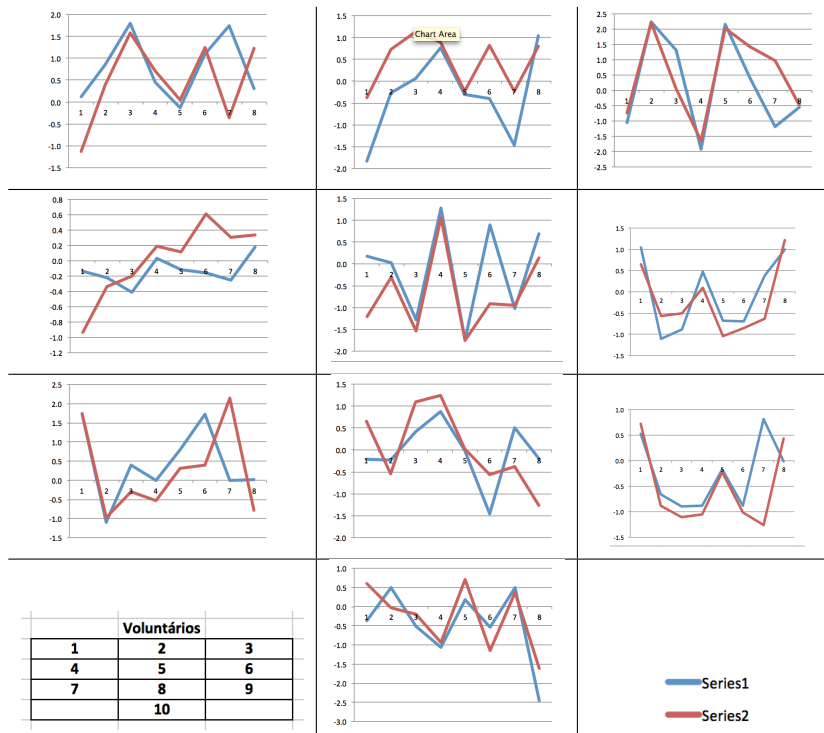
Tabela 2.4: Perfis estandardizados em jejum em Maio

Voluntário	glucose	triglicéridos	colesterol	HDL	LDL	LDLox	IL1	IL6
1	0.1	0.9	1.8	0.5	-0.1	1.1	1.7	0.3
2	-1.8	-0.3	0.1	0.8	-0.3	-0.4	-1.5	1.0
3	-1.0	2.2	1.3	-1.9	2.2	0.4	-1.2	-0.6
4	-0.1	-0.2	-0.4	0.0	-0.1	-0.2	-0.2	0.2
5	0.2	0.0	-1.3	1.3	-1.7	0.9	-1.0	0.7
6	1.0	-1.1	-0.9	0.5	-0.7	-0.7	0.4	1.0
7	1.7	-1.1	0.4	0.0	0.8	1.7	0.0	0.0
8	-0.2	-0.2	0.4	0.9	0.0	-1.5	0.5	-0.2
9	0.5	-0.7	-0.9	-0.9	-0.2	-0.9	0.8	0.0
10	-0.4	0.5	-0.5	-1.1	0.2	-0.5	0.5	-2.5

Tabela 2.5: Perfis estandardizados em jejum em Junho

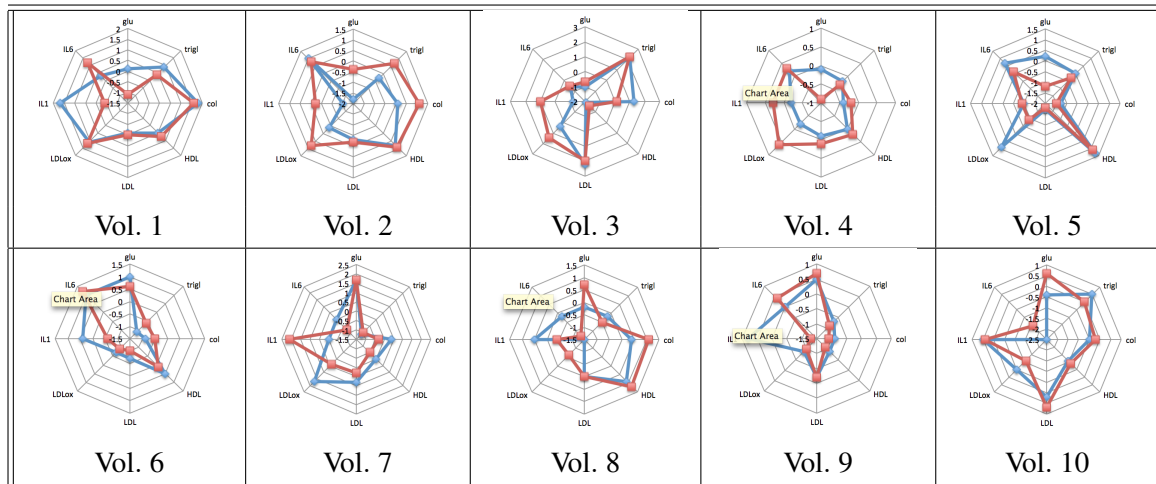
Voluntário	glucose	triglicéridos	colesterol	HDL	LDL	LDLox	IL1	IL6
1	-1.1	0.4	1.6	0.7	0.0	1.2	-0.4	1.2
2	-0.4	0.7	1.1	0.9	-0.2	0.8	-0.2	0.8
3	-0.7	2.2	0.1	-1.6	2.0	1.4	1.0	-0.5
4	-0.9	-0.3	-0.2	0.2	0.1	0.6	0.3	0.3
5	-1.2	-0.3	-1.5	1.1	-1.8	-0.9	-0.9	0.1
6	0.6	-0.6	-0.5	0.1	-1.0	-0.9	-0.6	1.2
7	1.7	-1.0	-0.3	-0.5	0.3	0.4	2.1	-0.8
8	0.7	-0.5	1.1	1.2	0.0	-0.6	-0.4	-1.3
9	0.7	-0.9	-1.1	-1.1	-0.2	-1.0	-1.3	0.4
10	0.6	0.0	-0.2	-0.9	0.7	-1.1	0.4	-1.6

Alteração dos perfis:



Apesar dos gráficos acima serem bastante expressivos, recorde-se que estamos a usar dados estandarizados para contrariar a compressão das interleucinas, que têm uma escala muito menor.

Porém o facto de se estar a usar dados estandarizados pode mascarar alguns aspetos, que as caras de Chernoff evidenciam quer nas dimensões correspondentes às diversas variáveis quer nas cores. Por isso, e para evitar a monotonia de usar repetidamente os mesmos instrumentos, vamos agora usar gráficos de radar para representar os perfis dos voluntários em Maio e em Junho, em jejum (dados de Maio a azul, de Junho a laranja; no sentido dos ponteiros do relógio, dados padronizados de Glucose, Triglicéridos, Colesterol, HDL, LDL, LDLox, IL-1, IL-6).



Observa-se que há grande diversidade de perfis dos 10 voluntários em jejum, quer em Maio quer em Junho, muito expressiva. Observam-se também em geral grandes alterações nos perfis antes/depois do tratamento, apenas o voluntário 6 parece ter mudado pouco..

2.3 Testes formais

Nas escassas indicações que nos foram dadas sobre o trabalho a realizar, foi explicitamente sugerido que se comparassem os dados de Maio com os de Junho.

Como cada indivíduo é o seu próprio controle, são dados correlacionados, e portanto em todas as situações em que não foi rejeitado um ajustamento normal (o que apenas aconteceu com GlucoseMaio2h com $p - value$ 0.025, TriglMaio6h com $p - value$ 0.024, IL-1Maio4h, com $p - value$ 0.023) vamos usar o teste t para dados emparelhados.

De qualquer forma como há um número muito escasso de dados comparo com os resultados que se obtêm usando o teste de Wilcoxon (com apenas 10 pares de observações pedimos cálculos exatos do valor- p).

Observe-se ainda que não estou a excluir nenhum caso, apesar de achar que o sujeito 3 seria de excluir.

Nos casos em que houve rejeição do ajustamento normal usamos apenas o teste de Wilcoxon, anotando que o teste t não é aplicável.

Os resultados adiante descritos foram obtidos usando SPSS 23 (faz-se reprodução dos *outputs* no fim desta secção). Por outro lado, em face dos comentários já feitos sobre as correções de Bonferroni — nomeadamente o risco de haver rejeições em número excessivo — decidimos não as usar.

2.3.1 Amostra

Temos repetidamente questionado a validade desta coleção de dados. Pareceu-nos por isso que seria interessante investigar a aleatoriedade, usando o teste de Friedman. Como afirmam Mosteller e Rourke (1993, p. 259, secção **Teste de Aleatoriedade; a estatística de Friedman**), “quando nos são fornecidas diversas ordenações do mesmo tipo de itens, esperamos que haja correlação entre as ordens atribuídas. [...] A estatística de Friedman avalia a discordância entre as somas de ordens [...]”. Para tal transformamos as magnitudes observadas na Tabela 2.1 em *ranks* (ascendentes exceto no caso de HDL, em que valores altos são mais saudáveis), obtendo a Tabela 2.6. A soma dos quadrados dos *ranks* atribuídos a cada voluntário é 20045, pelo que o valor observado da estatística de Friedman é $T_F = \frac{12 \times 20045}{8 \times 10 \times 11} - 3 \times 8 \times 11 = 9.341$. Basta observar que este valor é muito próximo do valor esperado 9 da variável χ_9^2 para decidir que não existe evidência para rejeitar a hipótese nula de aleatoriedade. Claro que isso é atribuível ao diminuto tamanho da “amostra”, mas é um ponto a favor a ter em conta.

Tabela 2.6: *Ranks* em jejum em Maio

Voluntário	glucose	triglicéridos	colesterol	HDL	LDL	LDLox	IL1	IL6
1	6	9	10	5	5	9	10	7
2	1	4	6	3	3	5	1	10
3	2	10	9	10	10	7	2	2
4	5	6	5	6	6	6	4	6
5	7	7	1	1	1	8	3	8
6	9	1	3	4	2	3	6	9
7	10	2	7	7	9	10	5	5
8	4	5	8	2	7	1	7	3
9	8	3	2	8	4	2	9	4
10	3	8	4	9	8	4	7	1

2.3.2 Comparações emparelhadas

Nesta fase apenas investigamos se houve variação, deixando para mais tarde, no âmbito do estudo longitudinal, o estudo da evolução das variáveis. Usamos a diferença dos valores entre Maio e Junho, por para este efeito ter a mesma informação que os incrementos entre Maio e Junho.

Mais detalhadamente, dado que podemos emparelhar dados, usamos $x_{iM} - x_{iJ} = d_i$, $i = 1, \dots, 10$ e, assumindo que as observações D_i são de uma população normal D , testamos $H_0 : \mu_D = 0$ vs. $H_1 : \mu_D \neq 0$ (pois excetuando o caso de LDLox, em que podemos esperar que o teste unilateral $H_1 : \mu_D > 0$, não temos conhecimentos para prever se o tratamento com framboesas vai aumentar ou diminuir o valor médio das variáveis). Nos casos em que a normalidade não foi rejeitada, a estatística de teste $T = \sqrt{n} \frac{\bar{D}}{S_D}$ tem sob H_0 distribuição t de Student com 9 graus de liberdade.

De facto, na maior parte dos casos o teste de Lilliefors não rejeitou normalidade; pareceu-nos no entanto que, com amostras de tão baixa dimensão, é aconselhável comparar com o resultado do teste de ranks com sinais de Wilcoxon, em que a hipótese nula é que as amostras emparelhadas têm a mesma localização (anote-se ainda que no caso de normalidade média e mediana, tomada esta como localização central, são iguais. O recurso ao SPSS 23 permite a obtenção imediata de valores- p , e as consequentes decisões que seguidamente indicamos.

Glucose

- Glucose em jejum
 $t = -2.521, g.l. = 9, p = 0.033$
Wilcoxon $p = 0.049$
- Glucose 1 hora após ingestão do pequeno almoço
 $t = -3.040, g.l. = 9, p = 0.014$
Wilcoxon $p = 0.027$
- Glucose 2 horas após ingestão do pequeno almoço
Teste t não aplicável
Wilcoxon $p = 0.557$

Há evidência da ingestão de framboesas ter efeito no teor sérico de glucose em jejum ou uma hora após ingestão de pequeno almoço, mas ao fim de duas horas não há evidência de haver efeito.

Triglicéridos

- Triglicéridos em jejum
 $t = -1.306, g.l. = 9, p = 0.224$
Wilcoxon $p = 0.607$
- Triglicéridos 1 hora após ingestão do pequeno almoço
 $t = -1.059, g.l. = 9, p = 0.317$
Wilcoxon $p = 0.432$
- Triglicéridos 2 horas após ingestão do pequeno almoço
 $t = -1.254, g.l. = 9, p = 0.241$
Wilcoxon $p = 0.557$
- Triglicéridos 4 horas após ingestão do pequeno almoço
 $t = 1.353, g.l. = 9, p = 0.209$
Wilcoxon $p = 0.232$
- Triglicéridos 6 horas após ingestão do pequeno almoço
Teste t não aplicável
Wilcoxon $p = 0.131$

Não há evidência de existência de diferenças, no teor sérico de triglicéridos, entre os meses de Maio e Junho em qualquer das fases da medição.

HDL

- HDL em jejum

$$t = -1.269, g.l. = 9, p = 0.236$$

$$\text{Wilcoxon } p = 0.275$$

- HDL 6 horas após ingestão do pequeno almoço

$$t = -8.176, g.l. = 9, p = 0.000$$

$$\text{Wilcoxon } p = 0.002$$

Conclui-se assim que se pode manter a hipótese nula de não haver efeito da adição de framboesas à dieta no que se refere aos níveis séricos em jejum, mas que há uma diferença altamente significativa no que se refere a níveis séricos após 6 horas de tomar o pequeno almoço.

LDL

- LDL em jejum

$$t = 0.409, g.l. = 9, p = 0.692$$

$$\text{Wilcoxon } p = 1.000$$

- LDL 6 horas após ingestão do pequeno almoço

$$t = -1.657, g.l. = 9, p = 0.132$$

$$\text{Wilcoxon } p = 0.193$$

É de manter a hipótese nula de que não há alteração devida à ingestão de framboesas, nem em jejum nem 6 horas após ingestão do pequeno almoço

LDLox

- LDLox em jejum

$$t = 5.358, g.l. = 9, p = 0.000$$

$$\text{Wilcoxon } p = 0.002$$

- LDLox 4 horas após ingestão do pequeno almoço

$$t = 7.427, g.l. = 9, p = 0.000$$

$$\text{Wilcoxon } p = 0.002$$

- LDLox 6 horas após ingestão do pequeno almoço

$$t = 5.334, g.l. = 9, p = 0.002$$

$$\text{Wilcoxon } p = 0.004$$

Há indicação da ingestão de framboesas ter efeito no teor sérico de LDLox, em qualquer das fases da medição.

IL-1

- IL-1 em jejum
 $t = 1.512$, $g.l. = 9$, $p = 0.165$
Wilcoxon $p = 0.152$
- IL-1 1 hora após ingestão do pequeno almoço
 $t = 2.600$, $g.l. = 9$, $p = 0.029$
Wilcoxon $p = 0.031$
- IL-1 2 horas após ingestão do pequeno almoço
 $t = 0.732$, $g.l. = 9$, $p = 0.483$
Wilcoxon $p = 0.652$
- IL-1 4 horas após ingestão do pequeno almoço
Teste t não aplicável
Wilcoxon $p = 0.980$
- IL-1 6 horas após ingestão do pequeno almoço
 $t = 1.959$, $g.l. = 9$, $p = 0.082$
Wilcoxon $p = 0.84$

Apenas no que se refere ao teor sérico de IL-1 uma hora após absorção do pequeno almoço se observa haver diferença indicativa do efeito da ingestão de framboesas.

IL-6

- IL-6 em jejum
 $t = 1.839$, $g.l. = 9$, $p = 0.099$
Wilcoxon $p = 0.105$
- IL-6 2 horas após ingestão do pequeno almoço
 $t = -3.541$, $g.l. = 9$, $p = 0.006$
Wilcoxon $p = 0.014$
- IL-6 4 horas após ingestão do pequeno almoço
 $t = 2.947$, $g.l. = 9$, $p = 0.016$
Wilcoxon $p = 0.020$
- IL-6 6 horas após ingestão do pequeno almoço
 $t = 5.235$, $g.l. = 9$, $p = 0.001$
Wilcoxon $p = 0.004$

No que se refere à medição em jejum não se observa uma diferença significativa, mas observa-se diferença significativa nos teores séricos depois de decorridas duas, quatro ou seis hora de ingestão do pequeno almoço.

Apêndice: outputs do SPSS23

Observa-se que a variável para a qual houve um efeito mais expressivo é LDLox, o que parece apontar no sentido de haver efeito benéfico das framboesas devido ao seu elevado potencial anti-oxidante:

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	GluMa0 - GluJu0	-7.35000	9.21982	2.91556	-13.94546	-.75454	-2.521	9	.033
Pair 2	GluMa1 - GluJu1	-15.23000	15.84060	5.00924	-26.56168	-3.89832	-3.040	9	.014
Pair 3	GluMa2 - GluJu2	-5.28000	21.92167	6.93224	-20.96182	10.40182	-.762	9	.466

Figura 2.5: Comparação da Glucose em Maio e Junho (1)

Test Statistics ^a			
	GluJu0 - GluMa0	GluJu1 - GluMa1	GluJu2 - GluMa2
Z	-1.988 ^b	-2.191 ^b	-.663 ^b
Asymp. Sig. (2-tailed)	.047	.028	.508
Exact Sig. (2-tailed)	.049	.027	.557
Exact Sig. (1-tailed)	.024	.014	.278
Point Probability	.006	.004	.032

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figura 2.6: Comparação da Glucose em Maio e Junho (2)

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	TriMa0 - TriJu0	-9.78000	23.67928	7.48805	-26.71914	7.15914	-1.306	9	.224
Pair 2	TriMa1 - TriJu1	-10.71000	31.97574	10.11162	-33.58406	12.16406	-1.059	9	.317
Pair 3	TriMa2 - TriJu2	-19.56000	49.32978	15.59945	-54.84840	15.72840	-1.254	9	.241
Pair 4	TriMa4 - TriJu4	30.80000	71.98447	22.76349	-20.69459	82.29459	1.353	9	.209
Pair 5	TriMa6 - TriJu6	-24.23000	44.24847	13.99259	-55.88345	7.42345	-1.732	9	.117

Figura 2.7: Comparação dos Triglicéridos em Maio e Junho (1)

Test Statistics ^a					
	TriJu0 - TriMa0	TriJu1 - TriMa1	TriJu2 - TriMa2	TriJu4 - TriMa4	TriJu6 - TriMa6
Z	-.561 ^b	-.866 ^b	-.663 ^b	-1.274 ^c	-1.580 ^b
Asymp. Sig. (2-tailed)	.575	.386	.508	.203	.114
Exact Sig. (2-tailed)	.607	.432	.557	.232	.131
Exact Sig. (1-tailed)	.304	.216	.278	.116	.065
Point Probability	.018	.028	.032	.020	.013

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

c. Based on positive ranks.

Figura 2.8: Comparação dos Triglicéridos em Maio e Junho (2)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	chol2Ma0 - chol2Ju0	2.38000	20.42731	6.45968	-12.23282	16.99282	.368	9	.721
Pair 2	chol2Ma6 - chol2Ju6	-16.98000	22.91583	7.24662	-33.37299	-.58701	-2.343	9	.044

Figura 2.9: Comparação do Colesterol em Maio e Junho (1)

Test Statistics ^a		
	chol2Ju0 - chol2Ma0	chol2Ju6 - chol2Ma6
Z	-.255 ^b	-2.090 ^c
Asymp. Sig. (2-tailed)	.799	.037
Exact Sig. (2-tailed)	.846	.037
Exact Sig. (1-tailed)	.423	.019
Point Probability	.038	.005

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Figura 2.10: Comparação do Colesterol em Maio e Junho (2)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	LDLMa0 - LDLJu0	1.38000	10.65925	3.37075	-6.24517	9.00517	.409	9	.692
Pair 2	LDLMa6 - LDLJu6	-6.94000	13.24464	4.18832	-16.41464	2.53464	-1.657	9	.132

Figura 2.11: Comparação das LDL em Maio e Junho (1)

Test Statistics ^a		
	LDLJu0 - LDLMa0	LDLJu6 - LDLMa6
Z	-.059 ^b	-1.376 ^c
Asymp. Sig. (2-tailed)	.953	.169
Exact Sig. (2-tailed)	1.000	.193
Exact Sig. (1-tailed)	.500	.097
Point Probability	.045	.017

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Figura 2.12: Comparação das LDL em Maio e Junho (2)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	HDLMa0 - HDLJu0	-2.60000	6.47868	2.04874	-7.23457	2.03457	-1.269	9	.236
Pair 2	HDLMa6 - HDLJu6	-12.73000	4.92365	1.55699	-16.25217	-9.20783	-8.176	9	.000

Figura 2.13: Comparação das HDL em Maio e Junho (1)

Test Statistics^a

	HDLJu0 - HDLMa0	HDLJu6 - HDLMa6
Z	-1.172 ^b	-2.805 ^b
Asymp. Sig. (2-tailed)	.241	.005
Exact Sig. (2-tailed)	.275	.002
Exact Sig. (1-tailed)	.138	.001
Point Probability	.021	.001

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figura 2.14: Comparação das HDL em Maio e Junho (2)

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	LDLoxMa0 - LDLoxJu0	9.99000	5.89566	1.86437	5.77250	14.20750	5.358	9	.000
Pair 2	LDLoxMa4 - LDLoxJu4	22.18000	9.44373	2.98637	15.42436	28.93564	7.427	9	.000
Pair 3	LDLoxMa6 - LDLoxJu6	12.39000	9.04034	2.85881	5.92293	18.85707	4.334	9	.002

Figura 2.15: Comparação das LDLox em Maio e Junho (1)

Test Statistics^a

	LDLoxJu0 - LDLoxMa0	LDLoxJu4 - LDLoxMa4	LDLoxJu6 - LDLoxMa6
Z	-2.803 ^b	-2.803 ^b	-2.703 ^b
Asymp. Sig. (2-tailed)	.005	.005	.007
Exact Sig. (2-tailed)	.002	.002	.004
Exact Sig. (1-tailed)	.001	.001	.002
Point Probability	.001	.001	.001

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figura 2.16: Comparação das LDLox em Maio e Junho (2)

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	IL1Ma0 - IL1Ju0	.59000	1.23419	.39028	-.29289	1.47289	1.512	9	.165
Pair 2	IL1Ma1 - IL1Ju1	.62000	.75395	.23842	.08065	1.15935	2.600	9	.029
Pair 3	IL1Ma2 - IL1Ju2	.25000	1.08038	.34165	-.52286	1.02286	.732	9	.483
Pair 4	IL1Ma4 - IL1Ju4	.30000	1.48773	.47046	-.76426	1.36426	.638	9	.540
Pair 5	IL1Ma6 - IL1Ju6	.65100	1.05109	.33238	-.10090	1.40290	1.959	9	.082

Figura 2.17: Comparação da IL-1 em Maio e Junho (1)

Test Statistics^a

	IL1Ju0 – IL1Ma0	IL1Ju1 – IL1Ma1	IL1Ju2 – IL1Ma2	IL1Ju4 – IL1Ma4	IL1Ju6 – IL1Ma6
Z	-1.479 ^b	-2.173 ^b	-.511 ^b	-.051 ^c	-1.784 ^b
Asymp. Sig. (2-tailed)	.139	.030	.610	.959	.074
Exact Sig. (2-tailed)	.152	.031	.652	.980	.084
Exact Sig. (1-tailed)	.076	.016	.326	.490	.042
Point Probability	.008	.008	.027	.021	.010

- a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Figura 2.18: Comparação da IL-1 em Maio e Junho (2)

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	IL6Ma0 – IL6Ju0	.23600	.40591	.12836	-.05437	.52637	1.839	9	.099
Pair 2	IL6Ma2 – IL6Ju2	.64000	.57147	.18071	.23120	1.04880	3.541	9	.006
Pair 3	IL6Ma4 – IL6Ju4	.80900	.86807	.27451	.18802	1.42998	2.947	9	.016
Pair 4	IL6Ma6 – IL6Ju6	1.87200	1.13075	.35757	1.06311	2.68089	5.235	9	.001

Figura 2.19: Comparação da IL-6 em Maio e Junho (1)

Test Statistics^a

	IL6Ju0 – IL6Ma0	IL6Ju2 – IL6Ma2	IL6Ju4 – IL6Ma4	IL6Ju6 – IL6Ju0
Z	-1.682 ^b	-2.395 ^b	-2.293 ^b	-2.701 ^c
Asymp. Sig. (2-tailed)	.093	.017	.022	.007
Exact Sig. (2-tailed)	.105	.014	.020	.004
Exact Sig. (1-tailed)	.053	.007	.010	.002
Point Probability	.011	.002	.003	.001

- a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Figura 2.20: Comparação da IL-6 em Maio e Junho (2)

Capítulo 3

Perfis

Neste capítulo vamos fazer uma análise preliminar com vista a esclarecer a questão

O tratamento com framboesas teve um efeito detetável sobre as variáveis em análise?

O primeiro passo vai ser a representação gráfica do conjunto dos perfis individuais, para observar se indicam globalmente alterações na evolução temporal das variáveis — e mais uma vez anotamos que não temos informação que nos permita distinguir variáveis resposta de covariáveis, pelo que trataremos qualquer uma das 8 variáveis disponíveis como variáveis resposta.

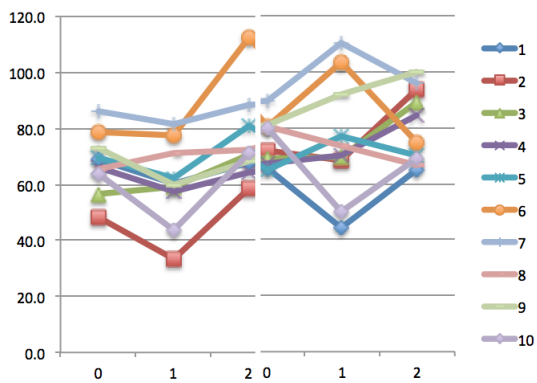
Os comentários sugeridos quer pelas Tabelas 1.1 a 1.9 quer pelos gráficos (que devem ser comparados com os gráficos 5.9 a 5.16) são os seguintes:

Glucose Enquanto em Maio parece haver um padrão geral de diminuição da glucose uma hora após ingestão do pequeno almoço seguida de retoma de valores próximos dos registados em jejum (exceto o voluntário 6, em que há acentuado aumento da glucose ao fim de duas horas), em Junho não se pode identificar um padrão.

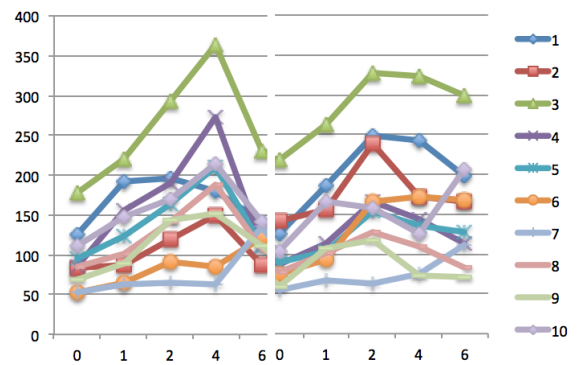
Triglicéridos A comparação dos dois períodos parece sugerir que os picos observados ao fim de quatro horas em Maio deixaram de existir em Junho. Observa-se também que o voluntário 3 tem valores fora do normal quer em Maio quer em Junho.

Colesterol Parece haver uma maior estabilização dos valores em Junho; a variabilidade individual não parece indicar uma alteração global de perfis no grupo.

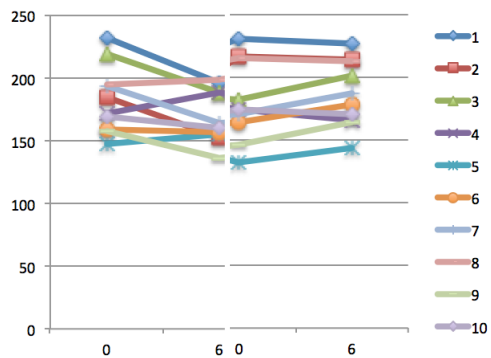
HDL Enquanto em Maio se observava uma consistente diminuição do HDL entre a fase de jejum e o final do período de 6 horas de observação, em Junho qualquer dos voluntários parece ter valores muito estáveis em todo o período. Por outro lado, os valores de HDL são genericamente mais elevados, o que é um indicador de que o tratamento produz benefícios. Também no que se refere a esta variável o voluntário 3 parece um *outlier*.



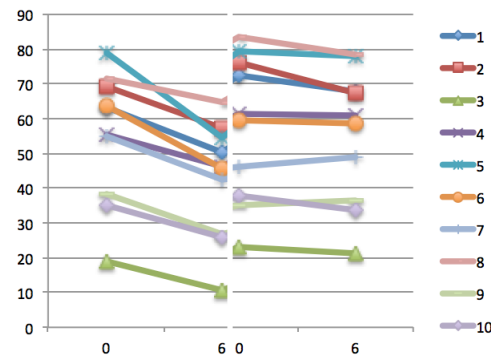
Perfis (Glucose)



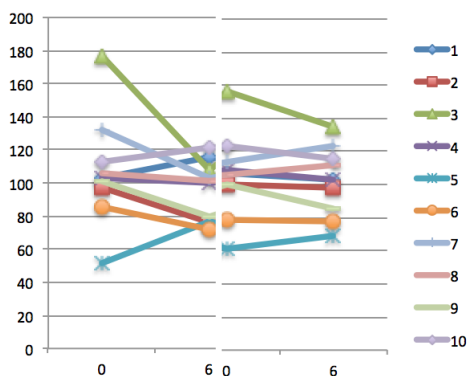
Perfis (Triglicéridos)



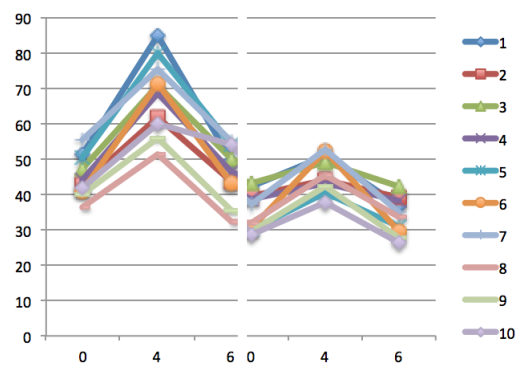
Perfis (Colesterol)



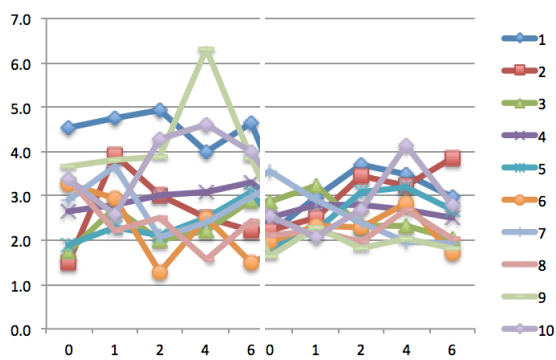
Perfis (HDL)



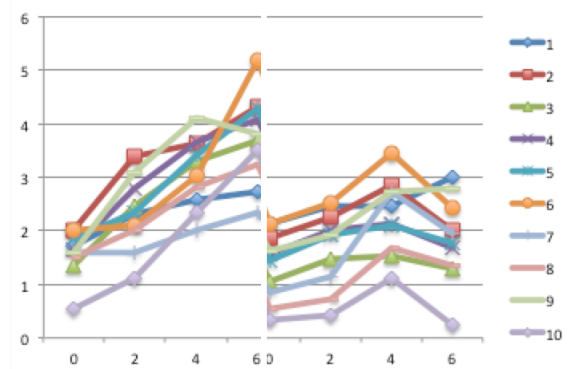
Perfis LDL)



Perfis (LDLox)



Perfis (IL – 1)



Perfis (IL – 6)

LDL Em Maio não há um padrão consistente; em Junho parece haver genericamente valores estáveis nos dois tempos de medição, e mais uma vez o voluntário 3 apresenta valores fora de controle.

LDLox No que se refere ao LDL oxidado, os valores parecem todos mais controlados em Junho, indicando que o tratamento com framboesas pode efetivamente melhorar o prognóstico no que se refere a inflamação das paredes arteriais, aterosclerose e eventuais crises vasculares.

IL-1 Observa-se uma regularização dos valores em Junho, após o período de tratamento, quer no que se refere a diminuição dos valores, quer à sua estabilização aos longo das seis horas de observação.

IL-6 Enquanto em Maio se observava um aumento gradual da IL-6 ao longo do período de observação, em Junho os valores, que são notoriamente mais baixos, têm variações muito menos acentuadas.

Como anteriormente vimos, o grupo de 10 voluntários é muito heterogéneo, e seria mesmo decerto de excluir o voluntário 3, muito atípico.

Os gráficos apresentados permitiram algumas observações gerais, mas não suportam a frase “um gráfico vale mil palavras”, por o excesso de informação os tornar confusos.

Como a observação das faces de Chernoff dos voluntários em Maio parece suportar a ideia de que há um grupo de voluntários (4, 5, 6, 9, 10) com perfis semelhante, adiante apresentamos os gráficos dos perfis desse subgrupo no que se refere às oito variáveis registadas.

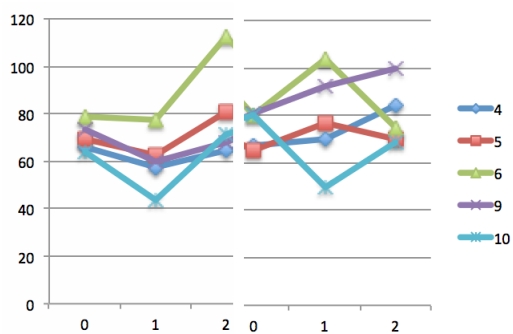
Para este subgrupo os comentários são praticamente a repetição dos anteriormente feitos, mas traduzem mais fielmente o que se observa:

Glucose Enquanto em Maio parece haver um padrão geral de diminuição da glucose uma hora após ingestão do pequeno almoço seguida de retoma de valores próximos dos em jejum (exceto o voluntário 6, em que há acentuado aumento da glicose ao fim de duas horas), em Junho não se pode identificar um padrão.

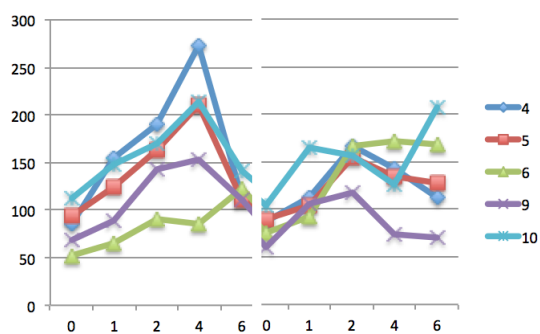
Triglicéridos A comparação dos dois períodos parece sugerir que os picos observados ao fim de quatro horas em Maio deixaram de existir.

Colesterol Parece haver uma maior estabilização dos valores em Junho; a variabilidade individual não parece indicar uma alteração global de perfis no grupo.

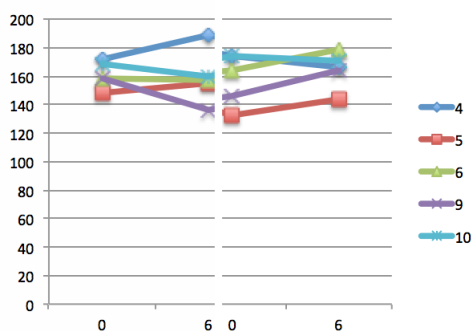
HDL Enquanto em Maio se observava uma consistente diminuição do HDL entre a fase de jejum e o final do período de 6 horas de observação, em Junho qualquer dos voluntários parece ter valores muito estáveis em todo o período. Por outro lado, os valores de HDL são genericamente mais elevados, indicando que o tratamento produz benefícios.



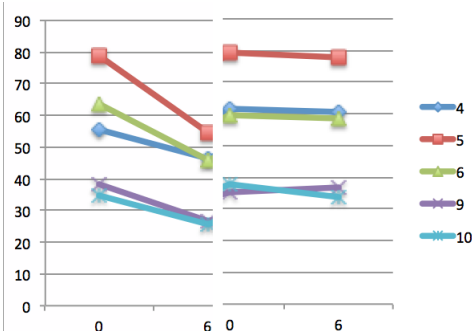
Perfis (Glucose)



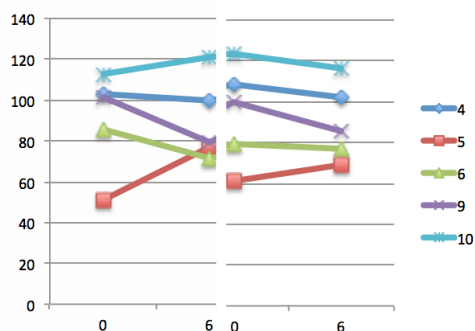
Perfis (Triglicéridos)



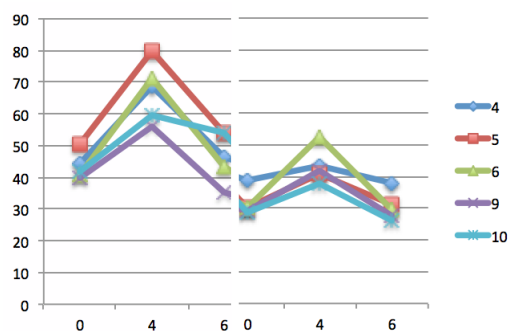
Perfis (Colesterol)



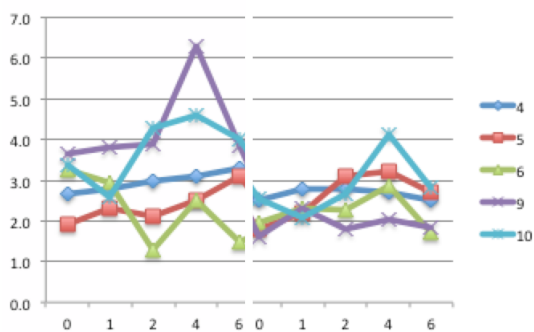
Perfis (HDL)



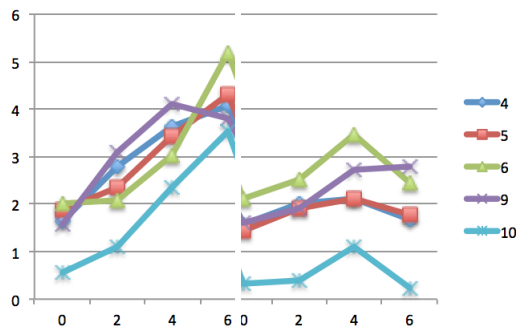
Perfis LDL)



Perfis (LDLox)



Perfis (IL – 1)



Perfis (IL – 6)

LDL Em Maio não há um padrão consistente; em Junho parece haver genericamente valores estáveis nos dois tempos de medição.

LDLox No que se refere ao LDL oxidado, os valores parecem todos mais controlados em Junho, indicando que o tratamento com framboesas pode efetivamente melhorar o prognóstico no que se refere a inflamação das paredes arteriais, aterosclerose e eventuais crises vasculares.

- IL-1** Observa-se uma regularização dos valores em Junho, após o período de tratamento, quer no que se refere a diminuição dos valores, quer à sua estabilização aos longo das seis horas de observação.
- IL-6** Enquanto em Maio se observava um aumento gradual da IL-6 ao longo do período de observação, em Junho os valores, que são notoriamente mais baixos, têm variações muito menos acentuadas.

Capítulo 4

Evolução de cada voluntário

4.1 Dados individuais

No Capítulo 2 investigámos a eventual alteração dos perfis individuais em jejum, uma das análises que nos tinha sido explicitamente solicitada. Agora, a fim de inspecionar o que se passa com cada voluntário, rerepresentamos nas Tabelas 4.1 a 4.10 os dados organizados por forma a mostrar a evolução de cada uma das variáveis em cada um deles. Como não há registo da medição destas em todos os instantes considerados, apenas no que se refere a 6h após ingestão do pequeno almoço há dados suficientes para fazer um estudo de perfis semelhante ao feito para os registos individuais em jejum (e para tal fazemos a imputação do valor da glicose, como adiante explicado).

Tabela 4.1: Voluntário 1

	Maio						Junho				
	0 h	1 h	2 h	4 h	6 h		0 h	1 h	2 h	4 h	6 h
Glicose	69.0	60.5	67.6				65.7	44.1	64.8		
Trigl.	125.7	193.0	196.9	180.7	141.7		126.3	186.1	249.4	241.9	198.2
Colesterol	232.0				196.1		230.2				227.0
LDL	102.7				115.3		106.7				103.4
HDL	63.4				50.3		72.3				67.8
LDLox	51.6			85.1	49.7		42.2			50.9	37.0
IL-1	4.6	4.8	4.9	4.0	4.7		2.1	3.0	3.7	3.5	3.0
IL-6	1.71		2.32	2.57	2.73		2.13		2.45	2.46	2.99

Tabela 4.2: Voluntário 2

	Maio						Junho				
	0 h	1 h	2 h	4 h	6 h		0 h	1 h	2 h	4 h	6 h
Glicose	48.3	33.4	58.8				71.9	68.4	93.8		
Trigl.	83.7	86.4	119.6	148.8	87.4		142.5	156.0	239.0	172.4	165.8
Colesterol	184.5				152.8		216.4				213.9
LDL	97.1				76.2		99.8				98.2
HDL	69.1				57.3		76.3				67.3
LDLox	42.8			61.7	43.7		39.8			44.0	39.2
IL-1	1.5	3.9	3.0	2.5	2.2		2.2	2.5	3.5	3.2	3.9
IL-6	2.02		3.40	3.62	4.31		1.86		2.26	2.84	2.00

Tabela 4.3: Voluntário 3

	Maio						Junho				
	0 h	1 h	2 h	4 h	6 h		0 h	1 h	2 h	4 h	6 h
Glicose	56.6	59.1	70.8				69.0	69.5	89.2		
Trigl.	177.4	220.6	293.3	363.6	231.0		219.3	263.5	328.3	322.9	299.8
Colesterol	219.0				188.6		182.2				201.0
LDL	176.3				107.4		156.7				134.9
HDL	18.8				10.5		23.0				21.5
LDLox	47.5			71.1	50.0		43.2			49.1	42.3
IL-1	1.7	2.7	2.0	2.2	2.9		2.9	3.2	2.3	2.3	2.0
IL-6	1.35		2.46	3.31	3.70		1.04		1.47	1.53	1.29

Tabela 4.4: Voluntário 4

	Maio						Junho				
	0 h	1 h	2 h	4 h	6 h		0 h	1 h	2 h	4 h	6 h
Glicose	66.2	57.5	64.3				67.3	70.3	84.4		
Trigl.	85.8	155.6	189.7	272.8	112.0		88.0	112.5	166.2	143.4	113.0
Colesterol	171.7				188.3		174.2				165.8
LDL	103.2				99.7		108.4				102.5
HDL	55.6				46.3		61.4				60.8
LDLox	44.2			68.6	46.3		38.6			43.3	38.0
IL-1	2.7	2.8	3.0	3.1	3.3		2.5	2.8	2.8	2.7	2.5
IL-6	1.66		2.80	3.65	4.08		1.56		2.02	2.12	1.67

Tabela 4.5: Voluntário 5

	Maio						Junho				
	0 h	1 h	2 h	4 h	6 h		0 h	1 h	2 h	4 h	6 h
Glicose	69.5	62.3	81.0				65.1	77.1	69.9		
Trigl.	94.6	124.0	163.3	210.5	111.6		89.7	103.6	154.0	134.8	127.3
Colesterol	147.9				154.7		132.0				143.8
LDL	51.2				77.1		61.2				69.2
HDL	78.8				54.6		79.6				78.0
LDLox	50.4			79.8	54.2		30.0			40.7	31.0
IL-1	1.9	2.3	2.1	2.5	3.1		1.8	2.2	3.1	3.2	2.7
IL-6	1.87		2.34	3.43	4.30		1.43		1.93	2.10	1.77

A primeira constatação tem que ver com a grande quantidade de dados omissos. De facto, apenas em jejum foi registada a série completa de medições das oito variáveis.

Admitimos que a glicose foi registada apenas nos instantes 0 h, 1 h e 2 h porque a partir daí o valor estabiliza (aparentemente nenhum dos voluntários é diabético).

Com este pressuposto vamos fazer a imputação no instante 6 h igual ao observado ao fim de duas horas, obtendo então registos “completos” das 8 variáveis 6 horas após ingestão do pequeno almoço, veja-se a Tabela 4.11.

Tabela 4.6: Voluntário 6

	Maio					Junho				
	0 h	1 h	2 h	4 h	6 h	0 h	1 h	2 h	4 h	6 h
Glicose	78.8	77.7	112.5			80.4	103.6	74.9		
Trigl.	52.4	65.7	90.5	85.7	122.2	75.5	92.7	166.7	172.2	168.8
Colesterol	158.8				156.9	164.5				178.9
LDL	85.4				71.6	79.2				77.3
HDL	63.7				45.7	59.4				58.5
LDLox	41.1			71.2	43.4	30.3			52.4	29.7
IL-1	3.3	2.9	1.3	2.5	1.5	2.0	2.3	2.3	2.9	1.7
IL-6	2.00		2.09	3.02	5.20	2.12		2.52	3.46	2.44

Tabela 4.7: Voluntário 7

	Maio					Junho				
	0 h	1 h	2 h	4 h	6 h	0 h	1 h	2 h	4 h	6 h
Glicose	85.9	81.5	88.5			89.4	110.5	95.9		
Trigl.	52.9	62.8	64.4	63.8	135.9	54.8	66.7	62.9	75.4	111.6
Colesterol	193.8				163.9	171.0				187.7
LDL	132.4				102.7	113.6				122.9
HDL	54.8				42.4	46.2				49.0
LDLox	55.3			75.3	55.0	37.4			52.6	35.0
IL-1	2.9	3.7	2.1	2.4	3.0	3.5	2.9	2.4	2.0	1.9
IL-6	1.59		1.60	2.00	2.34	0.85		1.13	2.74	1.96

Tabela 4.8: Voluntário 8

	Maio					Junho				
	0 h	1 h	2 h	4 h	6 h	0 h	1 h	2 h	4 h	6 h
Glicose	65.5	71.0	72.4			80.5	73.4	66.8		
Trigl.	85.4	101.5	141.0	187.7	107.4	77.3	101.0	126.7	108.9	82.8
Colesterol	194.3				199.2	215.1				213.4
LDL	106.3				101.3	106.3				111.3
HDL	71.4				64.8	83.7				78.3
LDLox	36.6			51.3	32.1	32.0			45.3	33.8
IL-1	3.4	2.2	2.5	1.6	2.4	2.1	2.2	2.0	2.7	2.1
IL-6	1.50		2.02	2.83	3.24	0.54		0.73	1.69	1.34

Mas como justificado no Capítulo 2, há que passar aos dados estandardizados para se tratar de representações gráficas, pois as diversas variáveis têm localizações centrais e escalas muito diversas. Os dados padronizados estão registados na Tabela 4.12.

Com base nesta Tabela 4.12 obtemos então os perfis de cada um dos 10 voluntários 6 horas após ingestão do pequeno almoço, veja-se a correspondente matriz dos gráficos.

Tabela 4.9: Voluntário 9

	Maio					Junho				
	0 h	1 h	2 h	4 h	6 h	0 h	1 h	2 h	4 h	6 h
Glicose	73.2	59.9	68.1			81.0	92.1	99.7		
Trigl.	69.3	89.5	143.7	152.4	111.0	59.7	106.6	117.9	73.4	70.6
Colesterol	158.4				136.3	145.6				164.0
LDL	101.5				79.4	99.8				85.7
HDL	38.5				26.7	35.3				36.7
LDLox	40.0			56.0	35.4	29.4			42.1	27.7
IL-1	3.7	3.8	3.9	6.3	3.9	1.6	2.3	1.8	2.1	1.8
IL-6	1.58		3.10	4.10	3.80	1.62		1.91	2.74	2.80

Tabela 4.10: Voluntário 10

	Maio					Junho				
	0 h	1 h	2 h	4 h	6 h	0 h	1 h	2 h	4 h	6 h
Glicose	63.8	43.7	71.3			80.0	49.9	68.7		
Trigl.	112.1	147.8	169.9	213.5	142.1	104.0	165.3	156.8	126.2	206.7
Colesterol	168.8				159.7	174.2				170.8
LDL	112.8				121.3	123.4				116.0
HDL	35.0				25.7	37.9				33.7
LDLox	42.0			59.8	54.0	28.7			37.7	26.2
IL-1	3.4	2.6	4.3	4.6	4.0	2.5	2.1	2.7	4.1	2.8
IL-6	0.55		1.10	2.34	3.52	0.32		0.41	1.10	0.24



Não há padrões discerníveis, apenas se pode constatar muito superficialmente que os voluntários 3 e 8 têm perfis idênticos em Maio e em Junho (e recorde-se que o voluntário 3 tem valores muito atípicos).

Tabela 4.11: Registos às 6 h (após imputação do valor da glicose).

		glicose	Trig.	colesterol	LDL	HDL	LDLox	IL-1	IL-6
vol. 1	Maio	67.6	141.7	196.1	115.3	50.3	49.7	4.7	2.73
	Junho	64.8	198.2	227.0	103.4	67.8	37.0	3.0	2.99
vol. 2	Maio	58.8	87.4	152.8	76.2	57.3	43.7	2.2	4.31
	Junho	93.8	165.8	213.9	98.2	67.3	39.2	3.9	2.00
vol. 3	Maio	70.8	231.0	188.6	107.4	10.5	50.0	2.9	3.70
	Junho	89.2	299.8	201.0	134.9	21.5	42.3	2.0	1.29
vol. 4	Maio	64.3	112.0	188.3	99.7	46.3	46.3	3.3	4.08
	Junho	84.4	113.0	165.8	102.5	60.8	38.0	2.5	1.67
vol. 5	Maio	81.0	111.6	154.7	77.1	54.6	54.2	3.1	4.30
	Junho	69.9	127.3	143.8	69.2	78.0	31.0	2.7	1.77
vol. 6	Maio	112.5	122.2	156.9	71.6	45.7	43.4	1.5	5.20
	Junho	74.9	168.8	178.9	77.3	58.5	29.7	1.7	2.44
vol. 7	Maio	88.5	135.9	163.9	102.7	42.4	55.0	3.0	2.34
	Junho	95.9	111.6	187.7	122.9	49.0	35.0	1.9	1.96
vol. 8	Maio	72.4	107.4	199.2	101.3	64.8	32.1	2.4	3.24
	Junho	66.8	82.8	213.4	111.3	78.3	33.8	2.1	1.34
vol. 9	Maio	68.1	111.0	136.3	79.4	26.7	35.4	3.9	3.80
	Junho	99.7	70.6	164.0	85.7	36.7	27.7	1.8	2.80
vol. 10	Maio	71.3	142.1	159.7	121.3	25.7	54.0	4.0	3.52
	Junho	68.7	206.7	170.8	116.0	33.7	26.2	2.8	0.24

Como os gráficos de perfis são pouco reveladores, vamos usar a Tabela 4.12 para fazer as caras de Chernoff de cada voluntário em Maio e em Junho, tal como nas Figuras 2.5 a 2.14 do Capítulo 2, veja-se adiante as Figuras 4.1 a 4.10.

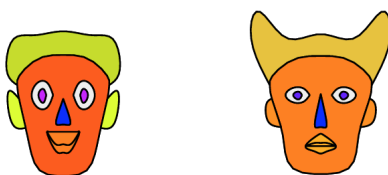


Figura 4.1: Voluntário 1 após 6h, Maio e Junho

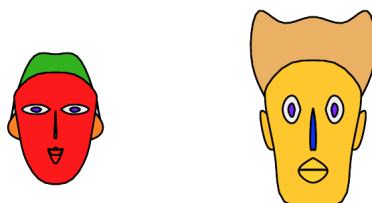


Figura 4.2: Voluntário 2 após 6h, Maio e Junho

As caras de Chernoff mostram que afinal apenas o voluntário 8 apresenta semelhança nos dois períodos de medições 6h após ingestão do pequeno almoço.

Tabela 4.12: Valores padronizados às 6 h (após imputação do valor da glicose).

		glicose	Trig.	colesterol	LDL	HDL	LDLox	IL-1	IL-6
vol 1.	Maio	-0.74	-0.08	0.71	0.90	0.08	1.04	2.20	-0.05
	Junho	-0.94	0.95	1.94	0.28	1.01	-0.35	0.26	0.16
vol. 2	Maio	-1.36	-1.07	-1.01	-1.13	0.45	0.39	-0.66	1.22
	Junho	1.10	0.36	1.42	0.01	0.99	-0.11	1.26	-0.63
vol. 3	Maio	-0.52	1.55	0.42	0.49	-2.04	1.08	0.15	0.73
	Junho	0.77	2.81	0.91	1.91	-1.45	0.23	-0.88	-1.20
vol. 4	Maio	-0.97	-0.62	0.40	0.09	-0.13	0.67	0.62	1.04
	Junho	0.44	-0.60	-0.49	0.23	0.64	-0.24	-0.32	-0.90
vol. 5	Maio	0.20	-0.63	-0.93	-1.08	0.31	1.54	0.39	1.21
	Junho	-0.58	-0.34	-1.37	-1.49	1.56	-1.01	-0.08	-0.82
vol. 6	Maio	2.41	-0.44	-0.84	-1.37	-0.17	0.35	-1.51	1.93
	Junho	-0.23	0.41	0.03	-1.07	0.52	-1.15	-1.23	-0.28
vol. 7	Maio	0.72	-0.19	-0.57	0.24	-0.34	1.63	0.23	-0.36
	Junho	1.24	-0.63	0.38	1.29	0.01	-0.57	-1.00	-0.67
vol. 8	Maio	-0.40	-0.71	0.84	0.17	0.85	-0.89	-0.39	0.36
	Junho	-0.80	-1.16	1.40	0.69	1.57	-0.70	-0.84	-1.16
vol. 9	Maio	-0.71	-0.64	-1.66	-0.97	-1.18	-0.53	1.28	0.81
	Junho	1.51	-1.38	-0.56	-0.64	-0.64	-1.37	-1.10	0.01
vol. 10	Maio	-0.48	-0.07	-0.73	1.21	-1.23	1.52	1.44	0.59
	Junho	-0.66	1.11	-0.29	0.93	-0.80	-1.54	0.05	-2.05

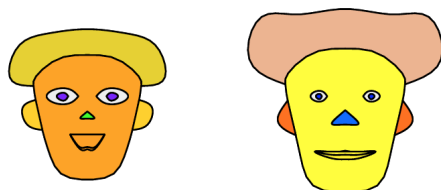


Figura 4.3: Voluntário 3 após 6h, Maio e Junho

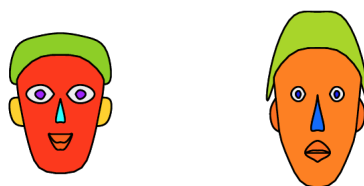


Figura 4.4: Voluntário 4 após 6h, Maio e Junho



Figura 4.5: Voluntário 5 após 6h, Maio e Junho

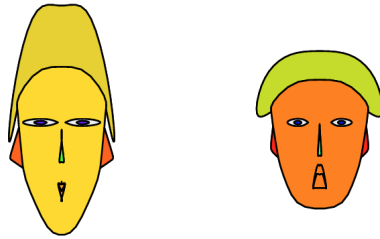


Figura 4.6: Voluntário 6 após 6h, Maio e Junho

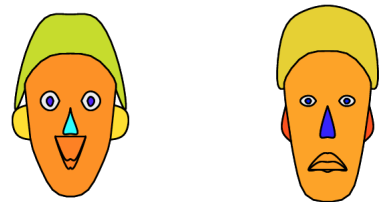


Figura 4.7: Voluntário 7 após 6h, Maio e Junho



Figura 4.8: Voluntário 8 após 6h, Maio e Junho

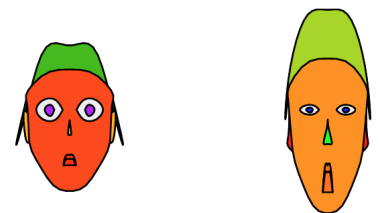


Figura 4.9: Voluntário 9 após 6h, Maio e Junho



Figura 4.10: Voluntário 10 após 6h, Maio e Junho

4.2 Correlação

Na análise de medições repetidas é fundamental estudar a dependência entre os dados, veja-se o capítulo sobre correlação serial de Ramsey and Schaffer (2002).

Mais uma vez, a qualidade dos dados não permite um estudo avançado como desejaríamos. De facto, no que se refere a colesterol, LDL e HDL apenas há duas observações em cada mês — e nessa situação fazemos meramente um estudo de correlações.

Por outro lado, a metodologia exposta em Ramsey and Schaffer tem como pressuposto que as

medições são identicamente espaçadas, o que acontece apenas no caso da glicose e da IL-1. Poderia eventualmente estudar-se nesses casos uma correlação média (semelhante à correlação ordinal média de Friedman), mas como não encontrei testes formais para investigar a aleatoriedade como o teste de Friedman faz, pareceu-me pouco interessante apresentar esses cálculos triviais.

Colesterol, LDL e HDL foram medidos apenas às 0h (jejum) e às 6h quer em Maio quer em Junho. É decerto interessante investigar a associação entre as medições nos dois períodos.

Assim, além dos registos, construímos tabelas com as variações, quer os incrementos entre as 0h e as 6h em Maio e em Junho, quer os incrementos em períodos homólogos de Maio e de Junho — isto é, Valor em Junho – Valor em Maio em jejum e Valor em Junho – Valor em Maio 6h após pequeno-almoço.

Colesterol					Variações				
Dados									
Maio			Junho		6h – 0h			Junho – Maio	
0 h	6 h		0 h	6 h	em Maio	em Junho		às 0h	às 6h
232.0	196.1		230.2	227.0	-35.9	-3.2		-1.8	30.9
184.5	152.8		216.4	213.9	-31.7	-2.5		31.9	61.1
219.0	188.6		182.2	201.0	-30.4	18.8		-36.8	12.4
171.7	188.3		174.2	165.8	16.6	-8.4		2.5	-22.5
147.9	154.7		132.0	143.8	6.8	11.8		-15.9	-10.9
158.8	156.9		164.5	178.9	-1.9	14.4		5.7	22.0
193.8	163.9		171.0	187.7	-29.9	16.7		-22.8	23.8
194.3	199.2		215.1	213.4	4.9	-1.7		20.8	14.2
158.4	136.3		145.6	164.0	-22.1	18.4		-12.8	27.7
168.8	159.7		174.2	170.8	-9.1	-3.4		5.4	11.1

LDL					Variações				
Dados									
Maio			Junho		6h – 0h			Junho – Maio	
0 h	6 h		0 h	6 h	em Maio	em Junho		às 0h	às 6h
102.7	115.3		106.7	103.4	12.6	-3.3		4.0	-11.9
97.1	76.2		99.8	98.2	-20.9	-1.6		2.7	22.0
176.3	107.4		156.7	134.9	-68.9	-21.8		-19.6	27.5
103.2	99.7		108.4	102.5	-3.5	-5.9		5.2	2.8
51.2	77.1		61.2	69.2	25.9	8.0		10.0	-7.9
85.4	71.6		79.2	77.3	-13.8	-1.9		-6.2	5.7
132.4	102.7		113.6	122.9	-29.7	9.3		-18.8	20.2
106.3	101.3		106.3	111.3	-5.0	5.0		0.0	10.0
101.5	79.4		99.8	85.7	-22.1	-14.1		-1.7	6.3
112.8	121.3		123.4	116.0	8.5	-7.4		10.6	-5.3

HDL					Variações				
Dados									
Maio				Junho		6h – 0h		Junho – Maio	
0 h	6 h			0 h	6 h	em Maio	em Junho	às 0h	às 6h
63.4	50.3			72.3	67.8	-13.1	-4.5	8.9	17.5
69.1	57.3			76.3	67.3	-11.8	-9.0	7.2	10.0
18.8	10.5			23.0	21.5	-8.3	-1.5	4.2	11.0
55.6	46.3			61.4	60.8	-9.3	-0.6	5.8	14.5
78.8	54.6			79.6	78.0	-24.2	-1.6	0.8	23.4
63.7	45.7			59.4	58.5	-18.0	-0.9	-4.3	12.8
54.8	42.4			46.2	49.0	-12.4	2.8	-8.6	6.6
71.4	64.8			83.7	78.3	-6.6	-5.4	12.3	13.5
38.5	26.7			35.3	36.7	-11.8	1.4	-3.2	10.0
35.0	25.7			37.9	33.7	-9.3	-4.2	2.9	8.0

Observam-se assim as correlações registadas na Tabela 4.13.

Tabela 4.13: Correlações de Pearson

	colesterol	LDL	HDL
(0h,6h), Maio	0.7188	0.5614	0.9640
(0h,6h), Junho	0.9450	0.9416	0.9884
incrementos Maio–Junho, (0h,6h)	-0.2912	0.5641	-0.1959
incrementos 0h–6h, (Maio,Junho)	0.3660	-0.7696	0.3178

Que se pode concluir? — e não esqueçamos que o que há que ter em conta é o coeficiente de determinação, interpretável como a fração da variabilidade de uma das variáveis em função da variabilidade da outra, assumindo relacionamento aproximadamente linear.

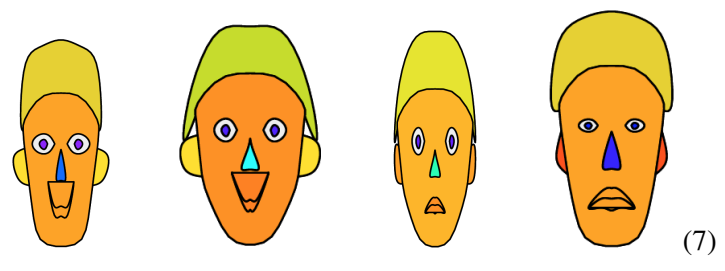
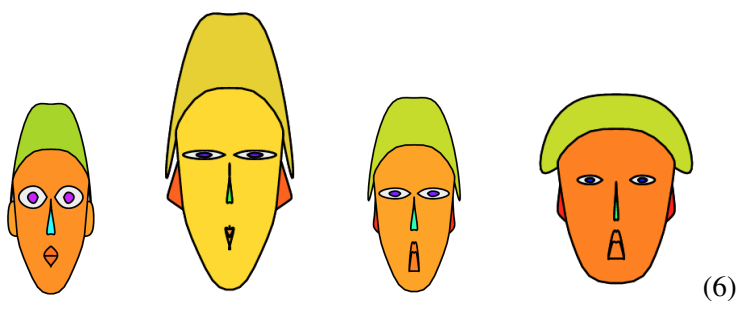
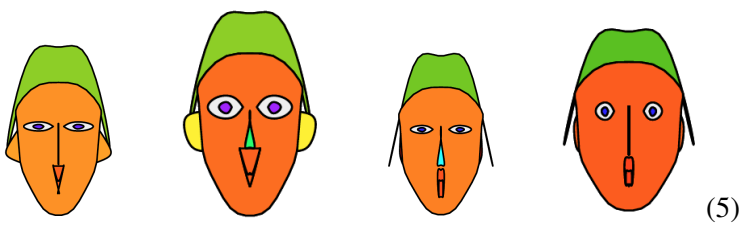
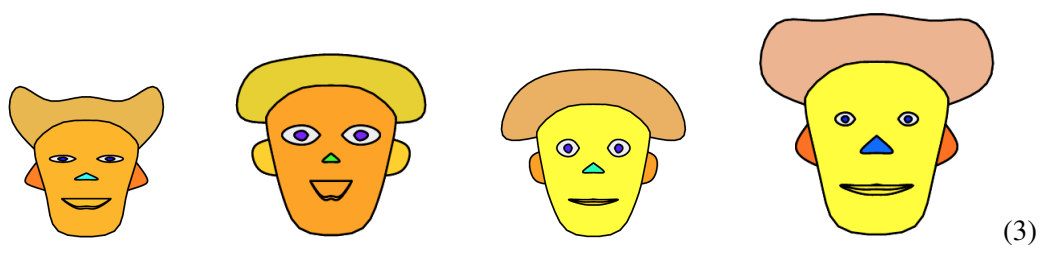
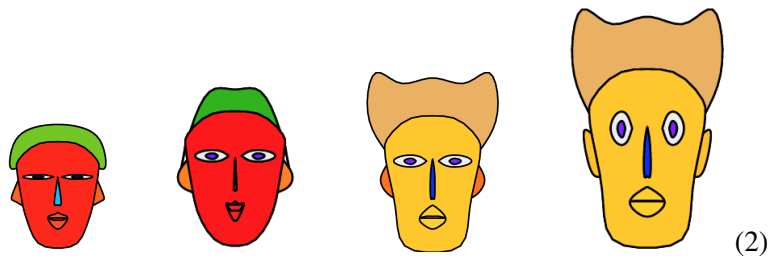
Aparentemente o tratamento reforçou a associação entre os valores em jejum e final de colesterol, LDL e HDL, pois em Junho as correlações entre as medições em jejum e 6 horas após pequeno almoço são todas superiores a 0.94 (e sempre mais próximas de 1 do que as calculadas para o mesmo par de observações em Maio). No que se refere a colesterol e a LDL houve assim uma estabilização notável (o HDL já estava estabilizado em Maio).

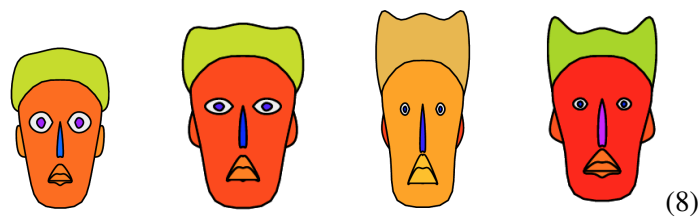
4.3 Retorno às Caras de Chernoff

Justapomos as caras de Chernoff que cada um dos voluntários apresenta em (da esquerda para a direita)

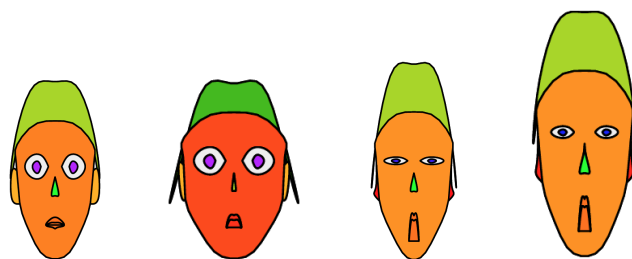
1. Maio em jejum;
2. Maio 6h após pequeno almoço
3. Junho em jejum;
4. Junho 6h após pequeno almoço.

Caras de Chernoff, 0h Maio, 6h Maio, 0h Junho, 6h Junho





(8)



(9)



(10)

Esta apresentação simultânea permite

1. reforçar a opinião de que existe excessiva variabilidade neste grupo de 10 voluntários, em qualquer dos quatro períodos observados, e nomeadamente que o voluntário 3 deveria ter sido excluído — repare-se na desconformidade da largura da correspondente linha, quando comparada com as outras;
2. observar que não há padrões definidos na variação — por exemplo, no voluntário 2 há grande diferença entre Maio e Junho, mas a variação do jejum para as 6h tem semelhanças, enquanto no voluntário 4 há grande semelhança em jejum em Maio e Junho, mas enquanto em Junho jejum e 6h são semelhantes, em Maio são muito distintos — , não se podendo concluir que existe um padrão definido na variação e em particular que haja um efeito consistente decorrente do tratamento com framboesas.

Capítulo 5

Gráficos das médias

Como Ramsey and Schafer (2002, p. 462) assinalam na nota introdutória ao capítulo sobre medições repetidas,

One highlighted methods involves reducing the several response variables to one or two summaries that address the question of interest, followed by performing analyses that use those summaries as responses. The summaries might consist of, for example, the average, the maximum, or the change in the repeated measure.

Vamos por isso focar-nos na análise de médias.

5.1 Gráficos: Médias com Barras de Confiança

Nos gráficos 5.1.0 a 5.8.0 em que se observam intervalos de confiança para a média de cada uma das variáveis em cada um dos tempos em que foram medidas (a metade esquerda corresponde a Maio, e a metade direita a Junho).

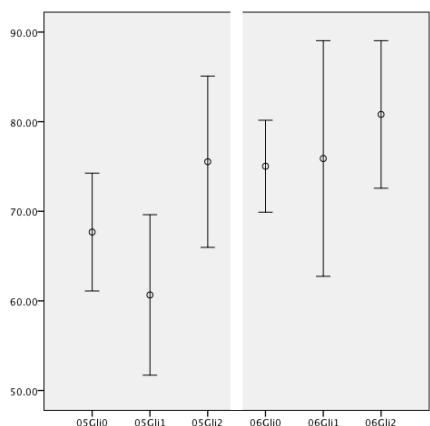


Figura 5.1.0: Glucose

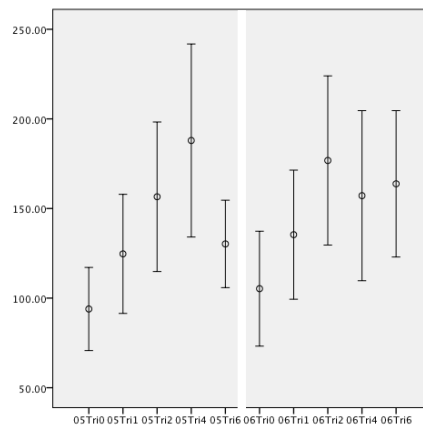


Figura 5.2.0: Triglicéridos

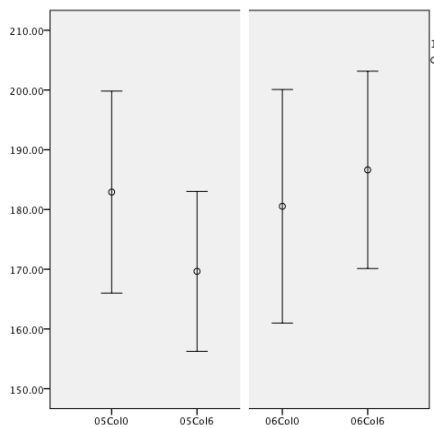


Figura 5.3.0: Colesterol

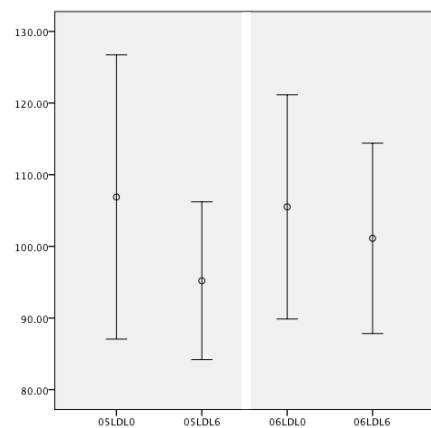


Figura 5.4.0: LDL

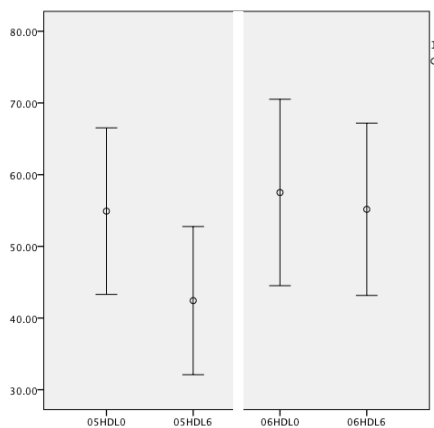


Figura 5.5.0: HDL

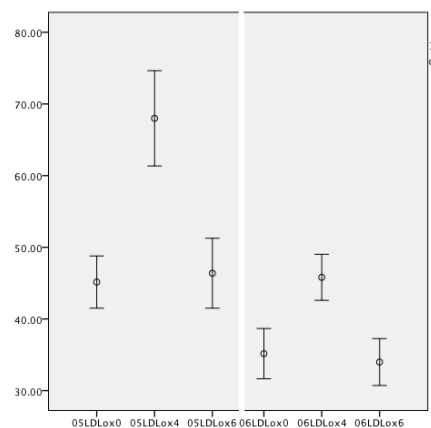


Figura 5.6.0: LDLox

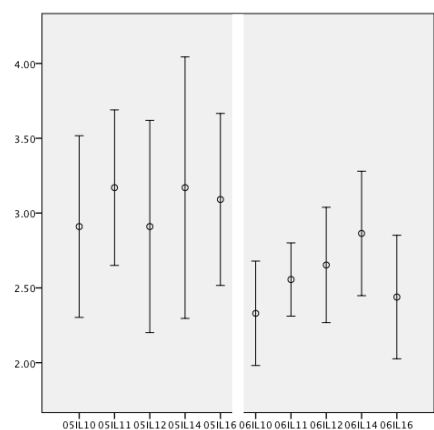


Figura 5.7.0: IL-1

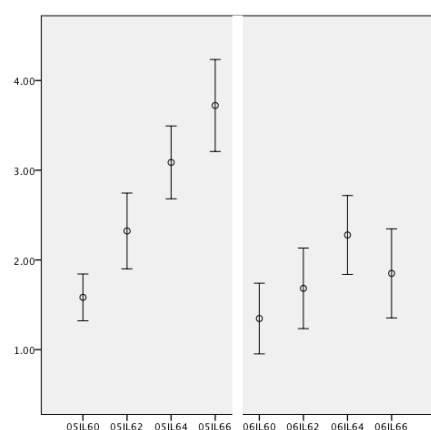


Figura 5.8.0: IL-6

Parece poder afirmar-se que houve subida do HDL e decréscimo do LDLox, do IL-1 e do IL-6, associados ao tratamento; por outro lado é mais difusa a associação entre tratamento e variação do colesterol e do LDL; também a interpretação no que se refere à associação do tratamento com a glicose e com triglicéridos não é óbvia.

Mas de novo temos que comentar que a coleção de voluntários é muito dispersa, e que na constituição de uma amostra adequada se devem usar critérios de exclusão.

5.2 Gráficos: Evolução das Variáveis nos Voluntários Cotejada com Médias e Limites de Confiança

Os gráficos 5.1 a 5.8 são uma reelaboração das anteriores Figuras 5.1.0 a 5.8.0, em que se representa a evolução de cada voluntário registando também, a negro e traço mais grosso, médias e limites do intervalo de confiança ($\bar{x} \pm t_{9,0.975} \frac{s}{\sqrt{10}}$). Observando que $t_{9,0.975} \frac{s}{\sqrt{10}} \approx 0.72$, a simples análise visual mostra que há um número elevado de observações fora do intervalo $\bar{x} - s, \bar{x} + s$, o que de novo nos leva a questionar o facto de não ter havido critérios de exclusão.

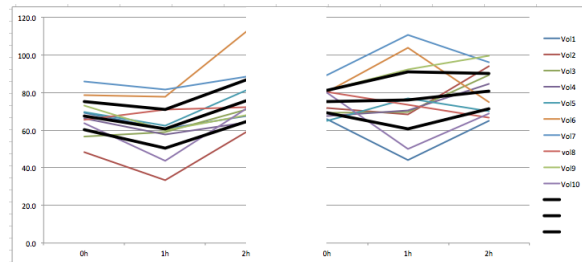


Figura 5.1: Glicose

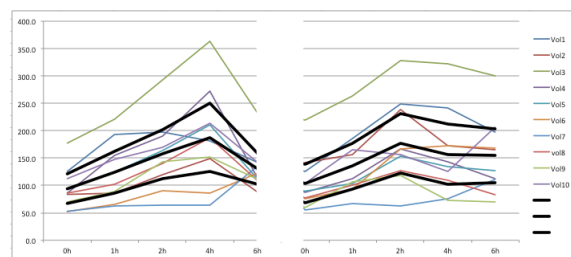


Figura 5.2: Triglicéridos

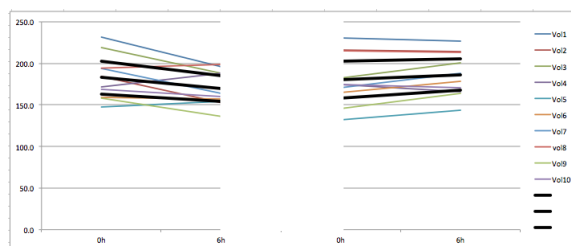


Figura 5.3: Colesterol

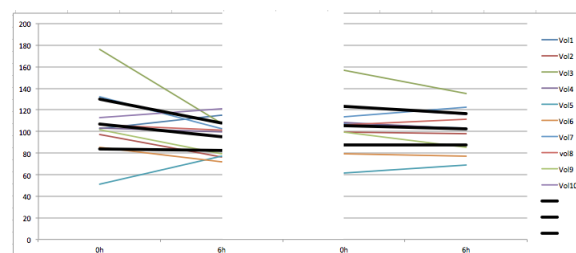


Figura 5.4: LDL

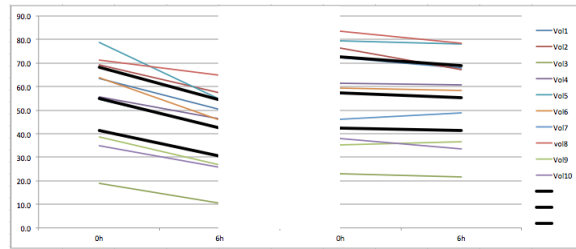


Figura 5.5: HDL

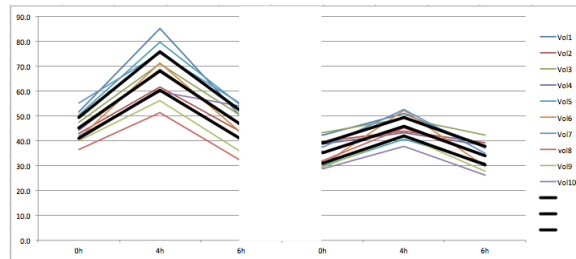


Figura 5.6: LDLox

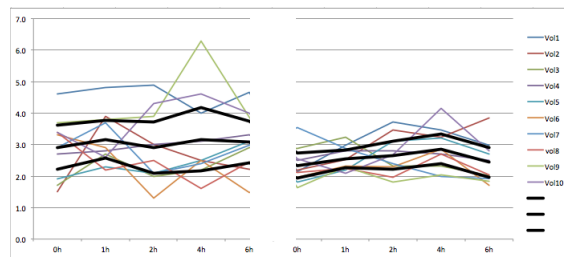


Figura 5.7: IL-1

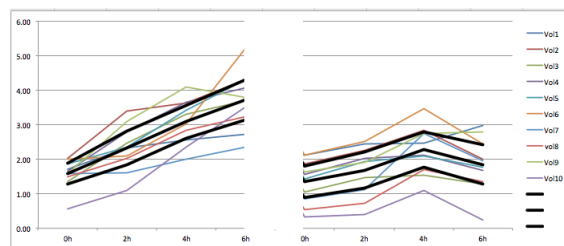


Figura 5.8: IL-6

Repare-se que estes gráficos pretendem mostrar padrões, pelo que optámos por usar o mesmo referencial para Maio e Junho, para observar se há alterações notáveis. Os gráficos reforçam a ideia de que há uma associação forte entre tratamento decréscimo substancial do LDLox e IL-6, mantendo o mesmo tipo de padrão. Mas de novo alertamos para a possibilidade de confundimento devido a hábitos alimentares e outra medicação tomada pelos voluntários.

Capítulo 6

Análise de Dados Longitudinais

Cabral e Gonçalves (2011) iniciam o curso de análise de dados longitudinais que guiou o meu estudo traduzindo Molenberghs and Verbeke:

Quando numa experiência os indivíduos são medidos repetidas vezes ao longo do tempo em relação a uma mesma característica, num ou mais grupos de tratamento, **sendo o próprio tempo um factor de interesse** diz-se que se têm dados longitudinais.

Citando ainda Cabral e Gonçalves (p. 3),

[...] o objetivo principal de um estudo longitudinal é o de caracterizar as alterações da variável resposta com o tempo [sendo ainda] objeto de interesse determinar se essas alterações se relacionam com um conjunto de covariáveis, isto é, com um conjunto de factores previamente escolhidos que não o tempo.

Como já repetidamente referimos, não há uma indicação clara do objetivo da modelação, nomeadamente qual a variável que interessaria modelar, e que covariáveis deveriam ser consideradas. Face à informação que pesquisei na internet (mas que por isso não sei se é fiável), seria levado a crer que o maior interesse do tratamento com framboesas, antioxidantes, seria ou controlar a inflamação das artérias causada pelo mau colesterol — e assim IL-1 e IL-6 — ou mesmo, diretamente, ver se a administração em grande escala daqueles antioxidante melhoravam o teor sérico do péssimo LDLox, diminuindo também o mau LDL e aumentando o bom HDL.

Face às anteriores considerações e à análise exploratória anteriormente apresentada, pareceria adequado passar à análise longitudinal das seguintes séries de dados:

- LDLox
- IL-1
- IL-6

visto serem as séries que tiveram maior alteração com a administração de uma dieta rica em framboesas, antioxidantes. Os dados são contínuos e a hipótese de ajustamento normal não foi rejeitada.

E de novo colocamos a questão — o que interessaria analisar, quais das variáveis deveríamos considerar como covariáveis?

Face à escassez de dados de LDLox, pareceu-nos pragmático abordar o estudo de IL-1 e de IL-6, isto é colocarmo-nos na perspectiva de o tratamento (ingestão de framboesas como antioxidante) tem como objetivo diminuir a inflamação vascular devida a depósitos de placas de LDL e sobretudo de LDLox.

Façamos porém alguns reparos:

1. Usar IL-1 e IL-6 como variáveis resposta e LDLox como covariável não parece fazer muito sentido, uma vez que os instantes temporais em que as três variáveis foram medidas não são os mesmos. Em particular, LDLox é medido apenas em jejum, 4h após pequeno almoço e 6h após pequeno almoço, falhando por isso a medição duas horas após pequeno almoço. Note-se também que os intervalos entre medições de LDLox e de IL-1 não são iguais.
2. IL-1 e IL-6 são medidas ambas em jejum, 2h, 4h e 6h após o pequeno almoço; descartando a leitura de IL-1 uma hora após o pequeno almoço ficaríamos com medições em instantes coincidentes, e a intervalos regulares, pelo que se poderia considerar a modelação conjunta das interleucinas.

Porém a análise conjunta de dados longitudinais é muito complexa (“*far from trivial*”), no dizer de Ramsey and Schafer (2002, p. 483) — conferir a citação Fig. 2 — e não encontramos nenhum

16.5.1 Two-Sample Analysis with More Than Two Responses

Extending the Hotelling's T^2 tools to situations involving more than two response variables is far from trivial. It is nearly impossible to tackle the calculations for T^2 without an understanding of linear algebra, including the use of matrix algebra. Nevertheless the Hotelling T^2 test statistic for comparing equality of two q -dimensional means from populations with the same variance and correlation structure, can be calculated with statistical computer packages. Suppose that there are q responses for each subject and that there are n_1 subjects in group 1 and n_2 subjects in group 2. A p -value is obtained by comparing

$$F\text{-statistic} = (n_1 + n_2 - q - 1)T^2 / [q(n_1 + n_2 - q)]$$

to an F -distribution with q and $(n_1 + n_2 - q - 1)$ degrees of freedom. Drawbacks of this tool are that it is not robust if the variances and correlations differ in the two populations and it is not robust if the population distributions are not normal.

Figura 6.1: Citação de Ramsey and Schaffer

exemplo que nos possa servir de guia.

3. Face ao acima referido, parece fazer sentido abordar separadamente a modelação de IL-1 e de IL-6, para o que usámos, de acordo com as recomendações de Cabral e Gonçalves (2012), o *package nlme* do *R*.

Mas antes de prosseguir para essa modelação parece adequado começar por uma visualização dos dados que seja reveladora de padrões.

6.1 IL-1

Para simplicidade de leitura, começamos por reproduzir o gráfico 6.2 dos perfis individuais no que se refere a IL-1, em Maio (controle) e Junho (dados experimentais), recordando que em negro mais carregado se representaram valores médios e limites de confiança (95%) no que refere a 0h, 1h, 2h, 4h e 6h.

É óbvio que em Junho o padrão ao longo das seis horas de observação é mais regular, e que houve não só um decréscimo dos valores, como a variabilidade inter indivíduos é menor.

Completamos essa informação com gráficos estelares que mostram a evolução do teor sérico de IL-1 ao longo do tempo no que respeita os 10 voluntários (na primeira linha Maio, na segunda Junho). É bastante evidente que não há padrões discerníveis na forma como IL-1 evolui nos voluntários, quer no período de controle quer no período experimental, e também que não há similitude entre esses dois períodos.

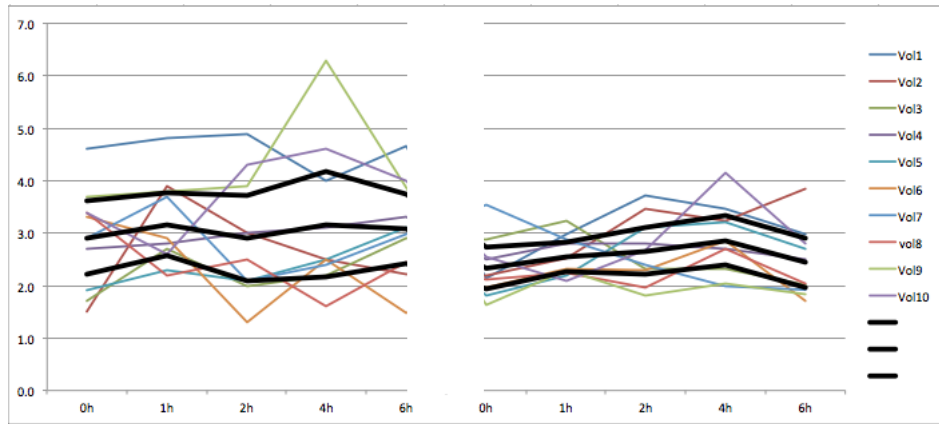
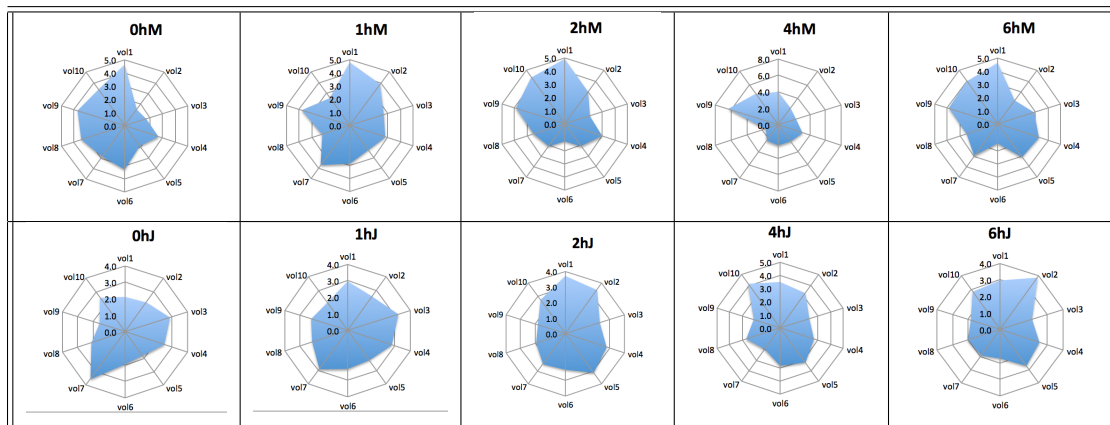
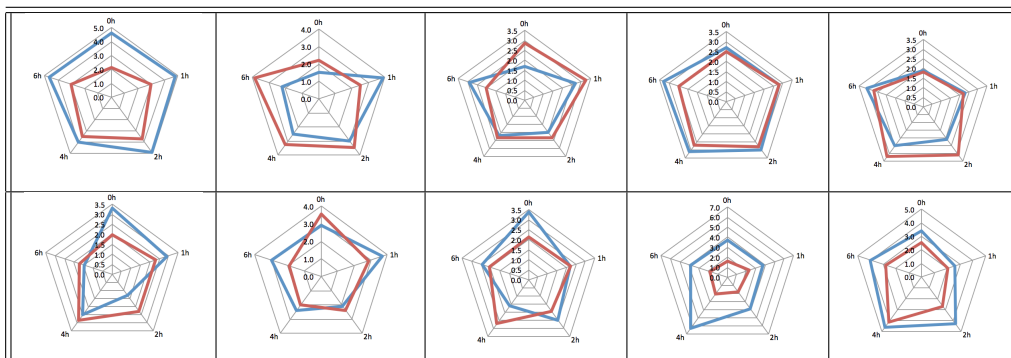


Figura 6.2: Variação de IL-1 ao longo do tempo. Além da representação do que se passa com cada voluntário, representa-se a negro grosso as linhas unindo os valores médios e limites dos intervalos de confiança para a média (95%).



Por outro lado, a mesma informação é fornecida nos gráficos de radar abaixo, mas desta vez centrados no que se passa com cada voluntário às 0h, 1h, 2h, 4h e 6h, em Maio (azul) e em Junho (laranja). Os voluntários estão ordenados de 1 a 5 na linha superior, e de 6 a 10 na inferior



A redução dos valores de IL-1 em Junho é bastante evidente nos voluntários 1, 4, 5 e 9.

Como facilmente se constata, a reação dos diversos voluntários ao tratamento com framboesas segue diversos padrões, embora se note que em alguns deles (1, 4, 5, 6, 9) se possa falar de perfis semelhantes, em que apenas há uma alteração de escala. Fica no entanto o alerta para a eventual pouca adequação da modelação que vai ser feita.

Pareceu-nos inicialmente adequado seguir de perto a abordagem de Cabral e Gonçalves (2012), exemplo 6.2 (repetimos aliás a análise em *RStudio*, para nos certificarmos de que usamos o código apropriado).

Para tal, carregamos as bibliotecas *tlck*, *nlme*, *lattice*, *ggplot2* e o ficheiro *DadAdilsM* contendo os registos das variáveis em Maio e em Junho, veja-se Tabela 1.9; codificamos os tempos de medição usando a variável *HOR*, com valores 0 (jejum), 1,2,4 e 6 (tempo em horas decorrido depois do pequeno-almoço, e o mês com a variável *MES* (0=Maio, 1=Junho).

Os gráficos acima levam a crer que há diferença entre os registos de casos (Junho) e controles (Maio), mas que em cada um desses períodos a hora de medição tem pouca influência.

No Capítulo introdutório referimos brevemente o formalismo da análise de dados longitudinais; recomendamos Cabral e Gonçalves (2011) para uma introdução esclarecedora; como já referimos no início deste estudo, a normalidade dos dados praticamente nunca foi rejeitada, mas essa decisão está longe de ser convincente, visto que cada uma das séries ajustadas tinha apenas 10 dados.

De qualquer forma, como parece mais adequado considerar modelos com feitos aleatórios porque — citando de novo Cabral e Gonçalves (2011, p. 17)

Os modelos com efeitos aleatórios são mais úteis quando o objetivo é descrever as alterações da resposta média de cada indivíduo e a relação destas alterações com as co-variáveis. Assim sendo, estes modelos têm como finalidade realizar inferências sobre o indivíduo e não sobre a população.

O ajustamento por normais não tem relevância, pois esses modelos “*permitem a acomodação de respostas quer sigam uma distribuição Gaussiana ou não*”.

Começamos por investigar uma variedade de modelos para *IL-1* em função de *MES*, ou *MES+HOR*, ou ainda *MES+HOR+MES*HOR* — pois qualquer modelo que não use *MES* como regressor não tem interesse para o que investigamos —, admitindo efeitos aleatórios no declive e ordenada na origem, ou apenas na ordenada na origem, ou apenas no declive, investigando por isso os seguintes modelos:

```
Mo1<-lme(IL1~MES, random=~0|VOL,data=dados,method="ML")
Mo2<-lme(IL1~MES, random=~1|VOL,data=dados,method="ML")
Mo3<-lme(IL1~MES, , random=~(0+MES)|VOL,data=dados,method="ML")
Mo4<-lme(IL1~MES, random=~(1+MES)|VOL,data=dados,method="ML")

Mo5<-lme(IL1~HOR+MES, random=~0|VOL,data=dados,method="ML")
Mo6<-lme(IL1~HOR+MES, random=~1|VOL,data=dados,method="ML")
Mo7<-lme(IL1~HOR+MES, random=~(0+MES)|VOL,data=dados,method="ML")
Mo8<-lme(IL1~HOR+MES, random=~(0+HOR)|VOL,data=dados,method="ML")
Mo9<-lme(IL1~HOR+MES, random=~(0+MES+HOR)|VOL,data=dados,method="ML")
Mo10<-lme(IL1~HOR+MES, random=~(1+MES)|VOL,data=dados,method="ML")
Mo11<-lme(IL1~HOR+MES, random=~(1+HOR)|VOL,data=dados,method="ML")
Mo12<-lme(IL1~HOR+MES, random=~(1+MES+HOR)|VOL,data=dados,method="ML")

Mo13<-lme(IL1~HOR+MES+HOR*MES, random=~1|VOL,data=dados,method="ML")
Mo14<-lme(IL1~HOR+MES+HOR*MES, random=~(0+MES)|VOL,data=dados,method="ML")
Mo15<-lme(IL1~HOR+MES+HOR*MES, random=~(0+HOR)|VOL,data=dados,method="ML")
Mo16<-lme(IL1~HOR+MES+HOR*MES, random=~(0+HOR*MES)|VOL,data=dados,method="ML")
```

```

Mo17<-lme(IL1~HOR+MES+HOR*MES, random=~(1+MES)|VOL,data=dados,method="ML")
Mo18<-lme(IL1~HOR+MES+HOR*MES, random=~(1+HOR)|VOL,data=dados,method="ML")
Mo19<-lme(IL1~HOR+MES+HOR*MES, random=~(1+HOR*MES)|VOL,data=dados,method="ML")
Mo20<-lme(IL1~HOR+MES+HOR*MES, random=~(1+MES+HOR)|VOL,data=dados,method="ML")
Mo21<-lme(IL1~HOR+MES+HOR*MES, random=~(1+MES+HOR*MES)|VOL,data=dados,method="ML")
Mo22<-lme(IL1~HOR+MES+HOR*MES, random=~(1+MES+HOR+HOR*MES)|VOL,data=dados,method="ML")

```

Uma breve formalização inicial foi feita na Secção 1.3. Comentemos porém que estes modelos estão ordenados por grau de complexidade quer no uso de regressores quer na eventual possibilidade de se considerarem efeitos aleatórios. Por exemplo

```

Mo1<-lme(IL1~MES, random=~0|VOL,data=dados,method="ML")
Mo2<-lme(IL1~MES, random=~1|VOL,data=dados,method="ML")
Mo3<-lme(IL1~MES, , random=~(0+MES)|VOL,data=dados,method="ML")
Mo4<-lme(IL1~MES, random=~(1+MES)|VOL,data=dados,method="ML")
...
Mo22<-lme(IL1~HOR+MES+HOR*MES, random=~(1+MES+HOR+HOR*MES)|VOL,data=dados,method="ML")

```

correspondem respetivamente a

- $IL1 = \beta_0 + \beta_1 MES + \varepsilon$
- $IL1_i = (\beta_0 + b_{0i}) + \beta_1 MES + \varepsilon_i$
- $IL1_i = \beta_0 + (\beta_1 + b_{1i})MES + \varepsilon$
- $IL1_i = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})MES + \varepsilon$
- ...
- $IL1_i = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})MES + (\beta_2 + b_{2i})HOR + (\beta_3 + b_{3i})HOR * MES + \varepsilon_i$

De forma condensada e genérica o modelo linear misto é descrito como

$$Y_i = X_i \beta + Z_i b_i + \varepsilon_i$$

onde Y_i é a matriz das variáveis respostas para o indivíduo i (no nosso caso, $IL1$ é a matriz 10×10 , visto que há 10 indivíduos e há 10 combinações de (Mês,Hora)), β é o vetor dos efeitos fixos, X_i é a matriz das covariáveis dos efeitos fixos, b_i é o vetor dos efeitos aleatórios, Z_i é a matriz das covariáveis dos efeitos aleatórios e ε_i é a matriz dos erros aleatórios dentro do grupo. Por exemplo, no Mo8 (adiante renomeado Mod.II, veja-se o output no apêndice 1), $X = (MES, HOR)$ e $b = HOR$. Para mais detalhes veja-se Cabral e Gonçalves (2011, 2.2).

$$b_i \sim N(0, D), \quad \varepsilon_i \sim N(0, \Sigma_i)$$

onde b_i e ε_i são independentes entre si e para os diferentes grupos.

No nosso estudo, o objetivo fundamental é investigar se houve uma efetiva descida de $IL1$ com o tratamento, isto é se é rejeitada a $H_0 : \beta = 0$ vs. $H_1 : \beta = 0 < 0$. A explicitação da interpretação dos outputs do que adiante renomeamos como Mod.VI e Mod.VIII, e dos modelos Mod.XI e Mod.XV, em

que juntamos como covariáveis TRI e LDLox, respectivamente, esclarece com mais detalhe o que se pretende.

Este trabalho inicial mostrou alguns problemas de convergência em alguns dos modelos (em particular no que se refere ao Mo4, adiante rebatizado como mod.VIII), o que nos obrigou a alterar alguns critérios no que respeita número de iteradas a considerar. Na segunda fase consideramos por isso o *script*

```
mod.I<-lme(IL1~HOR+MES, random=~1|VOL,data=dados,method="ML")
summary(mod.I)
mod.II<-lme(IL1~HOR+MES, random=~(0+HOR)|VOL,data=dados,method="ML")
summary(mod.II)
mod.III<-lme(IL1~HOR+MES+HOR*MES, random=~(0+HOR)|VOL,data=dados,method="ML")
summary(mod.III)
mod.IV<-lme(IL1~HOR+MES+HOR*MES, random=~1|VOL,data=dados,method="ML")
summary(mod.IV)
mod.V<-lme(IL1~HOR+MES+HOR*MES, random=~(1+HOR)|VOL,data=dados,method="ML")
summary(mod.V)
mod.VI<-lme(IL1~MES, random=~1|VOL,data=dados,method="ML")
summary(mod.VI)
mod.VII<-lme(IL1~MES, random=~(0+MES)|VOL,data=dados,method="ML")
summary(mod.VII)
mod.VIII<-lme(IL1~MES, random=~(1+MES)|VOL,data=dados,method="ML")
summary(mod.VIII)
mod.IX<-lme(IL1~MES+HOR*MES, random=~(1+MES)|VOL,data=dados,method="ML")
summary(mod.IX)
> mod.X<-lme(IL1~HOR+MES, random=~(1+MES)|VOL,data=dados,method="ML")
summary(mod.X)
```

O correspondente *output* está reproduzido no apêndice 1 deste capítulo. Limitamo-nos aqui a registar que

1. Como é natural, por IL-1 ser uma grandeza positiva, em qualquer dos modelos indica-se que o correspondente coeficiente em Maio é não nulo (altamente significativo)
2. Nos modelos mod.I a mod.VI considera-se significativa (valor- $p < 0.05$) uma descida do IL-1 que ronda -0.5
3. Nenhum coeficiente das outras variáveis consideradas é significativo.
4. Os modelos mod.VII a mod.X parecem melhorar o ajustamento (critério de Akaike), mas em contrapartida a alteração de IL-1 entre Maio e Junho devida ao efeito MES não é significativa ao nível 0.05.

e a tabular o coeficiente da alteração de IL-1 devido ao efeito MES e o correspondente p -value em cada um desses 10 modelos, bem como o AIC:

<i>modelo</i>	<i>coef. MES</i>	<i>valor - p</i>	<i>AIC</i>
<i>mod.I</i>	-0.4549896	0.0032	247.1762
<i>mod.II</i>	-0.4735782	0.0026	248.9941
<i>mod.III</i>	-0.4950724	0.0420	250.9799
<i>mod.IV</i>	-0.5118980	0.0330	249.0742
<i>mod.V</i>	-0.4950724	0.0420	250.9799
<i>mod.VI</i>	-0.4510544	0.0034	245.8660
<i>mod.VII</i>	-0.4347379	0.0128	258.4832
<i>mod.VIII</i>	-0.4515758	0.0936	230.3403
<i>mod.IX</i>	-0.5108365	0.1038	233.0579
<i>mod.X</i>	-0.4561743	0.0925	231.1879

Assim, o modelo *mod.VIII* $< -lme(IL1 \sim MES, random = \sim (1 + MES))$ destaca-se por ter **AIC=230.3403 inferior aos dos outros**, apresentando-se os modelos *mod.X* $< -lme(IL1 \sim HOR + MES, random = \sim (1 + MES))$ (AIC=231.1879) e *mod.IX* $< -lme(IL1 \sim MES + HOR * MES, random = \sim (1 + MES))$ (AIC=233.0579) como os competidores mais próximos.

No entanto, se adotarmos como **critério de avaliação principal uma valor-p < 0.05 no que respeita a alteração significativa de IL-1 devido ao efeito MES, deveremos considerar como preferível o modelo VI.**

Para além de IL-1, também a variável Triglicéridos foi medida em jejum (0h), e nos mesmos períodos 1h, 2h, 4h, 6h após o pequeno almoço. Será possível melhorar mais usando por exemplo TRI como covariável?

No Apêndice 2 registamos o *utput* de

```
> mod.XI<-lme(IL1~TRI+MES+HOR*MES, random=~(1+MES)|VOL,data=dados,method="ML")
> mod.XII<-lme(IL1~MES+HOR*MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
>mod.XIII<-lme(IL1~TRI+MES+HOR*MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
```

Tal como nos modelos anteriores, apenas a ordenada na origem (valor estimado do IL-1 em Maio) é significativamente diferente de 0. A redução de IL-1 de Maio para Junho não é detetada como significativa, veja-se o quadro abaixo:

<i>modelo</i>	<i>coef. MES</i>	<i>valor - p</i>	<i>AIC</i>
<i>mod.XI</i>	-0.5302166	0.0875	233.0079
<i>mod.XII</i>	-0.7849047	0.0798	234.3419
<i>mod.XIII</i>	-0.7849048	0.0798	234.3419

Pareceria mais interessante usar LDLox como covariável, mas isso exige que os registos de IL-1 correspondentes a 1h e 2h sejam apagados (pois não houve os correspondentes registos de LDLox),

como adiante explicamos com mais detalhe. Quer isso dizer que reduzimos de 100 para 60 dados de IL-1.

Por curiosidade, investigamos também o que se obtém quando se usa TRI como covariável usando apenas os 60 dados correspondentes a 0h, 4h, 6h:

```
> mod.XIV<-lme(IL1~TRI+MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
```

Este modelo trouxe finalmente algum sucesso no sentido em que há redução assinalável de AIC:

<i>modelo</i>	<i>coef. MES</i>	<i>valor - p</i>	<i>AIC</i>
<i>mod.XIV</i>	-0.6838579	0.2054	158.793

mas a redução de IL-1 de Maio para Junho não é assinalada como significativa, e esta “tortura”deliberada para os dados confessarem o que pretendemos não é defensável. O *output* completo está registado no Apêndice 3.

Atendendo ao que vimos na Internet sobre o contexto do problema, veja-se os capítulos de Introdução e de Análise Inicial, o que faria porventura sentido seria investigar se a alteração do LDLox entre Maio e Junho contribuía para explicar IL-1 naqueles dois períodos. Mas tal não é possível com o ficheiro de dados que estamos a usar, porque não há medições de LDLox 1h e 2h após pequeno almoço. Por isso, se dermos a instrução

```
> mod.XIVb<-lme(IL1~LDLox+MES+HOR*MES,
random=~(1+MES)|VOL, data=dados,method="ML")
```

obtemos simplesmente o aviso de erro

```
Error in na.fail.default(list(MES = c(OL, OL, OL, OL, OL, OL, OL, OL, :
missing values in object
```

Experimentamos por isso eliminar os dados de IL-1 registados 1h e 2h após o pequenos almoço, para investigarmos o modelo acima descrito:

```
> mod.XV<-lme(IL1~LDLox+MES+LDLox*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
```

com os resultados

<i>modelo</i>	<i>coef. MES</i>	<i>valor - p</i>	<i>AIC</i>
<i>mod.XV</i>	-1.6825549	0.0678	154.8474

(*Output* completo no Apêndice 4.)

Assim, em termos de AIC, houve uma melhoria. Mas o efeito MES, de maior magnitude, não é significtaivo.

Portanto, em minha opinião, o mod.VI apesar da sua simplicidade é o melhor.

6.2 Conclusões provisórias

Na coleção inicial de modelos (I a X) verificamos que a variável MES tem como efeito alterar a ordenada na origem e a variável HOR não é interessante. Nos modelos I a VII verifica-se que há um decréscimo (significativo, $p \approx 0.03$) que ronda -0.5 entre maio e Junho, enquanto o coeficiente de Hora e de Mês*Hora não são significativamente diferentes de 0. De facto, observando os gráficos de perfis, verifica-se que entre Maio e Junho IL-1 decresceu cerca de 0.5, e que a linha poligonal quer em Maio quer em Junho pode ser considerada uma ligeira variação amostral de uma linha horizontal.

Observa-se também que considerar efeitos aleatórios no que se refere à ordenada na origem baixa o AIC, o que faz sentido, face às críticas que fizemos à amostragem feita.

Observe-se que o melhor ajustamento, tomando o valor do AIC como critério, é obtido com o mod.VIII, sendo os mod.IX e mod.X competidores próximos; mas nestes curiosamente o valor- p associado ao coeficiente estimado para o Mês é 0.09, aparentemente os efeitos aleatórios no que se refere à ordenada na origem seriam suficientes para justificar o bom ajustamento do modelo. Note-se que naqueles três modelos se considere efeitos aleatórios no que se refere quer à ordenada na origem quer ao regressor MES.

As informações que encontramos na *www*, e que relatamos no capítulo introdutório, levam-nos a crer que triglicéridos e LDLox contribuem para a inflamação no sistema circulatório, e por isso não é de estranhar que aqueles e o mod.XV (em que mais uma vez se consideram efeitos aleatórios no que se refere quer à ordenada na origem quer ao regressor MES) sejam mais adequados — mas este último à custa de cortar informação, pois passamos a considerar apenas as 60 observações (10 voluntários, em jejum, ao fim de 4 horas e ao fim de 6 horas, em Maio e em Junho, que são as únicas comuns às duas variáveis que se pretende correlacionar). Encontramos também a informação de que alterações metabólicas da glicose teriam influência, mas como as medições desta variável são feitas apenas às 0h, 1h e 2h, para as podermos utilizar teríamos que deitar fora os dados de IL-1 das 4h e 6h, o que decerto não faz sentido.

Pensamos por isso que a repetida anotação de que num trabalho futuro mais planeado haverá que distinguir cuidadosamente variáveis e covariáveis a usar, e recolher os dados por forma a não ocorrerem estes problemas devidos a dados omissos.

6.3 Comentário crítico

Após este trabalho que não quisemos deixar de relatar, há que fazer os seguintes reparos:

- Enquanto no exemplo de Cabral e Gonçalves (2012, 6.2) há três conjuntos de dados independentes, e o objetivo será comparar os efeitos das dietas 2 e 3 com os da *baseline* dieta 1, no nosso caso pretende-se comparar séries emparelhadas de observações, considerando-se que as medições iniciais feitas em Maio servem de controle, sendo as medições feitas em Junho os resultados experimentais.
- Nos dados *bodyweight* analisados por Cabral e Gonçalves (2012, 6.2) há logo à partida a expectativa de que as sequências de dados individuais devem ter uma tendência crescente, por as medições serem efetuadas com um intervalo de tempo suficiente para eliminar flutuações aleatórias sem interesse. No estudo que nos ocupa, as medições são em tempos muito próximos, durante o processo de digestão, havendo assim padrões muito complexos determinados por fatores metabólicos.

Face a estas observações, fomos repensar na análise feita em Ramsey and Schafer (2012, cap. 16, *Repeated Measures and other Multivariate Responses*), nomeadamente porque a investigação sobre o papel do hipocampo nas memórias de curta e de longa duração dos primatas nos pareceu bastante informativo.

Porém a leitura atenta mostrou-nos que nos dados que temos que analisar há uma falha fatal: cada indivíduo é de facto o seu próprio controle (em Maio, relativamente ao resultado experimental de Junho) — mas não há um grupo experimental e outro de controle, para haver conclusões válidas seria necessário que o o grupo inicial fosse dividido em dois subgrupos, de preferência equilibrados, um dos quais tivesse a dieta com framboesas e outro não, para analisar se os padrões em Junho eram diferentes nos dois grupos. (Isto para além dos já insistentemente referidos problemas no que se refere a critérios de inclusão no estudo, ao facto de a amostra ser “de conveniência” e de tamanho muito reduzido, e não ter havido bloqueamento para limitar confundimento.)

Por outro lado, o interessante exemplo tratado no mesmo capítulo sobre o efeito de fibras de aveia no controle de colesterol mostra também que os dados cuja análise é pedida não são adequados, mais uma vez por não haver a possibilidade de comparação. Os dados seriam interessantes se depois das medições em Junho houvesse um período de *washout* em que os voluntários não comessem framboesas, e novas medições em Julho, a análise poderia ter algum interesse (com as limitações já referidas no que se refere à amostra e ao confundimento, havendo a este respeito que estabelecer um protocolo muito estrito sobre o tipo de alimentos e aditivos que poderiam ingerir).

Nestas condições, penso que é dispensável prosseguir para a análise das outras variáveis, pois do ponto de vista de metodologia não haveria qualquer adição, e os problemas atrás apontados são comuns a todas as variáveis.

Apêndice 1: *output mod.I a mod.X*

```
> dados<-read.table("DadAdilsM.txt",header=TRUE,dec=".")
> dados
```

	VOL	MES	HOR	IL1	IL6	COL	HDL	LDL	LDLox	TRI	GLU
1	1	0	0	4.60	1.7	232.0	63.4	102.7	51.6	125.7	69.0
2	2	0	0	1.50	2.0	184.5	69.1	97.1	42.8	83.7	48.3
3	3	0	0	1.70	1.4	219.0	18.8	176.3	47.5	177.4	56.6
4	4	0	0	2.70	1.7	171.7	55.6	103.2	44.2	85.8	66.2
5	5	0	0	1.90	1.9	147.9	78.8	51.2	50.4	94.6	69.5
6	6	0	0	3.30	2.0	158.8	63.7	85.4	41.1	52.4	78.8
7	7	0	0	2.90	1.6	193.8	54.8	132.4	55.3	52.9	85.9
8	8	0	0	3.40	1.5	194.3	71.4	106.3	36.6	85.4	65.5
9	9	0	0	3.70	1.6	158.4	38.5	101.5	40.0	69.3	73.2
10	10	0	0	3.40	0.6	168.8	35.0	112.8	42.0	112.1	63.8
11	1	0	1	4.80	NA	NA	NA	NA	NA	193.0	60.5
12	2	0	1	3.90	NA	NA	NA	NA	NA	86.4	33.4
13	3	0	1	2.70	NA	NA	NA	NA	NA	220.6	59.1
14	4	0	1	2.80	NA	NA	NA	NA	NA	155.6	57.5
15	5	0	1	2.30	NA	NA	NA	NA	NA	124.0	62.3
16	6	0	1	2.90	NA	NA	NA	NA	NA	65.7	77.7
17	7	0	1	3.70	NA	NA	NA	NA	NA	62.8	81.5
18	8	0	1	2.20	NA	NA	NA	NA	NA	101.5	71.0
19	9	0	1	3.80	NA	NA	NA	NA	NA	89.5	59.9

20	10	0	1	2.60	NA	NA	NA	NA	NA	147.8	43.7
21	1	0	2	4.90	2.3	NA	NA	NA	NA	196.9	67.6
22	2	0	2	3.00	3.4	NA	NA	NA	NA	119.6	58.8
23	3	0	2	2.00	2.5	NA	NA	NA	NA	293.3	70.8
24	4	0	2	3.00	2.8	NA	NA	NA	NA	189.7	64.3
25	5	0	2	2.10	2.3	NA	NA	NA	NA	163.3	81.0
26	6	0	2	1.30	2.1	NA	NA	NA	NA	90.5	112.5
27	7	0	2	2.10	1.6	NA	NA	NA	NA	64.4	88.5
28	8	0	2	2.50	2.0	NA	NA	NA	NA	141.0	72.4
29	9	0	2	3.90	3.1	NA	NA	NA	NA	143.7	68.1
30	10	0	2	4.30	1.1	NA	NA	NA	NA	169.9	71.3
31	1	0	4	4.00	2.6	NA	NA	NA	85.1	180.7	NA
32	2	0	4	2.50	3.6	NA	NA	NA	61.7	148.8	NA
33	3	0	4	2.20	3.3	NA	NA	NA	71.1	363.6	NA
34	4	0	4	3.10	3.7	NA	NA	NA	68.6	272.8	NA
35	5	0	4	2.50	3.4	NA	NA	NA	79.8	210.5	NA
36	6	0	4	2.50	3.0	NA	NA	NA	71.2	85.7	NA
37	7	0	4	2.40	2.0	NA	NA	NA	75.3	63.8	NA
38	8	0	4	1.60	2.8	NA	NA	NA	51.3	187.7	NA
39	9	0	4	6.30	4.1	NA	NA	NA	56.0	152.4	NA
40	10	0	4	4.60	2.3	NA	NA	NA	59.8	213.5	NA
41	1	0	6	4.65	2.7	196.1	50.3	115.3	49.7	141.7	NA
42	2	0	6	2.21	4.3	152.8	57.3	76.2	43.7	87.4	NA
43	3	0	6	2.90	3.7	188.6	10.5	107.4	50.0	231.0	NA
44	4	0	6	3.30	4.1	188.3	46.3	99.7	46.3	112.0	NA
45	5	0	6	3.10	4.3	154.7	54.6	77.1	54.2	111.6	NA
46	6	0	6	1.48	5.2	156.9	45.7	71.6	43.4	122.2	NA
47	7	0	6	2.97	2.3	163.9	42.4	102.7	55.0	135.9	NA
48	8	0	6	2.44	3.2	199.2	64.8	101.3	32.1	107.4	NA
49	9	0	6	3.86	3.8	136.3	26.7	79.4	35.4	111.0	NA
50	10	1	6	4.00	3.5	159.7	25.7	121.3	54.0	142.1	NA
51	1	1	0	2.10	2.1	230.2	72.3	106.7	42.2	126.3	65.7
52	2	1	0	2.20	1.9	216.4	76.3	99.8	39.8	142.5	71.9
53	3	1	0	2.90	1.0	182.2	23.0	156.7	43.2	219.3	69.0
54	4	1	0	2.50	1.6	174.2	61.4	108.4	38.6	88.0	67.3
55	5	1	0	1.80	1.4	132.0	79.6	61.2	30.0	89.7	65.1
56	6	1	0	2.00	2.1	164.5	59.4	79.2	30.3	75.5	80.4
57	7	1	0	3.50	0.9	171.0	46.2	113.6	37.4	54.8	89.4
58	8	1	0	2.10	0.5	215.1	83.7	106.3	32.0	77.3	80.5
59	9	1	0	1.60	1.6	145.6	35.3	99.8	29.4	59.7	81.0
60	10	1	0	2.50	0.3	174.2	37.9	123.4	28.7	104.0	80.0
61	1	1	1	3.00	NA	NA	NA	NA	NA	186.1	44.1
62	2	1	1	2.50	NA	NA	NA	NA	NA	156.0	68.4
63	3	1	1	3.20	NA	NA	NA	NA	NA	263.5	69.5
64	4	1	1	2.80	NA	NA	NA	NA	NA	112.5	70.3
65	5	1	1	2.20	NA	NA	NA	NA	NA	103.6	77.1
66	6	1	1	2.30	NA	NA	NA	NA	NA	92.7	103.6
67	7	1	1	2.90	NA	NA	NA	NA	NA	66.7	110.5
68	8	1	1	2.20	NA	NA	NA	NA	NA	101.0	73.4
69	9	1	1	2.30	NA	NA	NA	NA	NA	106.6	92.1
70	10	1	1	2.10	NA	NA	NA	NA	NA	165.3	49.9

71	1	1	2	3.70	2.5	NA	NA	NA	NA	249.4	64.8
72	2	1	2	3.50	2.3	NA	NA	NA	NA	239.0	93.8
73	3	1	2	2.30	1.5	NA	NA	NA	NA	328.3	89.2
74	4	1	2	2.80	2.0	NA	NA	NA	NA	166.2	84.4
75	5	1	2	3.10	1.9	NA	NA	NA	NA	154.0	69.9
76	6	1	2	2.30	2.5	NA	NA	NA	NA	166.7	74.9
77	7	1	2	2.40	1.1	NA	NA	NA	NA	62.9	95.9
78	8	1	2	2.00	0.7	NA	NA	NA	NA	126.7	66.8
79	9	1	2	1.80	1.9	NA	NA	NA	NA	117.9	99.7
80	10	1	2	2.70	0.4	NA	NA	NA	NA	156.8	68.7
81	1	1	4	3.50	2.5	NA	NA	NA	50.9	241.9	NA
82	2	1	4	3.20	2.8	NA	NA	NA	44.0	172.4	NA
83	3	1	4	2.30	1.5	NA	NA	NA	49.1	322.9	NA
84	4	1	4	2.70	2.1	NA	NA	NA	43.3	143.4	NA
85	5	1	4	3.20	2.1	NA	NA	NA	40.7	134.8	NA
86	6	1	4	2.90	3.5	NA	NA	NA	52.4	172.2	NA
87	7	1	4	2.00	2.7	NA	NA	NA	52.6	75.4	NA
88	8	1	4	2.70	1.7	NA	NA	NA	45.3	108.9	NA
89	9	1	4	2.10	2.7	NA	NA	NA	42.1	73.4	NA
90	10	1	4	4.10	1.1	NA	NA	NA	37.7	126.2	NA
91	1	1	6	3.00	3.0	227.0	67.8	103.4	37.0	198.2	NA
92	2	1	6	3.90	2.0	213.9	67.3	98.2	39.2	165.8	NA
93	3	1	6	2.00	1.3	201.0	21.5	134.9	42.3	299.8	NA
94	4	1	6	2.50	1.7	165.8	60.8	102.5	38.0	113.0	NA
95	5	1	6	2.70	1.8	143.8	78.0	69.2	31.0	127.3	NA
96	6	1	6	1.70	2.4	178.9	58.5	77.3	29.7	168.8	NA
97	7	1	6	1.90	2.0	187.7	49.0	122.9	35.0	111.6	NA
98	8	1	6	2.10	1.3	213.4	78.3	111.3	33.8	82.8	NA
99	9	1	6	1.80	2.8	164.0	36.7	85.7	27.7	70.6	NA
100	10	1	6	2.80	0.2	170.8	33.7	116.0	26.2	206.7	NA

Os sumários dos modelos acima indicados são:

```
> mod.I<-lme(IL1~HOR+MES, random=~1|VOL,data=dados,method="ML")
> summary(mod.I)
Linear mixed-effects model fit by maximum likelihood
Data: dados
      AIC      BIC    logLik
247.1762 260.202 -118.5881

Random effects:
Formula: ~1 | VOL
      (Intercept) Residual
StdDev:  0.4098028 0.7383248

Fixed effects: IL1 ~ HOR + MES
              Value Std.Error DF   t-value p-value
(Intercept) 2.9669497 0.19120970 88 15.516732 0.0000
HOR          0.0285366 0.03481928 88  0.819562 0.4147
MES         -0.4549896 0.15024043 88 -3.028410 0.0032
Correlation:
```

```
(Intr) HOR
HOR -0.461
MES -0.386 -0.032
```

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-2.01482814	-0.61662072	-0.04823236	0.67230265	4.05058330

Number of Observations: 100

Number of Groups: 10

```
> mod.II<-lme(IL1~HOR+MES, random=~(0+HOR)|VOL,data=dados,method="ML")
```

```
> summary(mod.II)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

AIC	BIC	logLik
248.9941	262.02	-119.4971

Random effects:

Formula: $\sim(0 + \text{HOR}) | \text{VOL}$
HOR Residual

StdDev: 0.1159498 0.7483546

Fixed effects: IL1 ~ HOR + MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9760757	0.14073406	88	21.146805	0.0000
HOR	0.0286728	0.05129881	88	0.558936	0.5776
MES	-0.4735782	0.15271199	88	-3.101120	0.0026

Correlation:

```
(Intr) HOR
HOR -0.436
MES -0.533 -0.022
```

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-1.97242816	-0.57124990	-0.05553288	0.57397476	3.92542792

Number of Observations: 100

Number of Groups: 10

```
> mod.III<-lme(IL1~HOR+MES+HOR*MES, random=~(0+HOR)|VOL,
```

```
data=dados,method="ML")
```

```
> summary(mod.III)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

AIC	BIC	logLik
250.9799	266.611	-119.49

Random effects:

Formula: $\sim(0 + \text{HOR}) \mid \text{VOL}$
 HOR Residual
StdDev: 0.1156857 0.7484321

Fixed effects: IL1 $\sim \text{HOR} + \text{MES} + \text{HOR} * \text{MES}$

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9870840	0.17007180	87	17.563664	0.0000
HOR	0.0242653	0.06383468	87	0.380128	0.7048
MES	-0.4950724	0.23986506	87	-2.063962	0.0420
HOR:MES	0.0083920	0.07180578	87	0.116870	0.9072

Correlation:

	(Intr) HOR	MES
HOR	-0.621	
MES	-0.710	0.443
HOR:MES	0.555	-0.591 -0.768

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.98693236	-0.55718021	-0.05347594	0.56276829	3.93373811

Number of Observations: 100

Number of Groups: 10

```
> mod.IV<-lme(IL1~HOR+MES+HOR*MES, random=~1|VOL,data=dados,method="ML")
```

```
> summary(mod.IV)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

	AIC	BIC	logLik
	249.0742	264.7052	-118.5371

Random effects:

Formula: $\sim 1 \mid \text{VOL}$
 (Intercept) Residual
StdDev: 0.4088625 0.7380558

Fixed effects: IL1 $\sim \text{HOR} + \text{MES} + \text{HOR} * \text{MES}$

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9960842	0.21333086	87	14.044308	0.0000
HOR	0.0169794	0.05085725	87	0.333864	0.7393
MES	-0.5118980	0.23630826	87	-2.166230	0.0330
HOR:MES	0.0220127	0.07030234	87	0.313114	0.7549

Correlation:

	(Intr) HOR	MES
HOR	-0.602	
MES	-0.558	0.544
HOR:MES	0.436	-0.726 -0.769

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-2.05577871	-0.61813915	-0.04123109	0.66776498	4.07474520

Number of Observations: 100

Number of Groups: 10

```
> mod.V<-lme(IL1~HOR+MES+HOR*MES, random=~(1+HOR)|VOL,data=dados,method="ML")
```

```
> summary(mod.V)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

	AIC	BIC	logLik
	250.9799	266.611	-119.49

Random effects:

Formula: $\sim(0 + \text{HOR}) | \text{VOL}$

	HOR	Residual
StdDev:	0.1156857	0.7484321

Fixed effects: IL1 ~ HOR + MES + HOR * MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9870840	0.17007180	87	17.563664	0.0000
HOR	0.0242653	0.06383468	87	0.380128	0.7048
MES	-0.4950724	0.23986506	87	-2.063962	0.0420
HOR:MES	0.0083920	0.07180578	87	0.116870	0.9072

Correlation:

	(Intr)	HOR	MES
HOR	-0.621		
MES	-0.710	0.443	
HOR:MES	0.555	-0.591	-0.768

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.98693236	-0.55718021	-0.05347594	0.56276829	3.93373811

Number of Observations: 100

Number of Groups: 10

```
> mod.VI<-lme(IL1~MES, random=~1|VOL,data=dados,method="ML")
```

```
> summary(mod.VI)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

	AIC	BIC	logLik
	245.866	256.2867	-118.933

Random effects:

Formula: $\sim 1 | \text{VOL}$

	(Intercept)	Residual
StdDev:	0.4092373	0.7411675

```

Fixed effects: IL1 ~ MES
              Value Std.Error DF   t-value p-value
(Intercept)  3.0391378 0.1689519 89 17.988179  0.0000
MES          -0.4510544 0.1499719 89 -3.007593  0.0034
Correlation:
  (Intr)
MES -0.453

```

```

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.10446667 -0.65811699 -0.07194745  0.66616287  4.09240900

```

```

Number of Observations: 100
Number of Groups: 10

```

```

> mod.VII<-lme(IL1~MES, random=~(0+MES)|VOL,
data=dados,method="ML")
> summary(mod.VII)
Linear mixed-effects model fit by maximum likelihood
Data: dados
      AIC      BIC    logLik
258.4832 268.9039 -125.2416

```

```

Random effects:
Formula: ~(0 + MES) | VOL
          MES Residual
StdDev: 0.0002510451 0.8466037

```

```

Fixed effects: IL1 ~ MES
              Value Std.Error DF   t-value p-value
(Intercept)  3.0308163 0.1221713 89 24.807931  0.0000
MES          -0.4347379 0.1710740 89 -2.541227  0.0128
Correlation:
  (Intr)
MES -0.714

```

```

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.0444233 -0.6447129 -0.1340031  0.5066378  3.8615278

```

```

Number of Observations: 100
Number of Groups: 10

```

```

> mod.VIII<-lme(IL1~MES, random=~(1+MES)|VOL,data=dados,method="ML")
> summary(mod.VIII)
Linear mixed-effects model fit by maximum likelihood
Data: dados

```

	AIC	BIC	logLik
	230.3403	245.9713	-109.1701

Random effects:

Formula: $\sim(1 + \text{MES}) \mid \text{VOL}$
 Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.7635072	(Intr)
MES	0.7315552	-0.949
Residual	0.6324411	

Fixed effects: IL1 \sim MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	3.0432030	0.2604715	89	11.683440	0.0000
MES	-0.4515758	0.2664347	89	-1.694883	0.0936

Correlation:

(Intr)	
MES	-0.9

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.8923270	-0.6256519	-0.1822294	0.5474607	3.4437932

Number of Observations: 100

Number of Groups: 10

```
> mod.IX<-lme(IL1~MES+HOR*MES, random=~(1+MES)|VOL,data=dados,method="ML")
```

```
> summary(mod.IX)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

	AIC	BIC	logLik
	233.0579	253.8992	-108.5289

Random effects:

Formula: $\sim(1 + \text{MES}) \mid \text{VOL}$
 Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.7657221	(Intr)
MES	0.7342739	-0.95
Residual	0.6279282	

Fixed effects: IL1 \sim MES + HOR * MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9918667	0.28536372	87	10.484398	0.0000
MES	-0.5108365	0.31078204	87	-1.643713	0.1038
HOR	0.0203936	0.04341183	87	0.469770	0.6397
MES:HOR	0.0211533	0.05982805	87	0.353569	0.7245

Correlation:

(Intr)	MES	HOR
--------	-----	-----

```
MES      -0.857
HOR      -0.383  0.352
MES:HOR  0.278 -0.498 -0.726
```

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.8198286	-0.5951425	-0.1628044	0.6206568	3.4170320

Number of Observations: 100

Number of Groups: 10

```
> mod.X<-lme(IL1~HOR+MES, random=~(1+MES)|VOL,data=dados,method="ML")
```

```
> summary(MOD.X)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

	AIC	BIC	logLik
	231.1879	249.4241	-108.594

Random effects:

Formula: $\sim(1 + \text{MES}) | \text{VOL}$

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.7665095	(Intr)
MES	0.7342710	-0.95
Residual	0.6282344	

Fixed effects: IL1 \sim HOR + MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9638424	0.27295678	88	10.858285	0.0000
HOR	0.0315315	0.02971725	88	1.061049	0.2916
MES	-0.4561743	0.26819572	88	-1.700901	0.0925

Correlation:

	(Intr)	HOR
HOR	-0.274	
MES	-0.862	-0.016

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.7733923	-0.5898820	-0.1638288	0.6256802	3.3909737

Number of Observations: 100

Number of Groups: 10

Seguidamente fizemos a comparação dos modelos, a título de exemplo:

```
> anova(mod.I,mod.II)
```

	Model	df	AIC	BIC	logLik
mod.I	1	5	247.1762	260.202	-118.5881
mod.II	2	5	248.9941	262.020	-119.4971

```

> anova(mod.I,mod.III)
      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
mod.I      1  5 247.1762 260.202 -118.5881
mod.III    2  6 250.9799 266.611 -119.4900 1 vs 2 1.803742 0.1793

> anova(mod.I,mod.IV)
      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
mod.I      1  5 247.1762 260.2020 -118.5881
mod.IV     2  6 249.0742 264.7052 -118.5371 1 vs 2 0.1020352 0.7494

> anova(mod.I,mod.VII)
      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
mod.I      1  5 247.1762 260.2020 -118.5881
mod.VII    2  4 258.4832 268.9039 -125.2416 1 vs 2 13.307 3e-04

> anova(mod.VII,mod.IX)
      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
mod.VII    1  4 258.4832 268.9039 -125.2416
mod.IX     2  8 233.0579 253.8992 -108.5289 1 vs 2 33.42533 <.0001

```

Apêndice 2: *output mod.XI a mod.XIII*

```

> mod.XI<-lme(IL1~TRI+MES+HOR*MES, random=~(1+MES)|VOL,data=dados,method="ML")
> summary(mod.XI)
Linear mixed-effects model fit by maximum likelihood
Data: dados
      AIC      BIC    logLik
233.0079 256.4544 -107.5039

Random effects:
Formula: ~(1 + MES) | VOL
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev  Corr
(Intercept) 0.7672353 (Intr)
MES          0.7102959 -0.972
Residual    0.6277587

Fixed effects: IL1 ~ TRI + MES + HOR * MES
      Value Std.Error DF  t-value p-value
(Intercept) 2.7620004 0.3269344 86  8.448179 0.0000
TRI          0.0019282 0.0013104 86  1.471512 0.1448
MES         -0.5302166 0.3067492 86 -1.728502 0.0875
HOR          0.0052204 0.0448523 86  0.116392 0.9076
MES:HOR     0.0250092 0.0601607 86  0.415706 0.6787

Correlation:
      (Intr) TRI    MES    HOR
TRI    -0.477
MES    -0.741 -0.042
HOR    -0.216 -0.232 0.358

```

MES:HOR 0.226 0.038 -0.508 -0.715

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-1.8125421	-0.6580391	-0.1689380	0.6048787	3.3520822

Number of Observations: 100

Number of Groups: 10

```
> mod.XII<-lme(IL1~MES+HOR*MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
> summary(mod.XII)
Linear mixed-effects model fit by maximum likelihood
```

```
Data: dados
      AIC      BIC  logLik
234.3419 260.3936 -107.171
```

Random effects:

```
Formula: ~(1 + MES) | VOL
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev  Corr
(Intercept) 0.7653385 (Intr)
MES          0.7011453 -0.974
Residual    0.6263850
```

Fixed effects: IL1 ~ MES + HOR * MES + TRI * MES

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.9302802	0.3916121	85	7.482609	0.0000
MES	-0.7849047	0.4426268	85	-1.773288	0.0798
HOR	0.0164863	0.0472008	85	0.349281	0.7277
TRI	0.0005151	0.0022268	85	0.231324	0.8176
MES:HOR	0.0095770	0.0634026	85	0.151051	0.8803
MES:TRI	0.0020838	0.0026198	85	0.795406	0.4286

Correlation:

	(Intr) MES	HOR	TRI	MES:HO
MES	-0.824			
HOR	-0.007	0.018		
TRI	-0.677	0.568	-0.375	
MES:HOR	0.012	-0.114	-0.741	0.269
MES:TRI	0.548	-0.723	0.303	-0.809

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-1.8305583	-0.6082204	-0.1387647	0.5653208	3.4187707

Number of Observations: 100

Number of Groups: 10

```
> anova (mod.IX,mod.X)
      Model df      AIC      BIC  logLik  Test  L.Ratio p-value
```

```

mod.IX      1  8 233.0579 253.8992 -108.5289
mod.X       2  9 233.0079 256.4544 -107.5040 1 vs 2 2.049963 0.1522

```

```
> anova(mod.IX,mod.XII)
```

```

      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
mod.IX     1  8 233.0579 253.8992 -108.5289
mod.XI     2 10 234.3419 260.3936 -107.1710 1 vs 2 2.715948 0.2572

```

```
> mod.XIII<-lme(IL1~TRI+MES+HOR*MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
```

```
> summary(mod.XIII)
```

Linear mixed-effects model fit by maximum likelihood

Data: dados

```

      AIC      BIC    logLik
234.3419 260.3936 -107.171

```

Random effects:

Formula: ~(1 + MES) | VOL

Structure: General positive-definite, Log-Cholesky parametrization

```

      StdDev  Corr
(Intercept) 0.7653389 (Intr)
MES          0.7011450 -0.974
Residual    0.6263851

```

Fixed effects: IL1 ~ TRI + MES + HOR * MES + TRI * MES

```

      Value Std.Error DF  t-value p-value
(Intercept) 2.9302802 0.3916122 85  7.482606 0.0000
TRI          0.0005151 0.0022268 85  0.231324 0.8176
MES         -0.7849048 0.4426266 85 -1.773289 0.0798
HOR          0.0164863 0.0472008 85  0.349281 0.7277
MES:HOR      0.0095770 0.0634026 85  0.151051 0.8803
TRI:MES      0.0020838 0.0026198 85  0.795406 0.4286

```

Correlation:

```

      (Intr) TRI    MES    HOR    MES:HO
TRI      -0.677
MES      -0.824  0.568
HOR      -0.007 -0.375  0.018
MES:HOR  0.012  0.269 -0.114 -0.741
TRI:MES  0.548 -0.809 -0.723  0.303 -0.307

```

Standardized Within-Group Residuals:

```

      Min      Q1      Med      Q3      Max
-1.8305585 -0.6082198 -0.1387642  0.5653204  3.4187709

```

Number of Observations: 100

Number of Groups: 10

```
> anova(mod.IX,mod.XIII)
```

```

      Model df      AIC      BIC    logLik  Test  L.Ratio p-value

```



```

mod.IX      1  8 233.0579 253.8992 -108.5289
mod.XII     2 10 234.3419 260.3936 -107.1710 1 vs 2 2.715948 0.2572

```

Apêndice 3: *output mod.XIV*

```

> mod.XIV<-lme(IL1~TRI+MES+TRI*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
> summary(mod.XIV)
Linear mixed-effects model fit by maximum likelihood
Data: dados
      AIC      BIC    logLik
158.793 175.5477 -71.39649

Random effects:
Formula: ~(1 + MES) | VOL
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev   Corr
(Intercept) 0.7960699 (Intr)
MES          0.7585741 -0.977
Residual    0.6876130

Fixed effects: IL1 ~ TRI + MES + TRI * MES
              Value Std.Error DF   t-value p-value
(Intercept)  2.9178294 0.4450804 47   6.555736 0.0000
TRI          0.0009612 0.0024339 47   0.394940 0.6947
MES         -0.6838579 0.5325508 47  -1.284118 0.2054
TRI:MES      0.0015690 0.0031383 47   0.499943 0.6194
Correlation:
      (Intr) TRI    MES
TRI    -0.754
MES    -0.803  0.613
TRI:MES 0.569 -0.755 -0.814

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-1.5683574 -0.6068216 -0.1321617  0.5143227  2.8929004

Number of Observations: 60
Number of Groups: 10

```

Apêndice 4: *output mod.XV*

```

> mod.XV<-lme(IL1~LDLox+MES+LDLox*MES,
random=~(1+MES)|VOL,data=dados,method="ML")
> summary(mod.XV)
Linear mixed-effects model fit by maximum likelihood
Data: dados
      AIC      BIC    logLik
154.8474 171.6021 -69.42368

```

Random effects:

Formula: $\sim(1 + \text{MES}) \mid \text{VOL}$

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.8045953	(Intr)
MES	0.7835884	-0.961
Residual	0.6524319	

Fixed effects: $\text{IL1} \sim \text{LDLox} + \text{MES} + \text{LDLox} * \text{MES}$

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.8089192	0.6414311	47	4.379144	0.0001
LDLox	0.0045593	0.0107578	47	0.423813	0.6736
MES	-1.6825549	0.9000930	47	-1.869312	0.0678
LDLox:MES	0.0329006	0.0193174	47	1.703156	0.0951

Correlation:

	(Intr)	LDLox	MES
LDLox	-0.890		
MES	-0.702	0.631	
LDLox:MES	0.492	-0.552	-0.928

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.5757689	-0.7189021	-0.1066649	0.6033227	2.9710440

Number of Observations: 60

Number of Groups: 10

Capítulo 7

Considerações Finais

No final desta análise parece promissor explorar a possibilidade de uma dieta rica em framboesas, de alto poder oxidante, ter efeitos benéficos sobre o colesterol, nomeadamente

1. diminuir o LDLox após o tratamento, e estabilizar os valores durante o período pós pequeno almoço, no sentido em que há menor dispersão;
2. aumentar o “bom” colesterol HDL entre os dois períodos, e diminuir a dispersão ao longo da digestão;
3. diminuir e estabilizar os mediadores de infeção Interleucina-1 e Interleucina-6; no que se refere a IL-1 observa-se também menor dispersão ao longo das 6 horas de observação em Junho, e no que se refere a IL-6 nota-se ainda que após 4 horas o valor deste indicador de inflamação diminui em Junho, ao contrário do que acontecia em Maio.

Assim, o tratamento com framboesas parece eficaz na redução de infeção, por aumentar o bom colesterol e diminuir o LDLox.

Mas os dados foram obtidos numa coleção de 10 voluntários, e por isso não se pode concluir que aquelas possibilidades são válidas para a população. Apesar de não ter sido rejeitada a hipótese de aleatoriedade daquela coleção de dados, isso pode dever-se apenas ao muito pequeno número de voluntários usados. Por outro lado não parece ter havido o uso de critérios de exclusão, algumas das medições indicam uma condição clínica preocupante em voluntários que podem estar a tomar medicamentos causadores de confundimento, não tinha havido “limpeza” dos dados (foi necessário alterar um dado 982 dividindo-o por 10), e o protocolo de obtenção dos dados pode ter tido deficiências — por exemplo no que se refere à medição de IL há dados com um decimal e dados com 2 decimais.

O pequeno tamanho da amostra pode também ser insuficiente para detetar alterações significativas, recorde-se o estudo feito sobre essa questão no capítulo de análise de dados longitudinais (e as dúvidas sobre se a manutenção da hipótese nula de normalidade não será apenas a consequência de o tamanho das amostras ser insuficiente para se conseguir rejeição).

Mas não é apenas a deficiente amostragem que está em causa. Há um outro erro fatal, no planeamento da experiência: para se conseguir tirar efetivamente conclusões sobre o efeito de framboesas nas variáveis recolhidas, seria necessário um planeamento que permitisse comparar de facto os efeitos do tratamento com framboesas, por exemplo

- fazer medições em Maio, haver o tratamento até ao fim de Maio, fazer novas medições no início de Junho, estarem até ao fim de Junho com dieta sem framboesas, e tornar a fazer medições em Julho; ou
- constituir dois grupos equilibrados, um dos quais recebia o tratamento e outro não (controle).

Qualquer que fosse o planeamento adequado, haveria ainda que ter o cuidado de estabelecer com rigor critérios de inclusão, nomeadamente para evitar confundimento. De facto, se os pacientes tomarem medicamentos que possam interferir com a metabolização de antioxidantes, isso pode comprometer irremediavelmente a qualidade do estudo.

Observando-se no entanto que o tratamento teve nesta coleção de voluntários um efeito benéfico — diminuição dos indicadores de inflamação IL-1 e IL-6 e das lipoproteínas oxidadas LDLox, e aumento das boas lipoproteínas

HDL — parece interessante prosseguir o estudo. Deve porém usar-se uma amostra aleatória de tamanho adequado, controlando tanto quanto possível para evitar confundimento, com um protocolo mais rígido, e com um delineamento experimental para investigar questões bem formuladas, nomeadamente distinguindo variáveis e covariáveis. Haverá evidentemente que ter em conta efeitos secundários do consumo de doses tão grandes de framboesas, nomeadamente a nível digestivo, e que avaliar se é a forma mais eficaz e barata de administrar antioxidantes aos doentes.

Bibliografia

Cabral, M. S., e Gonçalves, M. H. (2011). *Análise de Dados Longitudinais*, Sociedade Portuguesa de Estatística, Lisboa.

Mosteller, F., and Rourke, R. E. K. (1993). *Estatísticas Firmes*, Salamandra, Lisboa.

Pinheiro, J., and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*, Springer Verlag. New York.

Ramsey, F. L., and Schafer, D. W. (2002). *The Statistical Sleuth — A Course in Methods of Data Analysis*, 2nd ed., Duxbury / Thomson Learning.