



Análise Estatística para identificação de factores discriminantes de progressão da doença de Parkinson

Ana Rita Ribeiro Sousa

Dissertação para obtenção do Grau de Mestre em
Matemática e Aplicações

Júri

Presidente: Prof. Doutor António Manuel Pacheco Pires
Orientadores: Prof. Doutor João António Branco
Prof. Doutora Ana Maria Nobre Vilhena Nunes Pires de Melo Parente
Vogal: Prof. Doutor Joaquim José Coutinho Ferreira
Vogal: Prof. Doutora Maria da Conceição Esperança Amado

Dezembro de 2011

Resumo

A doença de Parkinson é a doença neurodegenerativa que mais afecta, a seguir à doença de Alzheimer, a população mundial, especialmente a mais idosa. Esta doença afecta o sistema motor (tremor, rigidez, lentidão de movimentos, instabilidade postural), não se conhecendo até à data nenhuma causa concreta do seu aparecimento. Sabe-se no entanto que existem doentes em que a doença progride mais rapidamente e outros em que a doença progride mais lentamente.

Neste trabalho pretendem-se encontrar as principais características que permitem distinguir estes dois grupos, usando um modelo de Regressão Logística. Dado o elevado número de variáveis em estudo é feita previamente uma selecção de variáveis, usando o método de Análise de Componentes Principais.

Uma outra questão importante suscitada por este estudo, é a forma como se codificam certas variáveis associadas à avaliação de inquéritos, nomeadamente quando há que transformar em variáveis numéricas as variáveis avaliadas de acordo com uma escala de Likert. Muitas vezes a característica que se pretende estudar nesses inquéritos depende do tipo de perguntas que são feitas e do próprio sujeito respondente. Pelo que neste trabalho também se apresenta uma descrição matemática da relação entre a característica que se pretende estudar e a apetência do sujeito, recorrendo ao modelo Rasch.

Palavras Chave: Doença de Parkinson, Progressão, Modelos Rasch, Regressão Logística, Escala de Likert, Análise de Componentes Principais.

Abstract

The Parkinson's disease is a neurodegenerative disease and after Alzheimer's disease, is the most common disease among the world population, specially the elderly. This disease affects the motor system (tremor, rigidity, slowness of movements, postural instability), however no precise cause about its origin is known at the moment. Nevertheless it is known that two groups of patients can be identified: one first group showing fast progression of the disease and a second group where the progression is slow.

In this dissertation we intend to find the main features that enable the discrimination between these groups, using a Logistic Regression Model. There are a huge number of variables, so a previous variables selection, using a Principal Components Analysis method, is done.

Another important question brought by this study is about the way some variables, associated with surveys evaluation, are transformed from a Likert Scale to numerical. Sometimes, the objective characteristic depends on the type of the questions as well as on the individuals who answer them. This work also describes a mathematical relation between the characteristic under study and the ability of the individual, using a Rasch Model.

Key words: Parkinson's Disease, Progression, Rasch Models, Logistic Regression, Likert Scale, Principal Components Analysis.

Agradecimentos

Até chegar aqui tive que passar longos e difíceis anos a trabalhar arduamente para concluir todas as cadeiras do curso e esta dissertação, e este caminho, só se tornou mais fácil graças a algumas pessoas que fui encontrando e outras que sempre estiveram comigo e me ajudaram a ser mais e melhor.

Aos meus pais por nunca terem deixado de acreditar em mim mesmo quando eu própria tinha dúvidas, por me terem dado um dos bens mais preciosos, os meus valores e a minha educação. E por lutarem todos os dias com muita dificuldade para que hoje eu possa estar onde estou.

Aos meus avós e restante família, por terem ajudado os meus pais na minha formação e pelo enorme coração que têm comigo e com os outros.

À minha querida mana Daniela, por ser mais que uma irmã, ser uma amiga e uma companheira, e por ser a melhor das cúmplices. Tenho muito orgulho em ti!

Ao Diogo, por estar comigo todos os dias, por me apoiar incondicionalmente, por me ajudar e encorajar nos momentos em que duvidei de mim própria durante este caminho e por ser o meu porto de abrigo. Contigo sou uma pessoa melhor, minha vida é muito mais feliz, e ao teu lado consigo ver um futuro melhor e cheio de coisas boas. "*Du bist meine Traummann*"

Aos meus orientadores, Professora Ana Pires e Professor João Branco, pelo tempo que dispenderam a ajudar-me, pela motivação e pelas orientações que me deram no desenvolvimento desta dissertação.

Aos membros do Centro de Estudos Egas Moniz, em especial ao Professor Doutor Joaquim Ferreira e à Dra. Leonor Guedes, por acreditarem em mim, pela disponibilidade, e pela ajuda na interpretação biológica necessária neste trabalho.

Ao BEST, e a todos os que conheci graças a esta organização, sem vocês esta incursão universitária não teria sido tão rica e maravilhosa. Obrigada a todos, pelo que aprendi e pelo que desenvolvi.

Finalmente, um enorme obrigada aos meus amigos, aqueles que estiveram sempre comigo, ou que apareceram mais tarde, mas que de uma forma ou de outra me marcaram, por me ajudarem e por partilharem comigo os momentos bons e menos bons. Em especial, Ana, Alberto, Patrício e João, muito obrigada.

Ana Rita Sousa

ÍNDICE

Resumo	i
Abstract	iii
Agradecimentos	v
Lista de figuras	viii
Lista de tabelas	xi
1 Introdução	1
2 Conceitos Introdutórios	5
2.1 Sintomas: Noções Básicas	5
2.2 Escalas: Noções Básicas	6
2.3 Outras noções importantes	8
3 Análise Preliminar de dados	11
3.1 Descrição das variáveis em estudo (204)	11
3.2 Análise Descritiva	32

4	Métodos	45
4.1	Análise de Componentes Principais	45
	Exemplo	46
4.2	Escala de Likert	47
	Exemplo	48
4.3	Modelo Rasch	48
4.3.1	Modelo de Rasch politómico	50
	Exemplo	50
4.4	Análise Discriminante	51
4.4.1	Regressão Logística Múltipla	52
	Exemplo	52
5	Apresentação de Resultados e Conclusões	55
5.1	Seleccção de variáveis	55
5.2	Modelo de Rasch politómico	56
5.3	Seleccção do modelo	57
5.4	Análise de componentes principais	58
5.5	Regressão logística múltipla	59
5.6	Conclusões	63
A	Macros em R	65
A.1	Modelos Rasch	65
A.1.1	Resultados para a UPDRS1 (sintomas não motores)	65
A.1.2	Resultados para a UPDRS2 (actividades da vida diária)	68
A.1.3	Resultados para a UPDRS3 (motilidade)	71
A.1.4	Resultados para a UPDRS4 (complicações do tratamento)	75
A.1.5	Resultados para a Escala de London	77
	Bibliografia	81

LISTA DE FIGURAS

3.1	Diagramas de caixa da idade do doente em relação à progressão	33
3.2	Histogramas da idade dos primeiros sintomas e da idade de diagnóstico . . .	33
3.3	Gráfico de barras do estado de demência do paciente em relação à progressão	36
3.4	Diagramas de caixa da duração dos vícios	36
3.5	Diagramas de caixa das características físicas para cada um dos grupos de progressão	37
3.6	Gráfico de barras do estadio HY em relação à progressão	38
3.7	Gráfico de barras da escala de Schwab e England em relação à progressão . .	39
3.8	Diagrama de caixa para a pontuação MMSE em cada um dos grupos de progressão	40
3.9	Progressão vs Actividade física regular	41
3.10	Diagramas de caixa para o total da Escala de Beck em relação à progressão da doença	42
3.11	Gráficos de barras dos itens da escala de London em relação aos grupos de progressão	43
4.1	Escala de valor da saúde	48

LISTA DE TABELAS

3.1	Sumário da idade em que os doentes tiveram os primeiros sintomas da doença	34
3.2	Sumário da idade em que foi feito o diagnóstico	34
3.3	Progressão vs Tremor	34
3.4	Progressão vs Rigidez	35
3.5	Progressão vs Bradicinesia	35
3.6	Progressão vs Instabilidade postural	35
3.7	Progressão vs Cirurgia (DBS)	38
3.8	Sumário da pontuação obtida no MMSE	39
3.9	Progressão vs Estado Civil	40
3.10	Progressão vs Dislipidemia	41
3.11	Progressão vs Ácido úrico elevado	42
4.1	Descrição das variáveis	47
4.2	Tabela das estimativas dos τ_{is}	51
4.3	Tabela das estimativas dos θ_v	51
4.4	Descrição das variáveis	53
4.5	Estimativas dos parâmetros	53
4.6	Matriz de Classificação	54

5.1	Tabelas das componentes principais de cada um dos subgrupos (ESC, NF, SINT, F, RESTO)	59
5.2	Frequências da progressão (variável resposta)	60
5.3	Estimativas dos parâmetros, estimativas dos erros padrão e valor-p para o Modelo 1	60
5.4	Estimativas dos parâmetros, estimativas dos erros padrão e valor-p para o Modelo 2	61
5.5	Estimativas de erro de predição para taxa aparente e validação cruzada . . .	61
5.6	Estimativas dos parâmetros associados às variáveis originais	62
5.7	Matriz de classificação para taxa aparente	62
5.8	Matriz de classificação para validação cruzada	63

CAPÍTULO 1

Introdução

A doença de Parkinson é uma doença degenerativa do sistema nervoso central de elevada prevalência, afectando entre 45 a 300 pessoas por 100.000 habitantes em todo mundo, e 2% da população mundial com mais de 65 anos (de Rijk *et al.* (2000)). Em Portugal, estima-se que existam cerca de 20.000 casos de doença de Parkinson, valor somente ultrapassado pela doença de Alzheimer, no domínio das doenças neurodegenerativas. Pelo conhecido envelhecimento da população ocidental, e por esta doença ter uma idade de início médio entre os 55 e 65 anos, espera-se que estes valores aumentem num futuro próximo.

Esta doença foi descrita pela primeira vez por James Parkinson em 1817, e desde então não se descobriu a sua causa. E apesar de em grupos restritos terem sido identificados vários factores genéticos e ambientais associados à doença, até à data não se encontrou um factor único ou um conjunto de factores que explique o seu aparecimento e evolução na totalidade dos doentes.

Dado o envelhecimento da população e o conseqüente aumento da prevalência da doença de Parkinson torna-se cada vez mais importante identificar as suas causas e aprofundar os conhecimentos ligados ao mecanismo da doença para que se possam encontrar soluções terapêuticas mais adequadas e desenvolver novos medicamentos.

No sentido de compreender melhor a evolução clínica da doença e de identificar possíveis novos mecanismos da mesma, foi desenvolvido o projecto “Doença de Parkinson de progressão rápida e lenta: determinantes moleculares e biomarcadores”, pelo Instituto de Medicina Molecular (Prof. Doutora Cristina Sampaio, Prof. Doutor Joaquim Ferreira, Dra. Leonor

C. Guedes, Dr. Mário Rosa, Dr. Miguel Coelho, Dr. Tiago Mestre e Doutor Tiago Outeiro), cujo objectivo é identificar factores epidemiológicos, clínicos e moleculares associados à evolução da doença de Parkinson que permitam distinguir um percurso mais rápido ou mais lento da doença.

A recolha de informação consistiu na avaliação clínica realizada por um médico, num questionário demográfico respondido pelo paciente ou pelo cuidador, se este estivesse incapacitado, e na análise clínica de uma amostra de sangue do paciente.

Em termos de elegibilidade, os critérios de inclusão no estudo foram: doentes com diagnóstico clínico de doença de Parkinson de acordo com o United Kingdom Brain Bank Criteria (UKBBC), com sintomas da doença há pelo menos 10 anos e que assinem o consentimento informado. Os critérios de exclusão foram: doentes com quadro clínico de parkinsonismo atípico.

Presentemente, a doença de Parkinson tem apenas tratamento sintomático, sendo o mais eficaz a levodopa. Dado não existirem formas de terapêutica curativa ou mesmo modeladora da doença, o estudo da história natural da mesma, da sua forma de evolução, e a identificação de subgrupos de doentes com velocidades de progressão distintos assume uma importância significativa pela possibilidade de se identificarem novos mecanismos de doença ou alvos de modelação que permitam a identificação de terapêuticas inovadoras.

No âmbito deste projecto desenvolveu-se a parceria entre a Unidade Neurológica de Investigação clínica do Instituto de Medicina Molecular (IMM) e a Secção de Probabilidades e Estatística do Instituto Superior Técnico com vista a pôr em prática métodos estatísticos adequados para estudar os factores clínicos associados à doença de Parkinson, ficando os factores genéticos e identificação de biomarcadores a cargo de outras entidades. Assim, nesta dissertação foram utilizados métodos e técnicas estatísticas, em particular a Análise Discriminante, para identificar factores epidemiológicos e clínicos que permitam fazer uma distinção entre os dois grupos de progressão da doença.

Dado o número elevado de variáveis foram utilizados diferentes métodos no intuito de se poder fazer uma selecção das mesmas. Mais concretamente foi efectuada uma análise preliminar de dados para identificar redundâncias, dados omissos e pouca variabilidade, e uma análise de componentes principais.

Neste trabalho procura-se ainda identificar formas para obter uma melhor modelação dos valores codificados nas diferentes escalas de medição consideradas nos questionários.

O protocolo, o consentimento, o questionário de diagnóstico médico e o questionário dirigido ao doente encontram-se reunidos em CD anexo.

No capítulo seguinte faz-se uma breve descrição de alguns termos médicos que serão utilizados ao longo do trabalho, assim como uma explicação das diferentes escalas usadas, para depois se passar, no capítulo 3, à análise exploratória dos dados, que será seguida da exposição de métodos estatísticos eventualmente mais esclarecedores da informação contida nos dados, no capítulo 4. No capítulo 5 são apresentados os resultados e conclusões deste trabalho.

CAPÍTULO 2

Conceitos Introdutórios

Neste capítulo apresentam-se algumas noções importantes de conceitos médicos e biológicos, necessários para a melhor compreensão do estudo, assim como a explicação das escalas de diagnóstico utilizadas.

2.1 Sintomas: Noções Básicas

Existem vários sinais e sintomas, ou queixas, referidas pelos doentes, associados à doença de Parkinson e como tal é importante compreender o seu significado. Nesta secção descreve-se o significado de sinais e sintomas avaliados no contexto deste projecto, com base em Kumar and Calne (2004).

- **Bradicinésia**

Lentificação e diminuição da frequência de realização de movimentos, que se pode expressar, por exemplo, na diminuição da mímica da face, ou na diminuição da velocidade da marcha.

- **Discinesia**

Movimento involuntário, espontâneo, anormal, não controlado por vontade do doente.

- **Distonia**

Movimento involuntário, anormal, caracterizado por contracções musculares mantidas, causando torções de partes do corpo, movimentos repetitivos ou posturas anormais.

- ***Freezing***

Bloqueio súbito de uma tarefa motora, transitório, ultrapassado após outras tentativas de concretização da tarefa.

- ***Sleep benefit motor***

Fenómeno em que após um período de sono os sintomas da doença melhoram.

- **Disfagia**

Dificuldade de deglutição.

- **Disartria**

Alteração na articulação das palavras.

- ***Punding***

Alteração do pensamento que leva o doente a realizar continuamente tarefas repetitivas, não necessárias.

- **Dislipidémia**

Presença de níveis alterados de lípidos no sangue (associado aos valores de colesterol e triglicéridos).

2.2 Escalas: Noções Básicas

No sentido de tornar mais objectivas e mensuráveis as observações clínicas dos doentes utilizaram-se diversas escalas clínicas neste estudo, cuja descrição é apresentada nesta secção.

- **Escala UPDRS**

A escala UPDRS (Unified Parkinson Disease's Rating Scale) visa fazer uma avaliação global detalhada do estado clínico do doente sobre quatro vertentes: (1) Actividade mental, comportamento e humor; (2) Actividades da vida diária; (3) Motilidade ¹ e (4) Complicações do tratamento. Cada vertente tem incluídos diversos itens, cada um dos quais apresentando uma gradação de 0 a 4, sendo 0 indicador de normalidade e 4 de maior gravidade (Virgilio (2003)).

- **Escala de Hoehn e Yahr**

Esta é uma escala de avaliação do estadio da doença em que se encontra o paciente, que varia entre 0 e 5. Sendo 0 correspondente à normalidade e 5 correspondente a um doente totalmente acamado (Virgilio (2003)).

- **Escala Schwab and England**

Esta é uma escala de incapacidade para actividades da vida diária (Virgilio (2003)). A escala de Schwab and England varia da seguinte forma: 100% (completamente independente e capaz de efectuar as suas tarefas), 90% (completamente independente e capaz de desempenhar as suas tarefas, mas com alguma lentidão), 80% (completamente independente na maioria das tarefas), 70% (não é completamente independente e tem maior dificuldade com algumas tarefas), 60% (alguma dependência, consegue executar a maioria das tarefas, mas com muita lentidão e esforço), 50% (maior dependência, necessidade de ajuda em metade do que faz), 40% (muito dependente, requer ajuda em todas as tarefas), 30% (por vezes e com esforço, inicia ou executa alguma tarefa), 20% (não faz nada sozinho, podendo ajudar em algumas tarefas), 10% (totalmente dependente e incapaz), 0% (acamado, compromisso de funções vegetativas, como a deglutição, micção e defecção).

- **Inventário de Depressão de Beck (BDI)**

A escala de depressão de Beck tem por objectivo medir a intensidade da depressão. Esta escala é constituída por 13 tópicos sobre sentimentos relacionados com depressão (por exemplo, tristeza ou desinteresse pelas coisas/pessoas), em que cada um destes é constituído por quatro frases, alusivas ao tópico, e de grau crescente de gravidade (Beck *et al.* (1961)).

¹Motilidade=Força motora

- **Escala de London**

A escala de London visa avaliar a incapacidade do doente imposta pela doença, face a actividades diárias. Esta escala avalia esta incapacidade de acordo com os seguintes itens: mobilidade, independência física, trabalho e lazer, integração social, orientação e independência económica, e varia entre 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente), 6 (completamente).

2.3 Outras noções importantes

Nesta secção são apresentados os significados de outros termos técnicos.

- **Coefficiente de parentesco de *Sewall Wright***

Este coeficiente que pode ser usado para medir o grau de parentesco entre dois indivíduos, é definido como a probabilidade de que, dado um *locus* escolhido ao acaso, os alelos dos dois indivíduos sejam idênticos por origem (Wright (1932)). Assumindo que não existe consanguinidade entre os progenitores, obtêm-se os seguintes valores: 100% (entre gémeos verdadeiros), 50% (entre pais e filhos e entre irmãos), 25% (entre avós e netos, entre meios-irmãos e entre tios e sobrinhos), 12.5% (entre bisavós e bisnetos ou entre primos diretos), 3.125% (entre primos em segundo grau).

- **Consanguinidade**

Diz-se que existe consanguinidade entre os indivíduos que possuem um ascendente próximo comum.

- **DBS**

Deep Brain Stimulation (DBS), ou estimulação cerebral profunda, é uma técnica cirúrgica que foi proposta por Benabid *et al.* (1987), e que consiste na estimulação eléctrica de regiões profundas do cérebro de modo a alterar circuitos neuronais. Na doença de Parkinson a estimulação cerebral profunda, entre outros efeitos, reduz o tremor, a bradicinésia e a rigidez do doente.

- **MMSE**

Mini Mental State Examination (MMSE), ou avaliação breve do estado mental, é um teste de 30 pontos (Folstein *et al.* (1975)), que procura avaliar o estado cognitivo do doente, incidindo nos seguintes domínios: orientação, retenção, atenção e cálculo, evocação, linguagem e apraxia ². A pontuação é tanto maior quanto melhor for o estado mental do doente.

Compreendidos todos os conceitos utilizados ao longo do trabalho, procede-se no Capítulo 3 à descrição das variáveis e ao melhor entendimento do que representam, fazendo-se uma análise preliminar de dados.

²Apraxia é um defeito cognitivo em que o doente, apesar de não apresentar defeito motor, sensitivo ou sensorial, não consegue realizar tarefas por ter desaprendido a realizá-las.

CAPÍTULO 3

Análise Preliminar de dados

Neste capítulo é feita uma apresentação dos dados, começando-se pela descrição de todas as variáveis. E prosseguindo-se com uma análise exploratória de modo a resumir as principais características dos dados numa forma mais fácil de entender, recorrendo-se para isso a gráficos e tabelas.

3.1 Descrição das variáveis em estudo (204)

O conjunto de dados deste estudo é constituído por informação demográfica e clínica, resultante do estudo de sintomas motores e não motores referentes a 78 pacientes com doença de Parkinson, sendo aproximadamente 50% do sexo feminino e 50% do sexo masculino, com sintomas da doença há pelo menos 10 anos e com idades compreendidas entre os 51 e os 88 anos, residentes em Portugal Continental.

De seguida apresenta-se uma descrição sucinta das variáveis assim como a nomenclatura usada para as definir.

- **Ácido úrico elevado (HighUricAcid)** - Indica se o paciente tem ácido úrico elevado (1) ou não (0);
- **Actividade física regular (RegularFisicalActivity)** - Indica se o paciente praticou regularmente alguma actividade física (1) ou não (0);

- **Actividade laboral (WorkActivity)** - Indica em que nível da actividade laboral se encontra o paciente, assumindo os valores: 0 para desempregado, 1 para activo e 2 para reformado;
- **Alcoolismo (AlcoholicHabits)** - Indica se nunca teve hábitos etílicos (0), deixou de beber (1) ou bebe bebidas alcoólicas regularmente (2);
- **Antecedentes familiares de DP (FamilialPD)** - Indica se o doente tem ou não outros membros da família com doença de Parkinson, assumindo o valor 1 no caso afirmativo e o valor 0 no caso negativo ;
- **Ataques de pânico no último mês (PanicAttacksLastMonth)** - Indica se o doente teve ataques de pânico no último mês (relativamente à data de entrada no estudo), assumindo o valor 1 em caso afirmativo, o valor 0 em caso negativo, e o valor 2 no caso de resposta desconhecida;
- **Benefício motor (MotorSleepBenefitEver)** - Indica se o doente alguma vez teve benefícios motores depois de dormir (1) ou não (0) ou se a resposta é desconhecida (2);
- **Bradicinésia como primeiro sintoma(BradikynesiaFMS)** - Indica se o primeiro sintoma da doença foi bradicinésia (1) ou não (0) ou se a resposta é desconhecida (2);
- **Cigarros diários (NumberCigarettesDay)** - Indica o número de cigarros que o paciente fuma ou fumava por dia;
- **Cirurgia de Parkinson (DBS)** - Indica se o doente já foi sujeito à cirurgia de estimulação cerebral profunda (Deep Brain Stimulation), assumindo o valor 1 em caso afirmativo e o valor 0 em caso negativo;
- **Comorbilidade (Comorbidity)** - Indica se o doente tem/teve outras doenças, assumindo o valor 1 em caso afirmativo e o valor 0 em caso negativo;
- **Consanguinidade (Consanguinity)** - Indica se existe (1) ou não (0) consanguinidade entre os ascendentes;
- **Contacto com herbicidas/adubos (HerbicideAdubosContact);**
- **Contacto com tintas/diluentes (InksDiluentesContact);**

- **Contacto com outros produtos tóxicos (ToxicsOtherContact);**
- **Copos de álcool diários (GlassesAlcoholDay);**
- **Data de Inclusão (InclusionDate)** - Indica o ano em que o paciente foi incluído no estudo;
- **Diagnóstico de demência segundo o melhor julgamento do médico (DementiaBestMedical)** - Indica se o médico ajuíza que o doente sofre de demência, assumindo o valor 1 em caso afirmativo, o valor 0 em caso negativo, e o valor 2 no caso de resposta desconhecida;
- **Dieta (Diet)** - Indica o tipo de dieta alimentar do paciente, assumindo os valores 0, no caso de ser uma dieta variada/comum, 1 no caso de ser uma dieta vegetariana, e 2 no caso de ser outra dieta;
- **Discinesias (DiskynesiasEver)** - Indica se o doente alguma vez teve discinésias (1) ou não (0) ou se a resposta é desconhecida (2);
- **Disfagia (DysphagiaEver)** - Indica se o doente alguma vez teve disfagia (1) ou não (0) ou se a resposta é desconhecida (2);
- **Dislipidémia (Dislipidemia)** - Indica se o pacientente tem dislipidémia (1) ou não (0);
- **Distonia (DystoniaEver)** - Indica se o doente alguma vez teve distonia (1) ou não (0) ou se a resposta é desconhecida (2);
- **Disartria (DysarthriaEver)** - Indica se o doente alguma vez teve disartria (1) ou não (0) ou se a resposta é desconhecida (2);
- **Discinésias na examinação (DyskinesiaIIIA)** - Indica se foram visíveis discinécias aquando da examinação (1) ou não (0);
- **Duração de alcoolismo (NumberYearsAlcoholicHabits);**
- **Duração de hábitos cafeínicos (NumberYearsCaffeinatedHabits);**

- **Duração de tabagismo (NumberYearsSmoking);**
- **BDIa tristeza (BDIa)** - Indica a resposta do paciente à pergunta a (tristeza) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIb desencorajamento em relação ao futuro (BDIb)** - Indica a resposta do paciente à pergunta b (desencorajamento em relação ao futuro) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIc fracasso (BDIc)** - Indica a resposta do paciente à pergunta c (fracasso) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDId satisfação (BDId)** - Indica a resposta do paciente à pergunta d (satisfação) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIe culpa (BDIe)** - Indica a resposta do paciente à pergunta e (culpa) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIf desilusão consigo próprio (BDIf)** - Indica a resposta do paciente à pergunta f (desilusão consigo próprio) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIg pensamento suicída (BDIg)** - Indica a resposta do paciente à pergunta g (pensamento suicída) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIh interesse pelos outros (BDIh)** - Indica a resposta do paciente à pergunta h (interesse pelos outros) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIi toma de decisões (BDIi)** - Indica a resposta do paciente à pergunta i (toma de decisões) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;

- **BDIj autoconfiança (BDIj)** - Indica a resposta do paciente à pergunta j (autoconfiança) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIk capacidade de trabalho (BDIk)** - Indica a resposta do paciente à pergunta k (capacidade de trabalho) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIl cansaço (BDIl)** - Indica a resposta do paciente à pergunta l (cansaço) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDIm apetite (BDIm)** - Indica a resposta do paciente à pergunta m (apetite) da Escala de depressão de Beck, que pode assumir os valores 0, 1, 2 ou 3. Sendo o sintoma tão mais grave quanto maior o número escolhido;
- **BDItotal (BDItotal)** - Indica o valor total da Escala de Beck;
- **Mobilidade, Escala de London (LondonMobility)** - Indica a opção escolhida pelo paciente referente à mobilidade de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);
- **Independência física, Escala de London (LondonPhysicalIndependance)** - Indica a opção escolhida pelo paciente referente à independência física de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);
- **Ocupação, Escala de London (LondonOccupation)** - Indica a opção escolhida pelo paciente referente à ocupação (trabalho e lazer) de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);
- **Integração social, Escala de London (LondonSocialIntegration)** - Indica a opção escolhida pelo paciente referente à integração social de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);

- **Orientação, Escala de London (LondonOrientation)** - Indica a opção escolhida pelo paciente referente à orientação (perceber o que se passa à sua volta) de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);
- **Independência económica, Escala de London (LondonEconomicIndependance)** - Indica a opção escolhida pelo paciente referente à independência económica de acordo com a Escala de London. Esta variável pode assumir os valores: 1 (de forma alguma), 2 (um pouco), 3 (bastante), 4 (muito), 5 (quase totalmente) , 6 (totalmente);
- **Estado clínico(UPDRSPatientClinicalState)** - Indica o estado clínico do paciente, podendo este ser ON se o paciente está bem (sob o efeito da medicação) ou OFF se o paciente não estava bem. Esta variável assume o valor 1 se o paciente está ON e assume o valor 0 caso contrário;
- **Estado Civil (EstadoCivil)** - Indica o estado civil do paciente, podendo assumir os valores 0 (solteiro), 1 (casado), 2 (divorciado), 3 (vive em união de facto) e 4 (viúvo);
- **Estado HY (HYstage)** - Indica o estado em que se encontra o paciente de acordo com a escala de Hoehn e Yahr. Esta variável pode assumir os valores: 0 (se não existirem sinais de doença), 1 (se a doença se manifestar apenas unilateralmente), 1.5 (se a doença se manifestar unilateralmente e houver envolvimento axial), 2 (manifestação da doença bilateralmente sem haver alteração do equilíbrio), 2.5 (doença bilateral ligeira, com recuperação na prova de equilíbrio), 3 (doença bilateral leve a moderada, alguma instabilidade postural, fisicamente independente), 4 (verifica-se uma incapacidade severa, mas consegue caminhar ou levantar-se sem ajuda), 5 (se o doente não for ajudado só consegue estar em cadeira de rodas ou acamado);
- **Estado Mental (MMSE)** - Indica a pontuação obtida no Mini Mental State Examination. Esta variável pode assumir valores inteiros entre 0 e 30;
- **Estado SE (SEscale)** - Indica a percentagem em que se encontra o paciente de acordo com a escala de Schwab e England. Esta variável pode assumir valores de 0 a 100
- **Etnia (Ethnicity)** - Indica a etnia do paciente: 1 (etnia europeia) e 0 (etnia africana);
- **Ensino (YearsAtSchool)** - Indica o número de anos em que o paciente frequentou o ensino escolar;

- **Filhos (SonsDaughters)** - Indica se o paciente tem filhos(as) (1) ou não (0);
- **Fim de alcoolismo (EndAlcoholicHabitsAge)** - Indica a idade em que o paciente deixou de ter hábitos etílicos;
- **Fim de tabagismo (SmokingUntilAge)** - Indica até que idade o paciente fumou;
- **Fim de hábitos cafeínicos (EndCaffeinatedHabitsAge)** - Indica a idade do paciente, quando este deixou de beber bebidas com cafeína;
- **Flutuações motoras (MotorFluctuationsEver)** - Indica se o doente alguma vez teve flutuações motoras (1) ou não (0), ou se a resposta é desconhecida (2);
- **Fonte de informação de UPDRS (UPDRSsource)** - Indica a pessoa que forneceu informação para as repostas referentes à escala UPDRS. Esta variável assume os valores 1, no caso da informação ter sido fornecida pelo paciente; 2, no caso de ter sido o cuidador a fornecer a informação; e 3, se tiverem sido ambos;
- **Freezing (FreezingEver)** - Indica se o doente alguma vez teve freezing (1) ou não (0) ou se a resposta é desconhecida (2);
- **Fumador passivo(em casa) (PassiveSmokerHome)**;
- **Fumador passivo(no trabalho) (PassiveSmokerWork)**;
- **Género(Gender)** - Indica o sexo do paciente: 1(feminino) e 0 (masculino);
- **Hábitos cafeínicos (PresentCaffeinatedBeverages)** - Indica o tipo de bebida com cafeína que o paciente bebe ou bebia com regularidade. Esta variável pode assumir os valores: 0 (café), 1 (chá) ou 2 (coca-cola);
- **Idade (DateBirth)** - Indica a idade do paciente à data de 01 de Janeiro de 2011;
- **Idade aquando dos primeiros sintomas motores (DateFirstMotorSymptoms)**;
- **Idade aquando do diagnóstico (DateDiagnosis)**;

- **Idade reforma/desemprego (AgeRetirementorUnempliement);**
- **Influência das discinésias (DyskinesiaIIIB)** - Indica se as discinésias afectaram a observação do doente (1) ou não (0);
- **Início de tabagismo (AgeOnsetSmoking)** - Indica a idade do paciente quando este começou a fumar;
- **Instabilidade postural como primeiro sintoma (PosturalInstabilityFMS)** - Indica se o primeiro sintoma da doença foi instabilidade postural (1) ou não (0) ou se a resposta é desconhecida (2);
- **Instabilidade postural (PosturalInstEver)** - Indica se o doente alguma vez teve instabilidade postural, assumindo o valor 1 no caso afirmativo, o valor 0 no caso negativo e o valor 2 em caso de resposta desconhecida;
- **Minutos desde a medicação(MinutesSinceLevodopa);**
- **Motivo de reforma/desemprego (MotiveRetirementorunempliment)** - Indica o motivo pela qual o doente se reformou ou ficou desempregado. Esta variável assume os valores: 0 se o motivo foi doença, 1 se o motivo foi a idade, 2 se o motivo foi uma reestruturação do local de trabalho ou 3 se se deveu a motivos familiares;
- **NMSQ1 excesso e queda de saliva (NMSQ1)** - Indica a resposta do paciente à pergunta 1 (excesso/queda de saliva) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ2 perda ou alteração da capacidade de saborear/cheirar (NMSQ2)** - Indica a resposta do paciente à pergunta 2 (perda ou alteração da capacidade de saborear/cheirar) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ3 dificuldade em engolir (NMSQ3)** - Indica a resposta do paciente à pergunta 3 (dificuldade em engolir) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ4 vômitos/náuseas (NMSQ4)** - Indica a resposta do paciente à pergunta 4 (vômitos/náuseas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);

- **NMSQ5 obstipação (NMSQ5)** - Indica a resposta do paciente à pergunta 5 (obstipação) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ6 incontinência de fezes (NMSQ6)** - Indica a resposta do paciente à pergunta 6 (incontinência de fezes) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ7 sensação de não defecar completamente (NMSQ7)** - Indica a resposta do paciente à pergunta 7 (sensação de não defecar completamente) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ8 sensação de urgência em urinar (NMSQ8)** - Indica a resposta do paciente à pergunta 8 (sensação de urgência em urinar) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ9 levantar-se regularmente para ir à casa de banho durante a noite (NMSQ9)** - Indica a resposta do paciente à pergunta 9 (levantar-se regularmente para ir à casa de banho durante a noite) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ10 dores inesperadas (NMSQ10)** - Indica a resposta do paciente à pergunta 10 (dores inesperadas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ11 alteração inesperada de peso (NMSQ11)** - Indica a resposta do paciente à pergunta 11 (alteração inesperada de peso) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ12 problemas em lembrar-se de acontecimentos/tarefas (NMSQ12)** - Indica a resposta do paciente à pergunta 12 (problemas em lembrar-se de acontecimentos/tarefas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);

- **NMSQ13 perda de interesse (NMSQ13)** - Indica a resposta do paciente à pergunta 13 (perda de interesse) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ14 ver/ouvir coisas que não estão lá (NMSQ14)** - Indica a resposta do paciente à pergunta 14 (ver/ouvir coisas que não estão lá) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ15 dificuldade de concentração (NMSQ15)** - Indica a resposta do paciente à pergunta 15 (dificuldade de concentração) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ16 tristeza (NMSQ16)** - Indica a resposta do paciente à pergunta 16 (tristeza) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ17 ansiedade (NMSQ17)** - Indica a resposta do paciente à pergunta 17 (ansiedade) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ81 interesse em sexo diferente do habitual (NMSQ18)** - Indica a resposta do paciente à pergunta 18 (interesse em sexo diferente do habitual) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ19 dificuldade em ter sexo (NMSQ19)** - Indica a resposta do paciente à pergunta 19 (dificuldade em ter sexo) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ20 tonturas ou fraqueza (NMSQ20)** - Indica a resposta do paciente à pergunta 20 (tonturas ou fraqueza) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ21 quedas (NMSQ21)** - Indica a resposta do paciente à pergunta 21 (quedas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);

- **NMSQ22 dificuldade em manter-se acordado (NMSQ22)** - Indica a resposta do paciente à pergunta 22 (dificuldade em manter-se acordado) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ23 dificuldade em adormecer (NMSQ23)** - Indica a resposta do paciente à pergunta 23 (dificuldade em adormecer) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ24 sonhos intensos (NMSQ24)** - Indica a resposta do paciente à pergunta 24 (sonhos intensos) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ25 falar/mexer-se durante o sono (NMSQ25)** - Indica a resposta do paciente à pergunta 25 (falar/mexer-se durante o sono) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ26 sensações desagradáveis nas pernas (NMSQ26)** - Indica a resposta do paciente à pergunta 26 (sensações desagradáveis nas pernas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ27 pernas inchadas (NMSQ27)** - Indica a resposta do paciente à pergunta 27 (pernas inchadas) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ28 suar em excesso (NMSQ28)** - Indica a resposta do paciente à pergunta 28 (suar em excesso) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ29 visão a dobrar (NMSQ29)** - Indica a resposta do paciente à pergunta 29 (visão a dobrar) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);
- **NMSQ30 acreditar que lhe acontecem coisas que não são verdade (NMSQ30)** - Indica a resposta do paciente à pergunta 30 (acreditar que lhe acontecem coisas que não são verdade) do questionário de sintomas não motores. Esta variável pode assumir os seguintes valores: 0 (sim), 1 (não sei) ou 2 (não);

- **Número diário de bebidas com cafeína (NumberCaffeinatedBeveragesDay);**
- **Número de traumatismos cranianos com perda de consciência (Number-HeadTraumaLossConscience);**
- **Número de Filhos (NumberSonsDaughters);**
- **Ocupação actual (PresentJob) - Esta variável é de resposta aberta;**
- **País de Nascimento(CountryBirth) - Esta variável é de resposta aberta;**
- **Perturbação do sono REM (RSBDLastMonth) - Indica se o doente teve perturbação do sono REM (Rapid Eye Movement) no último mês (relativamente à data de entrada no estudo), assumindo o valor 1 em caso afirmativo e o valor 0 em caso negativo;**
- **Peso (Weight);**
- **Progressão (RESPOSTA) - Esta é a variável resposta do estudo, e indica se o paciente tem progressão rápida (1) ou lenta (0) da doença, com base na pergunta UPDRS3.12, referente à estabilidade postural;**
- **Punding (Punding) - Indica se no último mês (em relação à data de entrada no estudo) o doente teve algum sintoma de punding (1) ou não (0);**
- **Quedas (FallsEver) - Indica se o doente alguma vez teve quedas (1) ou não (0) ou se a resposta é desconhecida (2);**
- **Raça(Race) - Indica a raça do paciente: 1 (raça branca) e 0 (raça negra);**
- **Relação familiar para com outro familiar com DP (FamilialPDrelation) - Indica o grau de parentesco entre o doente e o(s) seu(s) familiar(s) com doença de Parkinson. Os valores assumidos por esta variável são obtidos de acordo com a escala do coeficiente de parentesco de Sewall Wright;**
- **Rigidez como primeiro sintoma (RigidityFMS) - Indica se o primeiro sintoma da doença foi rigidez (1) ou não (0) ou se a resposta é desconhecida (2);**

- **Religião (Religion)** - Indica a religião do paciente: 0 (ateu), 1 (católico), 2 (outro), 3 (judeu), 4 (hindu), 5 (muçulmano) e 6 (budista);
- **Residência em propriedade agrícola (EverResidentFarm)** - Indica se o paciente alguma vez viveu numa propriedade agrícola (1) ou não (0) ;
- **Tabagismo (SmokingHabits)** - Indica o tipo de hábitos tabágicos do paciente: 0 (nunca fumou), 1 (deixou de fumar), 2 (fumador), 3 (fumador passivo);
- **Toma de Levodopa(UPDRSPatientOnLevodopa)** - Indica se o paciente toma Levodopa (1) ou não (0);
- **Tempo para demência (DementiaOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de demência;
- **Tempo para início de tremor dos membros (LimbTremorOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início do tremor;
- **Tremor como primeiro sintoma(TremorFMS)** - Indica se o primeiro sintoma da doença foi tremor (1) ou não (0) ou se a resposta é desconhecida (2);
- **Tempo para início de quedas (FallsOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de quedas;
- **Tempo para início de flutuações motoras (MotorFluctuationsOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de flutuações motoras;
- **Tempo para início de discinesias (DiskynesiasOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de discinesias;
- **Tempo para início de distonia (DystoniaOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de distonias;
- **Tempo para início de freezing (FreezingOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de freezing;

- **Tempo para início de benefício motor (MotorSleepBenefitOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de benefício motor;
- **Tempo para início de disfagia (DysphagiaOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de disfagia;
- **Tempo para início de disartria (DysarthriaOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de disartria;
- **Trabalhos anteriores (PreviousJobs)** - Esta pergunta é de resposta aberta;
- **Tremor dos membros (LimbTremorEver)** - Indica se o doente alguma vez teve tremor dos membros, se sim assume o valor 1 se não assume o valor 0, se a resposta for desconhecida assume o valor 2;
- **Último trabalho (LastJob)** - Indica em que nível da actividade laboral se encontra o paciente, assumindo os valores: 1 para engenheiro, 2 para empresário, 3 para doméstica, 4 para motorista, 5 para empregado de escritório/administrativo/secretário/funcionário público, 6 para professor, 7 para operador de máquinas (técnico/manobrador/lubrificador), 8 para funcionário do metro, 9 comerciante, 10 para pedreiro/pintor/electricista, 11 para trabalho ligado à agricultura, 12 para operário fabril/chefe de secção/operador central, 13 para porteiro, 14 para cabeleireiro/barbeiro, 15 para serralheiro/mecânico, 16 para pasteleiro, 17 para empregado de balcão, 18 para trabalho ligado a sondagem, 19 para costureiro, 20 para bancário, 21 para polícia, 22 para oficial de operações de voo;
- **UPDRS1.1 alterações cognitivas (UPDRS1.1)** - Indica a resposta à pergunta 1 (alterações cognitivas) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.2 alucinações e psicose (UPDRS1.2)** - Indica a resposta à pergunta 2 (alucinações e psicose) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.3 humor deprimido (UPDRS1.3)** - Indica a resposta à pergunta 3 (humor deprimido) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS1.4 humor ansioso (UPDRS1.4)** - Indica a resposta à pergunta 4 (humor ansioso) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.5 apatia (UPDRS1.5)** - Indica a resposta à pergunta 5 (apatia) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.6 sintomas relativos a síndrome de desregulação dopaminérgica (UPDRS1.6)** - Indica a resposta à pergunta 6 (sintomas relativos a síndrome de desregulação dopaminérgica) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.7 problemas com o sono (UPDRS1.7)** - Indica a resposta à pergunta 7 (problemas com o sono) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.8 sonolência diurna (UPDRS1.8)** - Indica a resposta à pergunta 8 (sonolência diurna) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.9 dor e outras sensações (UPDRS1.9)** - Indica a resposta à pergunta 9 (dor e outras sensações) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.10 problemas urinários (UPDRS1.10)** - Indica a resposta à pergunta 10 (problemas urinários) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.11 obstipação (UPDRS1.11)** - Indica a resposta à pergunta 11 (obstipação) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS1.12 tonturas ao levantar (UPDRS1.12)** - Indica a resposta à pergunta 12 (tonturas ao levantar) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS1.13 cansaço (UPDRS1.13)** - Indica a resposta à pergunta 13 (cansaço) da escala UPDRS referente a sintomas não motores. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.1 fala (UPDRS2.1)** - Indica a resposta à pergunta 1 (fala) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.2 saliva e babar-se (UPDRS2.2)** - Indica a resposta à pergunta 2 (saliva e babar-se) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.3 mastigar e engolir (UPDRS2.3)** - Indica a resposta à pergunta 3 (mastigar e engolir) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.4 comer (UPDRS2.4)** - Indica a resposta à pergunta 4 (comer) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.5 vestir (UPDRS2.5)** - Indica a resposta à pergunta 5 (vestir) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.6 higiene (UPDRS2.6)** - Indica a resposta à pergunta 6 (higiene) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.7 escrita (UPDRS2.7)** - Indica a resposta à pergunta 7 (escrita) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.8 ocupação de tempos livres (UPDRS2.8)** - Indica a resposta à pergunta 8 (ocupação de tempos livres) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS2.9 virar-se na cama (UPDRS2.9)** - Indica a resposta à pergunta 9 (virar-se na cama) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.10 tremor (UPDRS2.10)** - Indica a resposta à pergunta 10 (tremor) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.11 levantar-se da cama, de um carro, ou de uma cadeira profunda (UPDRS2.11)** - Indica a resposta à pergunta 11 (levantar-se da cama, de um carro, ou de uma cadeira profunda) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.12 (UPDRS2.12)** - Indica a resposta à pergunta 12 da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS2.13 freezing (UPDRS2.13)** - Indica a resposta à pergunta 13 (freezing) da escala UPDRS referente a aspectos motores de experiências do dia a dia. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.1 fala (UPDRS3.1)** - Indica a resposta à pergunta 1 (fala) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.2 expressão facial (UPDRS3.2)** - Indica a resposta à pergunta 2 (expressão facial) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.3 rigidez do pescoço (UPDRS3.3Neck)** - Indica a resposta à pergunta 3 (rigidez do pescoço) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.3 rigidez do membro superior direito (UPDRS3.3RUE)** - Indica a resposta à pergunta 3 (rigidez do membro superior direito) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS3.3 rigidez do membro superior esquerdo (UPDRS3.3LUE)** - Indica a resposta à pergunta 3 (rigidez do membro superior esquerdo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.3 rigidez do membro inferior direito (UPDRS3.3RLE)** - Indica a resposta à pergunta 3 (rigidez do membro inferior direito) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.3 rigidez do membro inferior esquerdo (UPDRS3.3LLE)** - Indica a resposta à pergunta 3 (rigidez do membro inferior esquerdo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.4 bater de dedos da mão direita (UPDRS3.4R)** - Indica a resposta à pergunta 4 (bater de dedos da mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.4 bater de dedos da mão esquerda (UPDRS3.4L)** - Indica a resposta à pergunta 4 (bater de dedos da mão esquerda) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.5 movimento de mãos- mão direita (UPDRS3.5R)** - Indica a resposta à pergunta 5 (movimento de mãos- mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.5 movimento de mãos- mão esquerda (UPDRS3.5L)** - Indica a resposta à pergunta 5 (movimento de mãos- mão esquerda) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.6 supinação das mãos- mão direita (UPDRS3.6R)** - Indica a resposta à pergunta 6 (supinação das mãos- mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS3.6 supinação das mãos-mão esquerda (UPDRS3.6L)** - Indica a resposta à pergunta 6 (supinação das mãos-mão esquerda) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.7 bater do pé direito no chão (UPDRS3.7R)** - Indica a resposta à pergunta 7 (bater do pé direito no chão) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.7 bater do pé esquerdo no chão (UPDRS3.7L)** - Indica a resposta à pergunta 7 (bater do pé esquerdo no chão) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.8 agilidade do membro inferior direito (UPDRS3.8R)** - Indica a resposta à pergunta 8 (agilidade do membro inferior direito) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.8 agilidade do membro inferior esquerdo (UPDRS3.8L)** - Indica a resposta à pergunta 8 (agilidade do membro inferior esquerdo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.9 levantar-se da cadeira (UPDRS3.9)** - Indica a resposta à pergunta 9 (levantar-se da cadeira) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.10 marcha (UPDRS3.10)** - Indica a resposta à pergunta 10 (marcha) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.11 freezing da marcha (UPDRS3.11)** - Indica a resposta à pergunta 11 (freezing da marcha) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS3.13 postura (UPDRS3.13)** - Indica a resposta à pergunta 13 (postura) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.14 bradicinésia do corpo (UPDRS3.14)** - Indica a resposta à pergunta 14 (bradicinésia do corpo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.15 tremor postural- mão direita (UPDRS3.15R)** - Indica a resposta à pergunta 15 (tremor postural- mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.15 tremor postural- mão esquerda (UPDRS3.15L)** - Indica a resposta à pergunta 15 (tremor postural- mão esquerda) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.16 tremor de acção da mão direita (UPDRS3.16R)** - Indica a resposta à pergunta 16 (tremor de acção da mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.16 tremor de acção da mão mão esquerda (UPDRS3.16L)** - Indica a resposta à pergunta 16 (tremor de acção da mão direita) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.17 amplitude do tremor de repouso- lábio e mandíbula (UPDRS3.17lipJaw)** - Indica a resposta à pergunta 17 (amplitude do tremor de repouso- lábio e mandíbula) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.17 amplitude do tremor de repouso do membro superior direito (UPDRS3.17RUE)** - Indica a resposta à pergunta 17 (amplitude do tremor do membro superior direito) escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS3.17 amplitude do tremor de repouso do membro superior esquerdo (UPDRS3.17LUE)** - Indica a resposta à pergunta 17 (amplitude do tremor do membro superior esquerdo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.17 amplitude do tremor de repouso do membro inferior direito (UPDRS3.17RLE)** - Indica a resposta à pergunta 17 (amplitude do tremor de repouso do membro inferior direito) escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.17 amplitude do tremor de repouso do membro inferior esquerdo (UPDRS3.17LLE)** - Indica a resposta à pergunta 17 (amplitude do tremor de repouso do membro inferior esquerdo) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS3.18 constância do tremor de repouso(UPDRS3.18)** - Indica a resposta à pergunta 18 (constância do tremor de repouso) da escala UPDRS referente ao exame motor. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **Tempo para início de instabilidade postural (PosturalInstOnsetYear)** - Indica o tempo (em anos) medido entre o aparecimento do primeiro sintoma da doença e o início de instabilidade postural;
- **UPDRS4.1 tempo de duração das discinésias (UPDRS4.1)** - Indica a resposta à pergunta 1 (tempo de duração das discinésias) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS4.2 impacto funcional das discinésias (UPDRS4.2)** - Indica a resposta à pergunta 2 (impacto funcional das discinésias) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS4.3 tempo de duração “OFF” (UPDRS4.3)** - Indica a resposta à pergunta 3 (tempo de duração “OFF”) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

- **UPDRS4.4 impacto funcional das flutuações (UPDRS4.4)** - Indica a resposta à pergunta 4 (impacto funcional das flutuações) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS4.5 complexidade das flutuações motoras (UPDRS4.5)** - Indica a resposta à pergunta 5 (complexidade das flutuações motoras) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);
- **UPDRS4.6 distonia do “OFF” dolorosa (UPDRS4.6)** - Indica a resposta à pergunta 6 (distonia do “OFF” dolorosa) da escala UPDRS referente a complicações motoras. Esta variável assume os valores 0 (normal), 1 (ligeiro), 2 (moderado), 3 (marcado), 4 (severo);

3.2 Análise Descritiva

Nesta secção faz-se uma análise descritiva das variáveis com recurso a histogramas, diagramas de caixa, tabelas de contingência, gráficos de barras, à ideia de correlação e a medidas descritivas como o mínimo, máximo, média, mediana, 1^o e 3^o quartis.

O trabalho computacional indispensável ao tratamento estatístico dos dados foi efectuado com recurso ao *software* estatístico R, R Core Team (2009), incluindo funções já existentes neste *software* e a funções construídas com base na linguagem fornecida.

- **Idade dos pacientes em relação à progressão da doença**

No sentido de perceber se a idade se revela diferente nos dois grupos, o de progressão rápida (1) e o de progressão lenta (0) fizeram-se os diagramas de caixa, representados na Figura 3.1.

Em termos de idade, não se observam diferenças significativas entre os grupos. Embora, se perceba que os doentes mais velhos pertencem ao grupo de progressão rápida, o que é expectável, já que se tem constatado que os efeitos da doença são mais severos nos doentes com idade mais avançada.

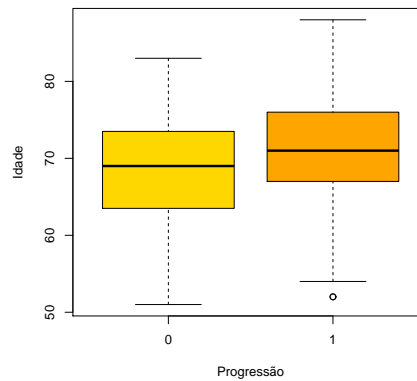


Figura 3.1: Diagramas de caixa da idade do doente em relação à progressão

- **Idade em que apareceram os primeiros sintomas e Idade em que foi diagnosticada a doença**

Na maioria dos doentes a idade em que surgiram os primeiros sintomas coincide com a idade em que foi diagnosticada a doença, tal deve-se ao facto de os pacientes não conseguirem precisar a idade em que tiveram os primeiros sintomas, ou de os terem desvalorizado quando estes surgiram. Na Figura 3.2 mostram-se os histogramas referentes a estas duas idades.

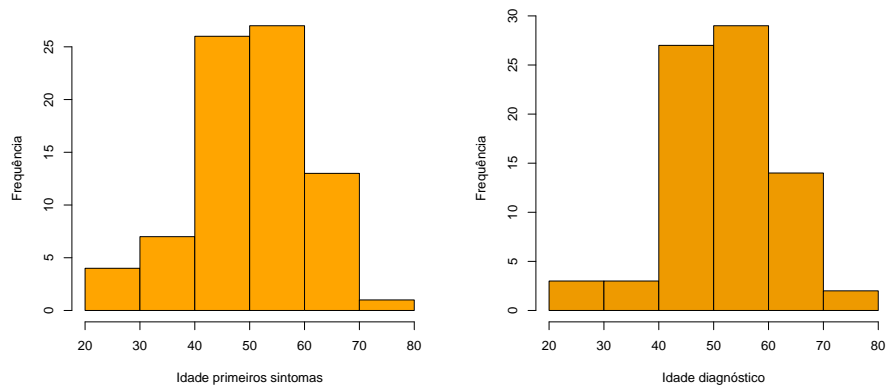


Figura 3.2: Histogramas da idade dos primeiros sintomas e da idade de diagnóstico

Como se pode observar a maioria dos doentes teve os primeiros sintomas e foi diagnosticado com a doença de Parkinson entre os 40 e os 60 anos, que é mais ou menos o que se esperaria, de acordo com a literatura. Outra questão interessante, que se observa nos dados brutos, é o facto de em muitos casos a data de diagnóstico e a data em que surgiram os primeiros sintomas coincidirem ou serem muito próximas. Tal dever-se-á não só ao facto de os pacientes não

se recordarem da data em que tiveram os primeiros sintomas, ou de os terem desvalorizado quando surgiram; mas também, ao diagnóstico ser feito no início da doença.

Nas tabelas 3.1 e 3.2 pode ver-se um resumo da distribuição das idades dos pacientes.

Tabela 3.1: Sumário da idade em que os doentes tiveram os primeiros sintomas da doença

	Min	1°Q	Média	Mediana	3°Q	Máx	NA
Idade	24	45	50.69	51	57.75	74	0

Tabela 3.2: Sumário da idade em que foi feito o diagnóstico

	Min	1°Q	Média	Mediana	3°Q	Máx	NA
Idade	25	46	52.22	52	58.75	77	0

Como seria de esperar pela observação dos gráficos de barras não existem diferenças significativas entre a idade de diagnóstico e as idades dos primeiros sintomas.

- **Primeiros Sintomas**

Em geral, os primeiros sintomas da doença de Parkinson são: tremor, rigidez, bradicinésia e instabilidade postural. Nas tabelas de contingência 3.3, 3.4, 3.5 e 3.6 que relacionam cada um destes sintomas com a pertença a cada um dos grupos, pode ver-se essa relação.

Tabela 3.3: Progressão vs Tremor

		Tremor			
		0 (Não)	1 (Sim)	2 (Desconhecido)	
Progressão	0 (lenta)	15	28	0	43
	1 (rápida)	13	22	0	35
		28	50	0	78

Os dados revelam uma certa homogeneidade nos grupos no que diz respeito aos sintomas: tremor, rigidez e bradicinésias. No entanto, a Tabela 3.6 mostra que a maioria dos pacientes

Tabela 3.4: Progressão vs Rigidez

		Rigidez			
		0 (Não)	1 (Sim)	2 (Desconhecido)	
Progressão	0 (lenta)	24	14	5	43
	1 (rápida)	18	15	2	35
		42	29	7	78

Tabela 3.5: Progressão vs Bradicínésia

		Bradicínésia			
		0 (Não)	1 (Sim)	2 (Desconhecido)	
Progressão	0 (lenta)	17	24	2	43
	1 (rápida)	11	22	2	35
		28	46	4	78

Tabela 3.6: Progressão vs Instabilidade postural

		Instabilidade Postural			
		0 (Não)	1 (Sim)	2 (Desconhecido)	
Progressão	0 (lenta)	39	0	4	43
	1 (rápida)	29	3	3	35
		68	3	7	78

não teve como primeiro sintoma a instabilidade postural, pelo que esta variável não trará informação muito relevante para o objectivo de estudo. Há ainda a referir que, de acordo com a literatura, tipicamente os doentes cujo primeiro sintoma é o tremor, tendem a progredir mais lentamente, e de facto, a maioria dos doentes pertencentes ao grupo de progressão lenta (0) tiveram como primeiro sintoma da doença o tremor.

- **Demência em relação à progressão da doença**

Tendo em conta que a demência revela um estado mais avançado da doença, é interessante

perceber até que ponto os doentes com demência se incluem no grupo dos que têm uma progressão rápida da doença ou dos que têm uma progressão lenta. Recorde-se que esta variável assume os valores 0 (não tem demência), 1 (tem demência) e 2 (resposta desconhecida), e observe-se o gráfico de barras da 3.3:

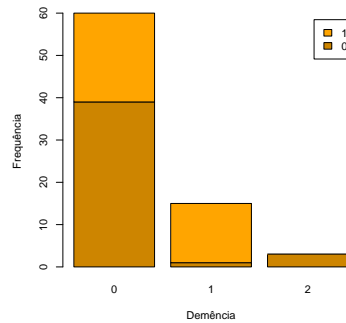


Figura 3.3: Gráfico de barras do estado de demência do paciente em relação à progressão

De facto, a esmagadora maioria dos doentes com demência pertence ao grupo de progressão rápida.

- **Vícios**

Neste ponto, vai observar-se a distribuição da duração dos vícios dos pacientes, nomeadamente, dos hábitos tabágicos, etílicos e cafeínicos que podem afectar a saúde da população.

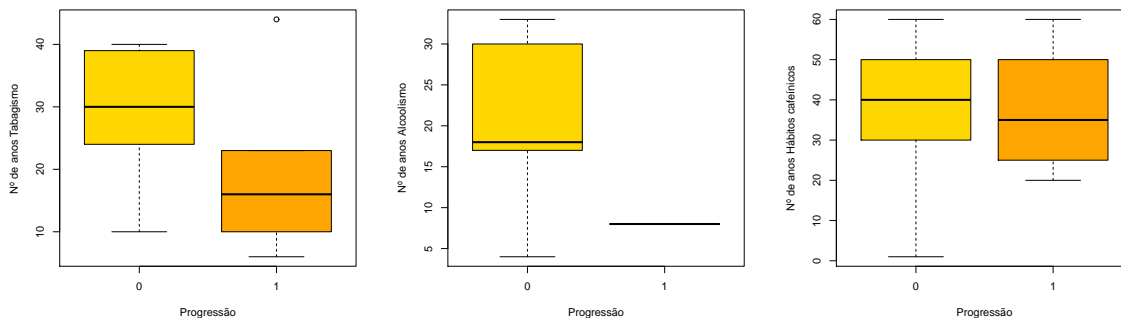


Figura 3.4: Diagramas de caixa da duração dos vícios

Não se nota uma diferença significativa entre os grupos no que diz respeito aos hábitos cafeínicos. No entanto, em relação ao tabaco e ao álcool, percebem-se diferenças entre os grupos, mas há que ter em conta, que estas variáveis têm poucas observações, o que resulta do facto

de apenas uma parte dos pacientes ter ou ter tido hábitos alcoólicos e/ou tabágicos, ou de não terem revelado essa informação.

- **Características físicas (peso e altura)**

Na Figura 3.5 mostram-se os diagramas de caixa referentes ao peso e altura relacionados com o progresso da doença.

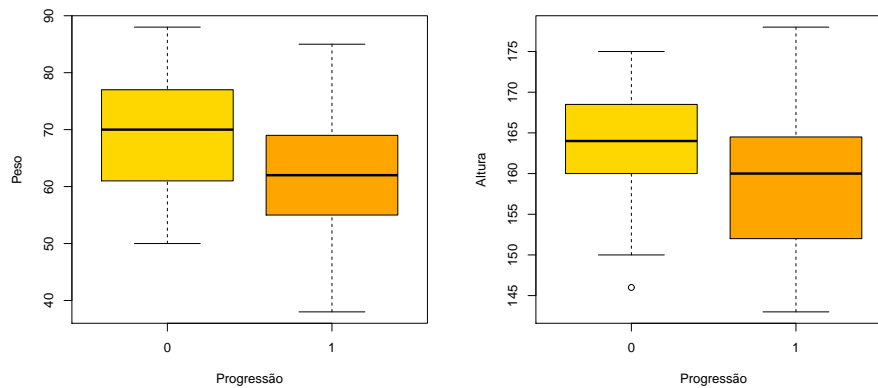


Figura 3.5: Diagramas de caixa das características físicas para cada um dos grupos de progressão

Estas características físicas não parecem evidenciar diferenças significativas entre os dois grupos de progressão.

- **Cirurgia**

Observe-se de seguida, na Tabela 3.7, o grupo de progressão a que pertencem os doentes já sujeitos à cirurgia de Parkinson.

Constata-se que a maioria dos doentes sujeitos a DBS pertence ao grupo de progressão lenta da doença, tal dever-se-á em parte ao enviesamento que advém da selecção de doentes para cirurgia, já que nos critérios para cirurgia se excluem, por exemplo, doentes com mais de 70 anos, dementes, e com algumas complicações terapêuticas.

Tabela 3.7: Progressão vs Cirurgia (DBS)

		DBS		
		0 (Não)	1 (Sim)	
Progressão	0 (lenta)	24	19	43
	1 (rápida)	29	6	35
		53	25	78

- **Estadio Hoehn e Yahr**

A escala de estadio da doença de Hoehn e Yahr é, como já referido no capítulo anterior, um método simples de estadiamento da doença. Na Figura 3.6 apresenta-se o gráfico de barras dos diferentes estadios mostrando-se a separação por cada um dos grupos de progressão da doença.

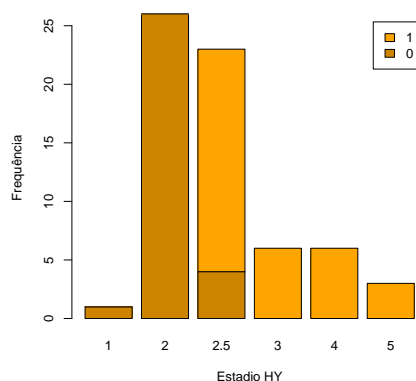


Figura 3.6: Gráfico de barras do estadio HY em relação à progressão

Como se pode observar a maioria dos doentes que apresenta uma progressão lenta da doença (grupo 0), encontra-se nos estadios mais baixos de Hohen e Yahr, o que faz todo o sentido, tendo em conta que o factor considerado para a separação dos grupos é a instabilidade postural. Por outro lado, a maioria dos doentes com progressão rápida (grupo 1), encontram-se nos estadios superiores, o que seria de esperar dado que são estadios de maior incidência da doença, e estes apresentam uma evolução mais rápida da doença.

- **Escala Schwab e England**

À semelhança do que foi feito no ponto anterior, observe-se na Figura 3.7 a relação entre cada um dos grupos de progressão e a escala de Schwab e England.

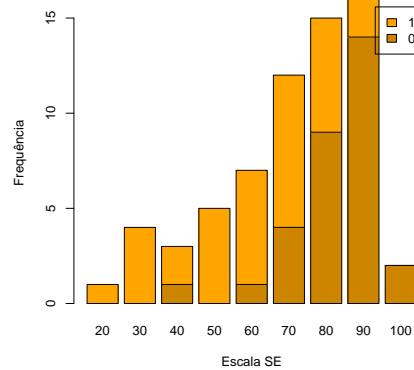


Figura 3.7: Gráfico de barras da escala de Schwab e England em relação à progressão

Neste caso é possível perceber que nas percentagens mais baixas da escala, relativas a uma maior incapacidade, não existem doentes com progressão lenta da doença, e nas percentagens mais elevadas da escala, relativas a uma maior independência, quase não existem doentes com progressão rápida da doença.

- **Exame do estado mental**

O MMSE é um exame rápido, com perguntas simples, que permite atribuir uma pontuação até 30 pontos, que avalia o estado mental do doente. Assim, veja-se a Tabela 3.8, que mostra um resumo dos valores dos pacientes (a sigla NA indica o número de valores não disponíveis):

Tabela 3.8: Sumário da pontuação obtida no MMSE

	Min	1ºQ	Média	Mediana	3ºQ	Max	NA
MMSE	2	23	28	25.56	29	30	15

Como se pode verificar, em média os doentes têm uma pontuação elevada (28) havendo, no entanto pontuações extremamente baixas (2), o que revelará doentes num estado avançado de doença. Veja-se também, na Figura 3.8 os diagramas de caixa para os diferentes grupos:

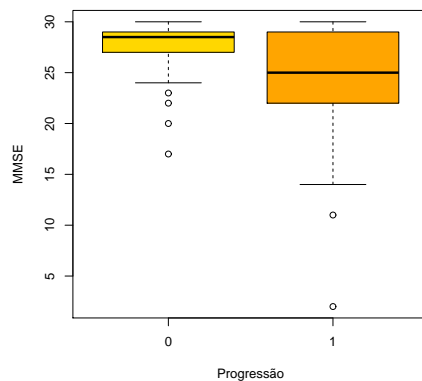


Figura 3.8: Diagrama de caixa para a pontuação MMSE em cada um dos grupos de progressão

Embora a diferença não seja muito acentuada, vê-se claramente que a pontuação do grupo com progressão lenta tende a ser superior relativamente ao grupo de progressão rápida da doença.

- **Estado Civil**

O estado civil é uma variável que *a priori* não parece estar muito relacionada com a evolução da doença de Parkinson, contudo, os médicos constataam que o facto dos doentes viverem sozinhos coloca-os maioritariamente no grupo de doentes com progressão rápida. Tal dever-se-á à falta de estimulação a nível mental, a alterações afectivas e também a nível físico.

Tabela 3.9: Progressão vs Estado Civil

		Estado Civil					
		0 (sol)	1 (cas)	2 (div)	3 (união)	4 (viú)	
Progressão	0 (lenta)	1	34	2	0	2	39
	1 (rápida)	0	24	1	0	2	27
		1	58	3	0	4	66

Neste caso não se podem tirar conclusões significativas na medida em que a esmagadora maioria dos doentes é casado.

- **Actividade física regular**

Actualmente sabe-se que a prática de exercício físico com regularidade é um factor muito importante para uma boa saúde. Neste sentido será interessante perceber a relação desta variável com a variável resposta (progressão).

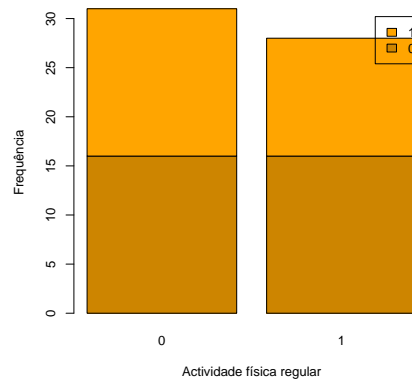


Figura 3.9: Progressão vs Actividade física regular

A Figura 3.9 mostra que o grupo de progressão lenta (0) é muito semelhante ao grupo de progressão rápida (1) no que diz respeito à actividade física regular, não havendo portanto evidência para considerar este factor como um factor de distinção entre estes dois grupos.

- **Factores internos**

Neste ponto, vai analisar-se a relação entre os dois grupos e a dislipidémia e ácido úrico elevados.

Tabela 3.10: Progressão vs Dislipidémia

		Dislipidémia		
		0 (Não)	1 (Sim)	
Progressão	0 (lenta)	19	16	35
	1 (rápida)	17	8	25
		36	24	60

Tabela 3.11: Progressão vs Ácido úrico elevado

		Ácido úrico elevado		
		0 (Não)	1 (Sim)	
Progressão	0 (lenta)	28	7	35
	1 (rápida)	25	1	26
		53	8	61

Infelizmente, não é possível aferir conclusões para nenhum dos casos uma vez que a maioria dos doentes não indica ter ácido úrico elevado (Tabela 3.11) nem dislipidémia (Tabela 3.10).

- **Escala de Beck**

A escala de Beck é usada para “medir” a intensidade da depressão. Assim, observa-se se existe alguma distinção entre os grupos progressão no que diz respeito aos níveis de depressão.

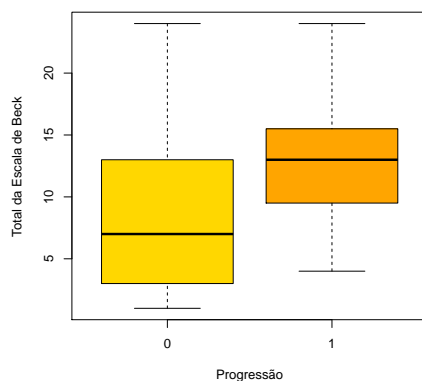


Figura 3.10: Diagramas de caixa para o total da Escala de Beck em relação à progressão da doença

Como se pode observar os doentes que pertencem ao grupo de progressão rápida (1) da doença apresentam valores de intensidade de depressão superiores em relação aos doentes que apresentam uma progressão lenta (0) da doença. Poder-se-ia pensar que este facto se deve ao facto de os doentes de progressão rápida da doença terem mais efeitos da doença, e portanto maior incapacidade, o que poderia conduzir à frustração e tristeza, porém esta depressão pode ser sintoma da própria doença, e não ser reactiva a esta.

• Escala de London

Como já referido no capítulo anterior esta escala procura medir em função de seis itens as limitações impostas pela doença, pelo que vamos observar os gráficos de barras para cada um destes itens em função do grupo de progressão.

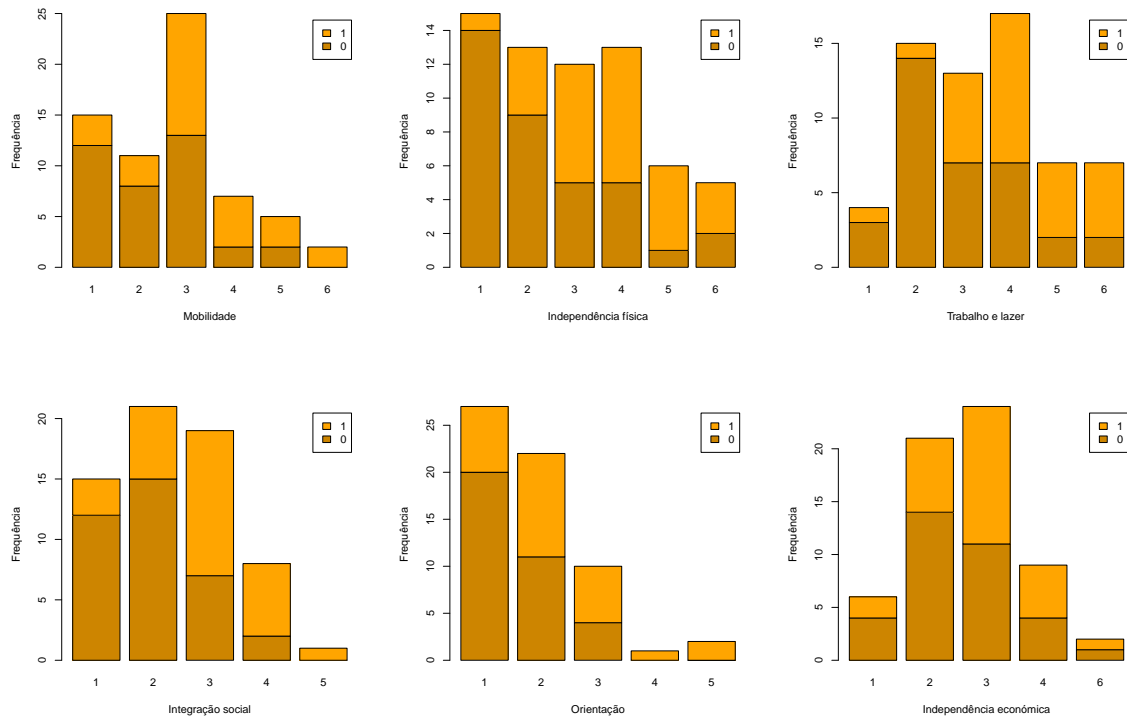


Figura 3.11: Gráficos de barras dos itens da escala de London em relação aos grupos de progressão

Os grupos estão mais ou menos equilibrados no que diz respeito aos seis factores da escala de London.

A título de resumo desta análise exploratória pode dizer-se que as variáveis relativas à idade do doente, à idade em que surgiram os primeiros sintomas, à demência, às características físicas (peso e altura), à cirurgia de Parkinson, à actividade física regular, aos factores internos (dislipidémia e ácido úrico elevado) e à Escala de London não se adivinham relevantes na distinção dos grupos, na medida em que não surgem diferenças significativas de progressão para estas características. As variáveis que dizem respeito a vícios e estado civil não devem ser consideradas na análise estatística que se segue, dado que apresentam poucas observações e pouca variabilidade. Por outro lado, existem variáveis que mostram diferenças entre os grupos: Estadio de Hohen e Yahr (menor estadio indica progressão mais lenta),

Escala de Schwab and England (menor percentagem, progressão rápida e maior percentagem, progressão lenta), MMSE (pontuação mais elevada indica progressão lenta) e Escala de depressão de Beck (menor pontuação indica progressão rápida).

Após este estudo exaustivo e descritivo que permitiu uma melhor compreensão das variáveis e das diferenças que estas apresentam nos respectivos grupos, vai fazer-se no, Capítulo 4, uma explicação dos métodos que se utilizaram para tratar os dados, e de seguida a análise dos resultados obtidos e das conclusões inerentes ao trabalho desenvolvido nesta dissertação.

CAPÍTULO 4

Métodos

Neste capítulo faz-se uma breve apresentação dos métodos utilizados no desenvolvimento deste trabalho. Juntamente com uma descrição sumária de cada método apresenta-se um exemplo ilustrativo para esclarecer o funcionamento do mesmo.

4.1 Análise de Componentes Principais

Nos dias que correm é muito comum a tarefa de explicar um conjunto de dados resultantes da medição de um conjunto de variáveis em cada um dos indivíduos de um dado grupo. E muitas vezes este conjunto de variáveis é muito grande em relação ao conjunto de indivíduos que se pretende estudar, pelo que é útil recorrer a técnicas estatísticas que permitam reduzir o número de variáveis sem que a informação global dos dados seja substancialmente perdida.

A Análise de Componentes Principais (ACP) é um dos métodos estatísticos mais utilizado na prática, com vista à redução do número de variáveis inicial (Jolliffe (2002)).

A ACP permite transformar o conjunto de variáveis originais correlacionadas entre si, num novo conjunto de variáveis que são combinações lineares das variáveis originais. A estas últimas que não estão correlacionadas, dá-se o nome de componentes principais. As novas variáveis são construídas de tal modo que é possível escolher um número mais reduzido de componentes principais que expliquem uma percentagem considerável da variabilidade dos dados.

A forma mais comum de obter as componentes principais (CPs) de um certo conjunto de dados é a seguinte: seja $\mathbf{X} = (X_1, \dots, X_p)^t$, o vector de p variáveis e seja Σ a matriz de covariâncias de \mathbf{X} , cujos valores próprios são $\lambda_1 \geq \dots \geq \lambda_p$ e cujos vectores próprios são $\gamma_1, \dots, \gamma_p$, respectivamente. A i -ésima componente principal é dada por $\gamma_i^t \mathbf{X}$.

Como $Var(X_i) = \lambda_i$, a variância total das p variáveis iniciais é $\sum_{i=1}^p Var(X_i) = \sum_{i=1}^p \lambda_i$, são usadas as relações (4.1) para determinar o número k de CPs a reter, sendo k tal que (4.1) corresponde a um valor elevado (em geral superior a 0,80).

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (4.1)$$

Se em vez do vector de variáveis \mathbf{X} se considerar uma matriz de dados, como acontece de facto na prática, pode aplicar-se o procedimento já descrito para obter CPs amostrais. Sendo neste caso a decomposição espectral feita a partir da matriz de covariâncias amostral \mathbf{S} . Obtém-se então $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_p$ e $\hat{\gamma}_1, \dots, \hat{\gamma}_p$. E a i -ésima componente principal é dada por $\hat{\gamma}_i^t \mathbf{X}$. A escolha do número k de CPs a reter baseia-se na relação (4.2)

$$\frac{\sum_{i=1}^k \hat{\lambda}_i}{\sum_{i=1}^p \hat{\lambda}_i} \quad (4.2)$$

De notar ainda que a obtenção das CPs pode igualmente fazer-se recorrendo tanto à matriz de correlações da população como à matriz de correlações amostrais.

Neste trabalho, dado o volume e natureza das variáveis, recorreu-se a esta análise usando o comando *PcaNA* do R (Serneels and Verdonck (2008) e Todorov and Filzmoser (2009)) que, com alguma fiabilidade, faz o cálculo de dados omissos permitindo assim a completação da matriz de dados. Desta forma é possível considerar mais observações no cálculo das componentes principais (com recurso ao comando *prcomp* do R (Mardia *et al.* (1979) e Becker *et al.* (1988))), o que permite que o estudo não perca relevância.

Exemplo

Considere-se um conjunto de dados sobre as detenções (por 100.000 habitantes) por assassinato, assalto e estupro, nos 50 estados dos EUA (McNeil (1977)), assim como a percentagem de população a residir nas áreas urbanas.

Na Tabela 4.1 apresenta-se a descrição das variáveis consideradas no estudo.

Tabela 4.1: Descrição das variáveis

Variáveis		
X_1	Murder	número de detenções por assassinato (em 100.000)
X_2	Assault	número de detenções por assalto (em 100.000)
X_3	Rape	número de detenções por estupro (em 100.000)
X_4	UrbanPop	percentagem de população urbana

Começa-se por calcular a matriz de covariâncias amostral, \mathbf{S} :

$$\mathbf{S} = \begin{pmatrix} 18.970 & 291.062 & 22.991 & 4.386 \\ 291.062 & 6945.317 & 519.269 & 312.275 \\ 22.991 & 519.269 & 87.729 & 55.768 \\ 4.386 & 312.275 & 55.768 & 209.519 \end{pmatrix}$$

Agora calculam-se os valores próprios e respectivos vectores próprios associados, para depois se calcularem as componentes principais. Obtendo-se assim $\hat{\lambda}_1 = 7011.115$, $\hat{\lambda}_2 = 201.992$, $\hat{\lambda}_3 = 42.113$ e $\hat{\lambda}_4 = 6.164$; e $\hat{\gamma}_1 = (-0.042, -0.995, -0.075, -0.046)^t$, $\hat{\gamma}_2 = (-0.045, -0.059, 0.201, 0.977)^t$, $\hat{\gamma}_3 = (0.080, -0.068, 0.974, -0.201)^t$ e $\hat{\gamma}_4 = (0.995, -0.040, -0.072, 0.058)^t$.

Ora $\sum_{i=1}^4 \hat{\lambda}_i = 7261.384$, e calculando a relação (4.2) para $k=1$, conclui-se que a primeira componente explica 96.6% da variabilidade dos dados, pelo que se deve reter esta componente. A primeira componente é dada por $-0.042X_1 - 0.995X_2 - 0.075X_3 - 0.046X_4$ (e as outras podem ser obtidas de forma análoga).

Como a variável assalto (Assault) tem o maior coeficiente na primeira componente retida pode interpretar-se este facto dizendo que a primeira CP representa o número de assaltos praticados nos vários estados dos EUA.

4.2 Escala de Likert

A escala de Likert é umas das mais comuns escalas na medição de atitude (Likert (1932)). Uma escala de Likert é unidimensional, ordenada e é constituída por um conjunto de frases (itens) em relação às quais o sujeito manifesta o seu grau de concordância, que pode variar entre 'discordo totalmente' (nível 1) até 'concordo totalmente' (nível 5, 7 ou 11).

A atitude do sujeito respondente é medida com base na soma ou na média dos níveis de resposta dado a cada item, o que matematicamente nem sempre é o mais correcto. Em consequência surge a necessidade de se encontrarem métodos que permitam traduzir a atitude do sujeito de forma mais adequada às frases que este tem de classificar. Este assunto é desenvolvido na secção seguinte.

Exemplo

Na figura 4.1 pode ver-se um exemplo de uma escala de Likert, a escala de valor da saúde, proposta por Lau *et al.* (1986), que é uma escala de quatro itens que visa avaliar o valor que a pessoa atribui à saúde. Esta escala tem sete categorias, desde 1 (discordo completamente) até 7 (concordo completamente).

Responda, numa escala de 1 (discordo completamente) a 7 (concordo completamente), qual o seu grau de concordância com cada uma das afirmações seguintes:

Item	1 – Discordo completamente	2	3	4	5	6	7 – Concordo completamente
1 - Se não se tiver saúde não se tem nada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 – Existem muitas coisas com que me preocupo mais do que com a minha saúde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 – Uma boa saúde é apenas um pormenor de menor importância numa vida feliz	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 – Não há nada mais importante do que ter uma boa saúde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figura 4.1: Escala de valor da saúde

4.3 Modelo Rasch

As escalas em geral, e as escalas de atitude em particular têm as suas limitações no que diz respeito às suas capacidades métricas, e por conseguinte é importante que se simule resultados de modo a averiguar a consistência das medições e a análise dos níveis de medição. A forma mais comum para fazer este tipo de simulações para escalas de Likert, são os modelos da *Item Response Theory (IRT)*. IRT são como que propostas de modelação de vários constructos, sejam eles atitudes, capacidades, aptidões permitindo a sua interpretação e integração na sua multidimensionalidade. O Modelo Rasch é uma particularização (mais simplificada) do modelo IRT (Kelderman and Rijkes (1994)).

Nos anos 60, os trabalhos de Georg Rasch, um matemático dinamarquês vieram trazer um novo fôlego à psicometria que começava a ficar estagnada ao mesmo tempo que era atacada

em vários dos seus procedimentos (Rasch (1960)). A verdadeira evolução aconteceu de facto quando Andrich aplicou o chamado modelo de Rasch à análise das medidas da escala de Likert, e outras semelhantes (Keats (1990)).

O modelo de Rasch proporia então novos métodos para o desenvolvimento e análise de testes, onde os itens e testes já não dependeriam da amostra da população e a habilidade do sujeito seria independente do conjunto de itens a que estes responderiam. Com efeito, este modelo é um método de construção de escalas em que o conhecimento de cada indivíduo é avaliado independentemente do conjunto de itens que as integram.

O objectivo deste modelo é portanto descrever a relação matemática entre a característica latente (a variável/ atitude que se quer medir com a escala) de um indivíduo e a sua performance num item.

Dito isto, temos as seguintes propriedades do modelo:

- (i) **unidimensionalidade**, o conjunto de itens mede apenas uma dimensão, uma variável latente;
- (ii) **independência local dos itens**, nenhum item deve conter informação que possa ser usada para responder a outro item;
- (iii) **monotonia**, quanto maior a habilidade do sujeito maior a probabilidade de responder correctamente ao item;
- (iv) **suficiência**, para estimar a dificuldade do item não é necessário saber que sujeitos responderam correctamente e para estimar a habilidade do sujeito não é necessário saber que itens foram respondidos.

A formulação do modelo Rasch assenta na seguinte equação, que tem por base uma função logística:

$$P[X_{vi} = 1] = c_i + \frac{(1 - c_i) \exp(\alpha_i(\theta_v - \sigma_i))}{1 + \exp(\alpha_i(\theta_v - \sigma_i))} \quad (4.3)$$

onde $P[X_{vi} = 1]$ representa a probabilidade do sujeito v responder correctamente ao item i , c_i é o parâmetro de aleatoriedade (probabilidade do sujeito, independente da sua capacidade, acertar o item i), α_i é o grau de discriminação do item i , θ_v é a capacidade do sujeito v e σ_i é o parâmetro de dificuldade do item i (corresponde à mediana da distribuição logística subjacente ao modelo).

Originalmente o modelo de Rasch foi desenvolvido para ser aplicado a dados binários (modelo dicotômico) sendo posteriormente alargado o seu uso a variáveis politômicas, como a seguir se descreve.

4.3.1 Modelo de Rasch politômico

O modelo de Rasch politômico deve ser usado quando para o item há mais do que duas respostas, sendo por exemplo de 0 (discordo totalmente) até m (concordo totalmente). De notar, que os valores $0, \dots, m$ são designados de categorias, no âmbito dos modelos Rasch (Fischer and Molenaar (1994)).

De (4.3), considera-se que todos os itens têm o mesmo poder de discriminação ($\alpha_i = 1, \forall i$), bem como o mesmo parâmetro de aleatoriedade ($c_i = 0, \forall i$) e generaliza-se para qualquer resposta (x). Donde, a probabilidade do sujeito v dar a resposta correcta no item i depende apenas da capacidade do sujeito e da barreira entre as categorias e é dada por:

$$P[X_{vi} = x] = \frac{\exp(x\theta_v - \sum_{s=1}^x \tau_{is})}{\eta_{vi}}, x = 0, \dots, m \quad (4.4)$$

Sendo $\eta_{vi} = \sum_{y=0}^m \exp(y\theta_v - \sum_{s=1}^y \tau_{is})$ e convencionou-se $\sum_{s=1}^0 \tau_{is} = 0$ dado que não há barreira para a categoria 0. E onde, θ_v é a capacidade do sujeito, τ_{ix} é o parâmetro de transição ou barreira do item i que separa a resposta na categoria $x-1$ para a categoria x .

Neste modelo há portanto $nm+1$ parâmetros a estimar: θ_v e os nm τ_{ix} ($i=1, \dots, n$ e $x=1, \dots, m$).

Exemplo

Para este exemplo vai considerar-se um conjunto genérico em que existem 20 observações em relação a 7 itens, cujas categorias variam de 0 a 3.

Na Tabela 4.2 mostram-se os τ_{is} (parâmetro de transição do item i da categoria $s-1$ para a categoria s) estimados. E na Tabela 4.3 podem ver-se os parâmetros de habilidade dos sujeitos, θ_v .

Pela observação das tabelas pode dizer-se, por exemplo, que para o item 5 a diferença entre a categoria 0 e a categoria 1 é maior (0.980) que a diferença entre a categoria 2 e a categoria 1 (-0.096). Além disso, o sujeito 7 é o que tem mais dificuldade em responder aos itens, dado que é o que tem menor habilidade (-2.190) enquanto que o sujeito 17 é o que tem menos dificuldade em responder aos itens, pois é o que tem maior habilidade (0.828).

Tabela 4.2: Tabela das estimativas dos τ_{is}

	Estimativa		Estimativa		Estimativa
τ_{12}	-0.937	τ_{33}	0.878	τ_{53}	-0.403
τ_{13}	0.330	τ_{41}	-0.423	τ_{61}	0.586
τ_{21}	-0.251	τ_{42}	0.263	τ_{62}	-0.504
τ_{22}	-0.065	τ_{43}	-0.140	τ_{63}	-1.100
τ_{23}	1.432	τ_{51}	0.980	τ_{71}	-0.849
τ_{31}	0.224	τ_{52}	-0.096	τ_{72}	0.298
τ_{32}	0.074				

Tabela 4.3: Tabela das estimativas dos θ_v

	Estimativas		Estimativas
θ_1	-0.450	θ_{11}	-0.165
θ_2	-0.165	θ_{12}	-0.270
θ_3	-0.773	θ_{13}	0.617
θ_4	-1.220	θ_{14}	0.119
θ_5	-0.450	θ_{15}	-0.025
θ_6	-0.450	θ_{16}	0.119
θ_7	-2.190	θ_{17}	0.828
θ_8	-0.305	θ_{18}	0.119
θ_9	-2.025	θ_{19}	-0.165
θ_{10}	-0.165	θ_{20}	0.305

4.4 Análise Discriminante

A análise discriminante parte do conhecimento de que n indivíduos observados em p variáveis pertencem a subgrupos distintos e procura determinar funções das p variáveis que melhor permitam distinguir/discriminar os indivíduos nesses grupos.

O objectivo principal desta análise consiste em encontrar funções, para a maioria dos métodos combinações lineares, das p variáveis observadas que melhor separem os indivíduos em causa. Através destas funções das variáveis será depois possível prever o grupo a que um indivíduo de origem desconhecida irá pertencer (McLachlan (2004)).

Uma das técnicas mais comuns de realizar uma análise discriminante é usar regressão logística, que pode ser regressão logística simples (se envolve apenas uma variável explicativa) ou regressão logística múltipla (quando estão presentes duas ou mais variáveis explicativas). Descreve-se esta última na secção que se segue.

4.4.1 Regressão Logística Múltipla

O modelo de Regressão Logística tem por objectivo descrever a relação entre uma variável resposta discreta e as variáveis explicativas numéricas (contínuas/discretas) e/ou categóricas (Hosmer and Lemeshow (2000)).

Considere-se a variável resposta Y binária (assume os valores 0 ou 1, com probabilidades π e $1 - \pi$, respectivamente), e $\mathbf{X} = X_1, \dots, X_{p-1}$ o vector de $(p-1)$ variáveis explicativas. O modelo de regressão logística múltipla é descrito pela seguinte equação:

$$E[Y|\mathbf{X}] = \pi(\mathbf{X}) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1})}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1})} \quad (4.5)$$

onde $\beta = (\beta_0, \beta_1, \dots, \beta_{p-1})^t$ é o vector de parâmetros.

À expressão (4.5) vai aplicar-se uma transformação muito importante nos modelos de regressão logística, a transformação *logit*, $\ln \frac{\pi}{1 - \pi}$, que conduz a uma nova expressão dos parâmetros, ou seja, a uma regressão linear múltipla:

$$\pi^* = \ln\left[\frac{\pi}{1 - \pi}\right] = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1} \quad (4.6)$$

Nestas condições recorre-se ao método de máxima verosimilhança para a estimação dos parâmetros, o que consiste em maximizar a função de máxima verosimilhança (4.7) que caracteriza os dados:

$$L(\beta|(x, y)) = \prod_{i=1}^n f_i(y_i) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (4.7)$$

onde, $\pi(x_i) = P(Y = 1|X = x_i)$ e $1 - \pi(x_i) = P(Y = 0|X = x_i)$.

Obtendo-se assim as estimativas pretendidas $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_{p-1})^t$.

Exemplo

Considere-se um conjunto de dados sobre o Factor de Risco Coronário obtidos num estudo de Rousseauw *et al.* (1983). Os dados representam observações em 462 homens brancos com

idades compreendidas entre os 15 e os 64 anos. O objectivo é determinar o risco de doença coronária, sendo a variável resposta a presença ou ausência de enfarte do miocárdio (chd). Na Tabela 4.4 apresenta-se a descrição das variáveis consideradas no estudo.

Tabela 4.4: Descrição das variáveis

Variáveis	
SBP	Pressão arterial sistólica, em mm de Hg
tabaco	Tabaco acumulado, em Kg
LDL	Lipoprotéina de baixa densidade colesterol, em mmol/L
adiposidade	Índice de adiposidade corporal
famhist	história familiar (0-ausente, 1-presente)
typeA	Resultado no teste de personalidade A
obesidade	Índice de massa corporal
alcool	Consumo actual de alcool
idade	Idade do paciente, em anos
chd	Doença cardíaca coronária (0-ausente, 1-presente)

Em primeiro lugar, determina-se o vector com as estimativas dos parâmetros $\hat{\beta}$, obtendo-se os valores que se mostram na Tabela 5.6.

Tabela 4.5: Estimativas dos parâmetros

i	Estimativa (β_i)
0	-0.508
1 sbp	0.001
2 tabaco	0.017
3 ldl	0.033
4 adiposidade	0.002
5 famhist	0.173
6 tipoA	0.006
7 obesidade	-0.011
8 alcool	-0.0002
9 idade	0.007

Pode-se observar que as variáveis que têm mais peso na doença cardíaca coronária são o

histórico familiar de doença e os valores de lipoproteína de baixa densidade colesterol elevados.

De seguida, faz-se a classificação dos indivíduos de acordo com os critérios já referidos, obtendo-se assim a matriz de classificação, apresentada na 4.6. Verifica-se assim que este método classifica correctamente 74.46% dos doentes.

Tabela 4.6: Matriz de Classificação

		Valores preditos	
		0	1
Valores verdadeiros	0	260	42
	1	76	84

Depois da explicação de todos os métodos utilizados no desenvolvimento deste trabalho, apresentam-se, no capítulo 5, os resultados obtidos da aplicação de cada um dos métodos.

CAPÍTULO 5

Apresentação de Resultados e Conclusões

Neste capítulo explica-se todo o processo de tratamento dos dados e apresentam-se os resultados obtidos na análise de componentes principais e na regressão logística múltipla.

5.1 Selecção de variáveis

Em primeiro lugar, e tendo em conta que existem 204 variáveis, e apenas 78 observações de doentes, fez-se uma “limpeza” das variáveis, eliminam-se aquelas que apresentavam um número muito reduzido de observações, que apresentavam pouca variabilidade, ou cujo interesse era demográfico e/ou de afectação no momento do diagnóstico. Na lista seguinte mostra-se a lista de variáveis eliminadas e o respectivo motivo:

- **Interesse demográfico e/ou afectação no momento do diagnóstico:** InclusionDate, UPDRSSource, UPDRSPacienteClinicalState, MinutesSinceLevodopa, DyskinesiaIIA, DyskinesiaIIB;
- **Poucas observações:** FamilialPDrelation, LimbTremorOnSetYear, PosturalInstOnSetYear, FallsOnSetYear, MotorFluctuationsOnSetYear, DiskynesiasOnSetYear, DystoniaOnSetYear, FreezingOnSetYear, MotorSleepBenefitOnSetYear, DysphagiaOnSetYear, DysarthriaOnSetYear, DementiaOnSetYear, MotiveRetirementorUnempliment, PreviousJobs, SmokingUntilAge, NumberYearsSmoking, NumberCigarettesDay, AgeOn-

setSmoking, PassiveSmokerHome, PassiveSmokerWork, EndAlcoholicHabitsAge, NumberYearsAlcoholicHabits, GlassesAlcoholDay;

- **Pouca variabilidade:** UPDRSPatientOnLevodopa, Race, Ethnicity, Religion, Diet.

Desta forma, houve uma redução de 204 para 170 variáveis. O que ainda assim é um número muito elevado em relação ao número de observações existentes.

Outra forma para reduzir o número de variáveis que se explorou foi a de calcular a correlação entre as 170 variáveis, de modo a perceber se existiam variáveis que forneciam informação muito semelhante. No entanto, não se encontraram variáveis com correlações muito elevadas, o que pode dever-se ao facto de a natureza das variáveis ser muito distinta. E como tal, não se usaram os valores de correlação como critério para remoção de qualquer variável.

Na secção que se segue, mostra-se o resultado obtido da aplicação do Modelo de Rasch politómico a grupos de variáveis que constituem escalas tipo de Likert, permitindo assim mais uma redução no número de variáveis.

5.2 Modelo de Rasch politómico

Como já explicado no capítulo anterior, o modelo Rasch permite encontrar a relação matemática entre a característica que se quer medir e a performance do sujeito.

Neste estudo, consideram-se as escalas UPDRS e de London. A escala UPDRS procura avaliar o estado do doente em quatro vertentes: actividade mental, comportamento e humor; actividades da vida diária; motilidade; complicações do tratamento, pelo que se aplica este método a cada uma destas vertentes, ficando assim com UPDRS1, UPDRS2, UPDRS3 e UPDRS4, respectivamente. A escala de London procura avaliar a incapacidade do doente imposta pela doença face a actividades diárias e é constituída por seis itens, aos quais será aplicado o modelo Rasch, ficando assim apenas a variável London.

A escala UPDRS, e cada uma das suas vertentes varia de 0 (normal) a 4 (severo), daí a necessidade de aplicar o Modelo de Rasch politómico, assim como a Escala de London que varia de 1 (de forma alguma) a 6 (completamente).

Assim, aplicou-se o comando *PCM* do pacote *eRm* (Mair and Hatzinger (2007)) do R, tendo-se obtido os resultados que podem ser consultados no apêndice A(A.1).

Usando estes resultados, obtiveram-se três novos conjuntos de dados, resultantes de três transformações diferentes operadas nos diferentes itens que permitem a obtenção de UPDRS1, UPDRS2, UPDRS3, UPDRS4 e London.

Em (5.1) pode ver-se a primeira transformação.

$$UPDRS(k)_v = \theta_v \times \sum_{i=1}^n UPDRS(k)_i \quad (5.1)$$

onde $UPDRS(k)_v$ se refere ao valor da UPDRSk ($k=1,2,3,4$) para o sujeito v e $UPDRS(k)_i$ se refere à observação do item i (existem $1, \dots, n$ itens) da vertente k da UPDRS.

Em (5.2) pode ver-se a segunda transformação.

$$UPDRS(k)_v = \sum_{i=1}^n \theta_{ix} \times \chi_x(UPDRS(k)_i) \times UPDRS(k)_i \quad (5.2)$$

sendo χ_x a função identidade da categoria x .

A terceira transformação consistiu na combinação das duas primeiras, e pode ver-se em (5.3).

$$UPDRS(k)_v = \theta_v \times \sum_{i=1}^n \theta_{ix} \times \chi_x(UPDRS(k)_i) \times UPDRS(k)_i \quad (5.3)$$

O processo aqui descrito foi usado de forma análoga para a escala de London.

Para cada uma destas modelações da escalas UPDRS e de London, são obtidos diferentes grupos de dados, como já foi referido. Estes diferentes conjuntos foram sujeitos a análise de componentes principais e regressão logística, para se fazer a selecção do melhor modelo, questão mais desenvolvida na secção que se segue.

5.3 Seleccção do modelo

Na construção do modelo estatístico que permite fazer a classificação do doente num dos grupos de progressão teve-se em conta duas questões, nomeadamente: i) encontrar o menor número de variáveis que melhor expliquem a variabilidade dos dados, e ii) também o menor número de variáveis que traduz um modelo interpretável e numericamente estável.

Assim sendo, neste trabalho, as técnicas de análise de componentes principais e de regressão foram aplicados a diferentes agrupamentos dos dados. Em primeiro lugar fez-se uma filtragem dos dados eliminando-se todas as variáveis que se referiam ao momento do diagnóstico (ficando 39 variáveis). Em segundo lugar, traduziram-se as perguntas de cada vertente da UPDRS, na respectiva média da vertente, e mantiveram-se todas as outras variáveis, ficando um total de 108 variáveis. Depois, usaram-se os resultados obtidos com o Modelo Rasch para as escalas UPDRS e London, do qual resultaram os três conjuntos já explicados na secção 5.2, havendo em cada um destes um total de 103 variáveis.

Neste processo aquele que produziu melhores resultados, do ponto de vista de classificação correcta dos doentes e estabilidade numérica, foi aquele em que as quatro vertentes da escala UPDRS e a escala de London resultaram do uso da habilidade do sujeito e dos parâmetros de dificuldade dos itens.

Então, pegando neste caso fez-se um refinamento das variáveis (o que se mostra na secção seguinte) de modo a melhorar os resultados obtidos.

5.4 Análise de componentes principais

Usando o conjunto de dados obtido na secção anterior, analisaram-se as variáveis e fizeram-se 5 agrupamentos: ESC (constituído por todas as variáveis que representam escalas), NF (constituído por variáveis que representam factores não físicos), SINT (constituído por variáveis que representam sintomas da doença), F (constituído pelos restantes factores físicos) e RESTO (as restantes variáveis).

Para cada um destes subgrupos foram calculadas as componentes principais, das quais foram retidas as primeiras componentes que no seu todo explicavam pelo menos 60% da variabilidade dos dados. Os resultados obtidos podem ser consultados na Tabela 5.1.

Como se pode ver, pelas tabelas, houve uma redução significativa das variáveis, ficando apenas 7 componentes principais, em termos significativos 18 das variáveis originais, o que facilita muito o processo interpretativo dos resultados. No entanto, é ainda necessário perceber quais destas componentes têm efectivamente um contributo importante na discriminação dos doentes nos grupos de progressão lenta e rápida da doença de Parkinson. Na secção que se segue faz-se esta análise através de uma regressão logística múltipla.

Tabela 5.1: Tabelas das componentes principais de cada um dos subgrupos (ESC, NF, SINT, F, RESTO)

	ESC			SINT	
	ESC1	ESC2		SINT1	SINT2
UPDRS1	-	-0.1	FreezingEver	-0.5	0.4
UPDRS2	-	0.8	MotorSleepBenefitEver	-0.1	-0.8
UPDRS3	1	-	Dysphagia	-0.7	0.2
SEscale	0.1	-0.6	Dysarthria	-0.6	-0.4
MMSE	-	-0.1			
BDItotal	-	0.1			
London	0.1	-0.1			

	F		NF
	F1		NF1
Weight	0.9	AgeRetirementorUnempliement	1
Height	0.5		

	RESTO
	RESTO1
DateBirth	-0.4
DateFirstMotorSymptoms	-0.7
DateDiagnosis	-0.6

5.5 Regressão logística múltipla

Nesta secção faz-se finalmente a regressão que permite perceber as variáveis que contribuem para a distinção entre os dois grupos de progressão da doença.

Recorde-se que a variável progressão é a variável resposta (Y) e é definida da seguinte forma:

$$Y = \begin{cases} 0, & \text{a doença progride lentamente} \\ 1, & \text{a doença progride rapidamente} \end{cases} \quad (5.4)$$

Na amostra de dados, a variável resposta Y , apresenta frequências que constam na Tabela 5.2.

Tabela 5.2: Frequências da progressão (variável resposta)

	Frequência	Percentagem
Progressão Lenta	43	55%
Progressão Rápida	35	45%
Total	78	100%

Assim, considerou-se o Modelo 1, aquele em que se consideram como explicativas todas as componentes principais obtidas na secção anterior (ESC1, ESC2, NF1, SINT1, SINT2, F1, RESTO1) para a regressão logística, com recurso ao comando *lm* do R (Chambers (1992)), tendo-se obtido os resultados representados na Tabela 5.3.

Tabela 5.3: Estimativas dos parâmetros, estimativas dos erros padrão e valor-p para o Modelo 1

	Estimativa	Erro Padrão	Valor-p
Ordenada na origem	2.7768	0.7922	0.0008
ESC1	-0.0014	0.0015	0.3437
ESC2	0.0085	0.0025	0.0009
NF1	-0.0028	0.0057	0.6195
SINT1	-0.0111	0.0908	0.9029
SINT2	-0.2096	0.1182	0.0807
F1	-0.0111	0.0048	0.0236
RESTO1	0.0035	0.0040	0.3813

Da observação desta tabela pode ver-se que as componentes com menor valor-p são ESC2, SINT2 e F1, sendo portanto estas as mais significativas para o modelo completo (Modelo 1). A componente F1 é constituída pelas variáveis peso (Weight) e altura (Height), mas como existem alguns dados omissos referentes a estas variáveis, optou-se por não as considerar. Assim sendo, na Tabela 5.4, mostram-se os resultados para o Modelo 2, em que se consideram as variáveis ESC2 e SINT2.

Tendo os dois modelos é importante compará-los e escolher aquele que melhor se adequa aos dados e que produz menos incerteza.

Tabela 5.4: Estimativas dos parâmetros, estimativas dos erros padrão e valor-p para o Modelo 2

	Estimativa	Erro Padrão	Valor-p
Ordenada na origem	0.7979	0.1040	4.98e-11
ESC2	0.0088	0.0022	0.0002
SINT2	-0.1844	0.1202	0.1293

Note-se que a classificação dos indivíduos cujas observações foram utilizadas para estimar o modelo de regressão produz estimativas do erro de classificação (taxas de erro aparente) que são enviesadas no sentido otimista, pelo que é importante usar um procedimento que forneça resultados mais realistas. Optou-se neste caso pela validação cruzada. Este processo consiste em estimar o modelo de regressão com todas as observações menos uma, sendo essa apenas utilizadas para ser classificada. Este processo é repetido de modo a que cada observação na amostra é usada uma vez na validação dos dados.

Na Tabela 5.5 apresentam-se as estimativas de erro de classificação calculadas a partir das taxas aparente e de validação cruzada para ambos os modelos. Estas taxas foram obtidas a partir das matrizes de classificação de cada um dos modelos que se apresentam a seguir.

Tabela 5.5: Estimativas de erro de predição para taxa aparente e validação cruzada

	Taxa aparente	Validação cruzada
Modelo 1	0.255	0.283
Modelo 2	0.27	0.282

O Modelo 2 é o que tem menor estimativa do erro de classificação com base na validação cruzada, e é também aquele que apresenta um menor *over-fitting* (avaliado pela diferença entre os dois tipos de estimativa), pelo que se considera ser este o melhor modelo, e será este o escolhido para efeitos de interpretação. Na Tabela 5.6 apresentam-se das estimativas dos parâmetros do modelo, já considerando as variáveis originais que constituem as componentes ESC2 e SINT2.

Então pode dizer-se que os sintomas que o doente refere ter relativamente a sentir-se melhor após o sono (MotorSleepBenefitEver) e disartria (Dysarthria) são indicadores a considerar no sentido de progressão rápida da doença, enquanto que os sintomas de freezing e disfagia (Dysphagia) são indicadores de progressão lenta. Por outro lado, em relação às escalas de avaliação do estado do doente, como já seria de esperar os doentes mais independentes

Tabela 5.6: Estimativas dos parâmetros associados às variáveis originais

	Estimativa
Ordenada na origem	0.798
UPDRS1	-0.001
UPDRS2	0.006
SEscale	-0.005
MMSE	-0.001
BDItotal	0.001
London	-0.001
FreezingEver	-0.086
MotorSleepBenefitEver	0.172
DysphagiaEver	-0.043
DysarthriaEver	0.086

nas actividades diárias (SEscale) e com melhor estado mental (MMSE) tendem a ter uma progressão lenta da doença, enquanto que os doentes com maior nível de depressão (BDI) tendem a pertencer ao grupo de doentes em que a doença progride mais rapidamente.

É também importante consultar as matrizes de classificação dos doentes no sentido de perceber a sensibilidade do método (tabelas 5.7 e 5.8).

Tabela 5.7: Matriz de classificação para taxa aparente

		Grupo predito	
		0(lenta)	1(rápida)
Grupo Real	0(lenta)	37	6
	1(rápida)	15	20

Como se pode observar o método parece ter maior facilidade a classificar os doentes com progressão lenta da doença do que de progressão rápida.

Tabela 5.8: Matriz de classificação para validação cruzada

		Grupo predito	
		0(lenta)	1(rápida)
Grupo Real	0(lenta)	36	7
	1(rápida)	15	20

Na secção que se segue são apresentadas as conclusões deste trabalho.

5.6 Conclusões

Nesta secção apresentam-se as conclusões do trabalho realizado.

Nesta dissertação fez-se inicialmente uma análise descritiva e exploratória dos dados, onde se procuraram observar as diferenças existentes nos grupos de progressão da doença para as variáveis em estudo. Desta análise, anteviu-se que o total da escala de Beck, que mede o índice de depressão do doente seria um indicativo para fazer a discriminação entre os dois grupos de progressão.

Em seguida, experimentaram-se diversas formas de modelar a escala UPDRS e a escala de London, que são escalas de Likert, usando os modelos Rasch, concebidos para a construção de escalas de atitudes, de modo a fazer uma redução das variáveis (imperativa neste conjunto de dados), mas principalmente a encontrar uma forma mais científica de codificar como numéricas, variáveis qualitativas.

Depois de experimentadas diferentes modelações das variáveis qualitativas, e tendo-se verificado bons resultados na modelação das escalas pelos modelos Rasch, encontrou-se o conjunto que permitiu apurar melhores resultados na regressão logística, de modo a encontrar um modelo que não só explicasse a variabilidade dos dados, mas também que fosse interpretável no contexto do problema.

Este estudo envolveu um elevado número de variáveis, que inevitavelmente teve que ser reduzido, dado existirem apenas 78 observações. E como tal, recorreu-se à análise de componentes principais, que permite obter novas variáveis (as componentes principais) que são combinações lineares das variáveis originais, e que explicam 60% da variabilidade dos dados. Sendo que esta análise foi feita em cinco subgrupos de variáveis, organizados de acordo com

o significado destas, para que se possa interpretar mais facilmente, e para que os resultados obtidos tenham um sentido biológico.

A regressão logística múltipla e dicotómica é considerada um método adequado na determinação de relações entre uma variável dicotómica que se pretende explicar, e as respectivas variáveis explicativas (qualitativas e/ou quantitativas). Após o cálculo das componentes principais, fez-se esta regressão, resultando que um grupo de escalas (UPDRS, MMSE, BDI, de Schwab and England, e de London) e de sintomas (benefício após o sono, disfagia, disartria e freezing) seriam os principais factores clínicos que permitem distinguir a progressão da doença de Parkinson.

Finalmente, espera-se que este trabalho tenha dado mais um contributo na compreensão desta doença neurodegenerativa, principalmente no que à distinção dos grupos de progressão diz respeito, e que desta forma se possa antever mais precocemente a que grupo de progressão o doente vai pertecer, e assim encontrar a terapêutica que melhor se lhe adequa.

Outra questão importante, e que introduziu um factor importante no desenvolvimento deste trabalho é o uso dos modelos de Rasch na modelação de escalas qualitativas, e até que ponto será este o mais adequado, ficando em aberto a ideia de se aplicarem outros métodos, porventura de mais simples entendimento e com melhores resultados.

Em termos de classificação seria interessante definir outros sugrupos de variáveis para análise de componentes principais, e aplicar outros métodos de classificação para comparar com a regressão logística.

APÊNDICE A

Macros em R

A.1 Modelos Rasch

Este apêndice contém o código R, implementado no *software* estatístico R, relativo aos procedimentos referidos ao longo deste estudo. De notar que o código não se encontra comentado, assumindo-se que o leitor está familiarizado com os comandos básicos do *software* em causa.

A.1.1 Resultados para a UPDRS1 (sintomas não motores)

Parâmetros de dificuldade do item-categoria

Results of PCM estimation:

Call: PCM(X = updrs1.full)

Conditional log-likelihood: -768.8401

Number of iterations: 56

Number of parameters: 43

Item (Category) Difficulty Parameters (eta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
UPDRS1.1.c2	-0.794	0.338	-1.456	-0.132
UPDRS1.1.c3	-0.458	0.444	-1.329	0.412
UPDRS1.1.c4	-0.417	0.519	-1.434	0.600
UPDRS1.2.c1	2.020	0.516	1.009	3.032

UPDRS1.2.c2	1.891	0.594	0.727	3.055
UPDRS1.2.c3	1.329	0.543	0.264	2.394
UPDRS1.3.c1	-0.916	0.326	-1.555	-0.277
UPDRS1.3.c2	-0.847	0.393	-1.618	-0.076
UPDRS1.3.c3	-1.589	0.379	-2.332	-0.846
UPDRS1.4.c1	-0.330	0.318	-0.952	0.293
UPDRS1.4.c2	-0.675	0.347	-1.356	0.005
UPDRS1.4.c3	-0.628	0.415	-1.440	0.185
UPDRS1.5.c1	0.256	0.317	-0.367	0.878
UPDRS1.5.c2	0.005	0.355	-0.691	0.702
UPDRS1.5.c3	0.669	0.539	-0.389	1.726
UPDRS1.5.c4	1.836	1.005	-0.133	3.806
UPDRS1.6.c1	2.851	0.714	1.451	4.250
UPDRS1.6.c2	3.162	0.996	1.209	5.115
UPDRS1.6.c3	2.230	0.732	0.794	3.665
UPDRS1.7.c1	-0.415	0.332	-1.065	0.236
UPDRS1.7.c2	-1.141	0.328	-1.784	-0.498
UPDRS1.7.c3	-0.683	0.435	-1.536	0.170
UPDRS1.8.c1	-0.574	0.343	-1.247	0.098
UPDRS1.8.c2	-1.582	0.320	-2.209	-0.954
UPDRS1.8.c3	-0.384	0.516	-1.395	0.627
UPDRS1.9.c1	-0.579	0.308	-1.182	0.024
UPDRS1.9.c2	-0.897	0.335	-1.555	-0.240
UPDRS1.9.c3	0.381	0.616	-0.826	1.587
UPDRS1.9.c4	0.119	0.632	-1.119	1.357
UPDRS1.10.c1	-0.156	0.335	-0.813	0.501
UPDRS1.10.c2	-0.568	0.361	-1.275	0.139
UPDRS1.10.c3	-0.905	0.386	-1.661	-0.149
UPDRS1.10.c4	0.103	0.626	-1.124	1.330
UPDRS1.11.c1	-0.383	0.313	-0.995	0.230
UPDRS1.11.c2	-0.666	0.347	-1.346	0.014
UPDRS1.11.c3	-0.495	0.430	-1.339	0.348
UPDRS1.12.c1	0.414	0.303	-0.181	1.008
UPDRS1.12.c2	1.067	0.478	0.129	2.005
UPDRS1.12.c3	1.277	0.611	0.079	2.475
UPDRS1.13.c1	-1.002	0.317	-1.624	-0.380
UPDRS1.13.c2	-1.005	0.371	-1.733	-0.277
UPDRS1.13.c3	-0.662	0.482	-1.606	0.283
UPDRS1.13.c4	-0.775	0.530	-1.813	0.263

Parâmetros de capacidade do sujeito

Estimation of Ability Parameters

Collapsed log-likelihood: -289.4472

Number of iterations: 11

Number of parameters: 21

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 1	-0.12223917	0.2540763	-0.62021964	0.375741304
theta 2	-0.39266907	0.2677160	-0.91738289	0.132044738
theta 3	-1.62606039	0.4717498	-2.55067298	-0.701447789
theta 4	-0.39266907	0.2677160	-0.91738289	0.132044738

theta 5 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 6 -1.26518278 0.3858679 -2.02146993 -0.508895626
theta 7 -0.12223917 0.2540763 -0.62021964 0.375741304
theta 8 -0.32222522 0.2632322 -0.83815093 0.193700480
theta 9 0.44655043 0.2558480 -0.05490244 0.948003306
theta 10 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 11 -0.54205782 0.2795349 -1.08993621 0.005820577
theta 12 0.13017369 0.2496532 -0.35913762 0.619484994
theta 13 0.25503808 0.2504213 -0.23577864 0.745854796
theta 14 -0.18738866 0.2564955 -0.69011062 0.315333310
theta 15 -1.12686551 0.3588958 -1.83028842 -0.423442598
theta 16 -0.32222522 0.2632322 -0.83815093 0.193700480
theta 17 -0.79856327 0.3076185 -1.40148435 -0.195642177
theta 18 -0.62231091 0.2872356 -1.18528238 -0.059339439
theta 19 -0.89729142 0.3211861 -1.52680456 -0.267778282
theta 20 -1.12686551 0.3588958 -1.83028842 -0.423442598
theta 21 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 22 -0.32222522 0.2632322 -0.83815093 0.193700480
theta 23 -0.62231091 0.2872356 -1.18528238 -0.059339439
theta 24 -1.42760997 0.4217130 -2.25415223 -0.601067713
theta 25 -0.05817817 0.2522126 -0.55250579 0.436149460
theta 26 -0.54205782 0.2795349 -1.08993621 0.005820577
theta 27 -1.88312205 0.5470723 -2.95536410 -0.810879997
theta 28 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 29 -0.05817817 0.2522126 -0.55250579 0.436149460
theta 30 0.58095040 0.2632069 0.06507427 1.096826528
theta 31 -0.79856327 0.3076185 -1.40148435 -0.195642177
theta 32 -0.54205782 0.2795349 -1.08993621 0.005820577
theta 33 -0.70742078 0.2964730 -1.28849718 -0.126344386
theta 34 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 35 -0.39266907 0.2677160 -0.91738289 0.132044738
theta 36 -0.18738866 0.2564955 -0.69011062 0.315333310
theta 37 -0.89729142 0.3211861 -1.52680456 -0.267778282
theta 38 -0.12223917 0.2540763 -0.62021964 0.375741304
theta 39 0.13017369 0.2496532 -0.35913762 0.619484994
theta 40 -0.32222522 0.2632322 -0.83815093 0.193700480
theta 41 -0.62231091 0.2872356 -1.18528238 -0.059339439
theta 42 -0.70742078 0.2964730 -1.28849718 -0.126344386
theta 43 -0.89729142 0.3211861 -1.52680456 -0.267778282
theta 44 -1.26518278 0.3858679 -2.02146993 -0.508895626
theta 45 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 46 -0.18738866 0.2564955 -0.69011062 0.315333310
theta 47 -0.54205782 0.2795349 -1.08993621 0.005820577
theta 48 -1.62606039 0.4717498 -2.55067298 -0.701447789
theta 49 -0.39266907 0.2677160 -0.91738289 0.132044738
theta 50 -0.62231091 0.2872356 -1.18528238 -0.059339439
theta 51 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 52 -0.89729142 0.3211861 -1.52680456 -0.267778282
theta 53 -1.00572184 0.3379115 -1.66801622 -0.343427461
theta 54 -0.79856327 0.3076185 -1.40148435 -0.195642177
theta 55 -1.12686551 0.3588958 -1.83028842 -0.423442598
theta 56 -0.18738866 0.2564955 -0.69011062 0.315333310
theta 57 -0.79856327 0.3076185 -1.40148435 -0.195642177
theta 58 -0.05817817 0.2522126 -0.55250579 0.436149460
theta 60 -1.62606039 0.4717498 -2.55067298 -0.701447789
theta 61 -0.12223917 0.2540763 -0.62021964 0.375741304
theta 62 -0.25393429 0.2595234 -0.76259085 0.254722277
theta 64 -0.18738866 0.2564955 -0.69011062 0.315333310

```
theta 65 -0.25393429 0.2595234 -0.76259085 0.254722277
theta 66 -1.26518278 0.3858679 -2.02146993 -0.508895626
```

A.1.2 Resultados para a UPDRS2 (atividades da vida diária)

Parâmetros de dificuldade do item-categoria

Item (Category) Difficulty Parameters (eta) with 0.95 CI:

Item (Category)	Estimate	Std. Error	lower CI	upper CI
UPDRS2.1.c2	-1.446	0.473	-2.372	-0.519
UPDRS2.1.c3	-0.641	0.516	-1.652	0.370
UPDRS2.1.c4	2.530	0.864	0.837	4.223
UPDRS2.2.c1	0.141	0.401	-0.644	0.927
UPDRS2.2.c2	0.132	0.415	-0.682	0.945
UPDRS2.2.c3	0.505	0.443	-0.363	1.374
UPDRS2.2.c4	4.535	1.079	2.420	6.651
UPDRS2.3.c1	0.449	0.311	-0.161	1.059
UPDRS2.3.c2	2.874	0.640	1.619	4.129
UPDRS2.3.c3	3.048	0.554	1.963	4.133
UPDRS2.4.c1	-0.609	0.363	-1.320	0.103
UPDRS2.4.c2	-0.059	0.420	-0.882	0.764
UPDRS2.4.c3	1.187	0.515	0.177	2.197
UPDRS2.5.c1	-1.476	0.428	-2.315	-0.636
UPDRS2.5.c2	-1.257	0.466	-2.170	-0.343
UPDRS2.5.c3	0.290	0.578	-0.843	1.423
UPDRS2.5.c4	1.609	0.662	0.311	2.908
UPDRS2.6.c1	-1.951	0.418	-2.770	-1.132
UPDRS2.6.c2	-0.252	0.552	-1.334	0.830
UPDRS2.6.c3	0.585	0.607	-0.606	1.775
UPDRS2.6.c4	1.650	0.663	0.351	2.950
UPDRS2.7.c1	-1.762	0.473	-2.689	-0.835
UPDRS2.7.c2	-1.579	0.517	-2.593	-0.566
UPDRS2.7.c3	-1.040	0.543	-2.105	0.024
UPDRS2.7.c4	0.839	0.668	-0.470	2.148
UPDRS2.8.c1	-1.110	0.413	-1.918	-0.301
UPDRS2.8.c2	-0.706	0.473	-1.634	0.222
UPDRS2.8.c3	0.120	0.540	-0.938	1.179
UPDRS2.8.c4	0.636	0.562	-0.467	1.738
UPDRS2.9.c1	-1.756	0.473	-2.683	-0.829
UPDRS2.9.c2	-1.695	0.509	-2.692	-0.697
UPDRS2.9.c3	-0.699	0.567	-1.811	0.413
UPDRS2.9.c4	0.541	0.638	-0.710	1.793
UPDRS2.10.c1	-0.950	0.371	-1.678	-0.223
UPDRS2.10.c2	-0.146	0.441	-1.009	0.718
UPDRS2.10.c3	1.013	0.526	-0.018	2.044
UPDRS2.10.c4	3.853	0.864	2.159	5.547
UPDRS2.11.c1	-2.552	0.557	-3.643	-1.461
UPDRS2.11.c2	-2.303	0.592	-3.463	-1.142
UPDRS2.11.c3	-0.929	0.668	-2.238	0.381
UPDRS2.11.c4	-0.146	0.688	-1.494	1.203
UPDRS2.12.c1	-2.394	0.508	-3.389	-1.399
UPDRS2.12.c2	-1.485	0.584	-2.629	-0.341
UPDRS2.12.c3	-1.127	0.592	-2.288	0.033
UPDRS2.12.c4	0.063	0.654	-1.218	1.345
UPDRS2.13.c1	-0.226	0.346	-0.905	0.452

UPDRS2.13.c2	1.056	0.500	0.076	2.037
UPDRS2.13.c3	1.149	0.481	0.206	2.092
UPDRS2.13.c4	2.769	0.625	1.544	3.993

Parâmetros da capacidade do sujeito

Estimation of Ability Parameters

Subject NA Group: 1

NA pattern: x x x x x x x x x x x x

Collapsed log-likelihood: -396.0235

Number of iterations: 15

Number of parameters: 28

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 1	1.70469703	0.3272368	1.06332463	2.34606942
theta 2	2.65141707	0.5008849	1.66970073	3.63313340
theta 3	-0.51280395	0.3528959	-1.20446721	0.17885930
theta 4	0.01573023	0.3019455	-0.57607200	0.60753247
theta 5	-0.64240613	0.3673100	-1.36232053	0.07750826
theta 6	-0.64240613	0.3673100	-1.36232053	0.07750826
theta 7	0.01573023	0.3019455	-0.57607200	0.60753247
theta 8	0.10482833	0.2951726	-0.47369934	0.68335599
theta 9	2.42535976	0.4518993	1.53965339	3.31106614
theta 10	0.10482833	0.2951726	-0.47369934	0.68335599
theta 11	0.19019941	0.2893170	-0.37685145	0.75725027
theta 12	1.60179636	0.3147243	0.98494808	2.21864464
theta 13	0.27242758	0.2843075	-0.28480493	0.82966008
theta 14	0.10482833	0.2951726	-0.47369934	0.68335599
theta 15	0.65293713	0.2702224	0.12331099	1.18256326
theta 16	1.50605321	0.3044549	0.90933263	2.10277378
theta 17	-0.17640460	0.3185890	-0.80082760	0.44801841
theta 18	1.60179636	0.3147243	0.98494808	2.21864464
theta 19	-1.10640766	0.4216099	-1.93274779	-0.28006754
theta 20	-0.28108538	0.3286634	-0.92525382	0.36308307
theta 21	-0.64240613	0.3673100	-1.36232053	0.07750826
theta 22	0.27242758	0.2843075	-0.28480493	0.82966008
theta 23	1.41599504	0.2960187	0.83580903	1.99618106
theta 24	-4.00263016	1.0497116	-6.06002712	-1.94523320
theta 25	1.33047221	0.2891074	0.76383212	1.89711230
theta 26	0.65293713	0.2702224	0.12331099	1.18256326
theta 28	-2.00644500	0.5375581	-3.06003958	-0.95285042
theta 29	0.19019941	0.2893170	-0.37685145	0.75725027
theta 30	0.35203100	0.2800840	-0.19692350	0.90098549
theta 31	-4.00263016	1.0497116	-6.06002712	-1.94523320
theta 32	0.57954616	0.2716893	0.04704490	1.11204743
theta 33	1.01764514	0.2727550	0.48305520	1.55223508
theta 34	-1.50251496	0.4701266	-2.42394617	-0.58108375
theta 35	0.01573023	0.3019455	-0.57607200	0.60753247
theta 36	0.19019941	0.2893170	-0.37685145	0.75725027
theta 37	-0.39282316	0.3400576	-1.05932379	0.27367747
theta 38	0.65293713	0.2702224	0.12331099	1.18256326
theta 39	1.01764514	0.2727550	0.48305520	1.55223508
theta 40	0.35203100	0.2800840	-0.19692350	0.90098549

```
theta 42 0.01573023 0.3019455 -0.57607200 0.60753247
theta 43 0.27242758 0.2843075 -0.28480493 0.82966008
theta 44 -1.73765885 0.5004374 -2.71849805 -0.75681965
theta 45 -0.51280395 0.3528959 -1.20446721 0.17885930
theta 46 1.70469703 0.3272368 1.06332463 2.34606942
theta 47 0.10482833 0.2951726 -0.47369934 0.68335599
theta 48 -0.51280395 0.3528959 -1.20446721 0.17885930
theta 49 0.42947466 0.2765969 -0.11264536 0.97159468
theta 50 0.35203100 0.2800840 -0.19692350 0.90098549
theta 51 0.50518208 0.2738085 -0.03147263 1.04183678
theta 52 -2.00644500 0.5375581 -3.06003958 -0.95285042
theta 54 0.10482833 0.2951726 -0.47369934 0.68335599
theta 55 0.27242758 0.2843075 -0.28480493 0.82966008
theta 56 0.27242758 0.2843075 -0.28480493 0.82966008
theta 57 -1.50251496 0.4701266 -2.42394617 -0.58108375
theta 58 0.65293713 0.2702224 0.12331099 1.18256326
theta 61 -1.29369737 0.4442700 -2.16445055 -0.42294420
theta 62 -1.73765885 0.5004374 -2.71849805 -0.75681965
theta 64 1.41599504 0.2960187 0.83580903 1.99618106
theta 66 -0.39282316 0.3400576 -1.05932379 0.27367747
```

Subject NA Group: 2

NA pattern: x x x x x x x x x x NA

Collapsed log-likelihood: -13.08991

Number of iterations: 8

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 41	1.997126	0.3856216	1.241321	2.752930

Subject NA Group: 3

NA pattern: x x x x x x x x NA x x

Collapsed log-likelihood: -12.05070

Number of iterations: 6

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 53	-0.7541428	0.3933522	-1.525099	0.01681335

Subject NA Group: 4

NA pattern: x x x x x x x x NA x x x

Collapsed log-likelihood: -9.898949

Number of iterations: 8

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 60	-2.242055	0.5933217	-3.404944	-1.079166

A.1.3 Resultados para a UPDRS3 (motilidade)

Parâmetros de dificuldade do item-categoria

Results of PCM estimation:

Call: PCM(X = updrs3.full)

Conditional log-likelihood: -1796.754

Number of iterations: 197

Number of parameters: 106

Item (Category) Difficulty Parameters (eta) with 0.95 CI:

Item (Category)	Estimate	Std. Error	lower CI	upper CI
UPDRS3.1.c2	-1.508	0.520	-2.526	-0.490
UPDRS3.1.c3	-0.450	0.606	-1.637	0.738
UPDRS3.1.c4	0.916	0.739	-0.533	2.365
UPDRS3.2.c1	-1.671	0.459	-2.571	-0.771
UPDRS3.2.c2	-2.253	0.465	-3.165	-1.341
UPDRS3.2.c3	-0.090	0.641	-1.347	1.167
UPDRS3.2.c4	1.740	0.911	-0.046	3.527
UPDRS3.3Neck.c1	-1.276	0.546	-2.345	-0.206
UPDRS3.3Neck.c2	-2.689	0.508	-3.685	-1.693
UPDRS3.3Neck.c3	-2.114	0.553	-3.198	-1.030
UPDRS3.3Neck.c4	0.293	0.797	-1.269	1.855
UPDRS3.3RUE.c1	-2.551	0.554	-3.637	-1.464
UPDRS3.3RUE.c2	-2.787	0.573	-3.910	-1.663
UPDRS3.3RUE.c3	0.005	0.820	-1.602	1.613
UPDRS3.3LUE.c1	-1.144	0.417	-1.962	-0.327
UPDRS3.3LUE.c2	-1.747	0.419	-2.568	-0.926
UPDRS3.3LUE.c3	0.692	0.654	-0.590	1.973
UPDRS3.3LUE.c4	3.176	1.190	0.843	5.508
UPDRS3.3RLE.c1	-0.633	0.372	-1.362	0.097
UPDRS3.3RLE.c2	-0.851	0.390	-1.615	-0.088
UPDRS3.3RLE.c3	0.842	0.552	-0.240	1.924
UPDRS3.3RLE.c4	2.894	0.873	1.183	4.605
UPDRS3.3LLE.c1	-0.636	0.372	-1.366	0.094
UPDRS3.3LLE.c2	-0.813	0.393	-1.582	-0.043
UPDRS3.3LLE.c3	0.663	0.529	-0.375	1.701
UPDRS3.3LLE.c4	2.854	0.869	1.151	4.557
UPDRS3.4R.c1	-2.527	0.548	-3.600	-1.453
UPDRS3.4R.c2	-2.485	0.579	-3.620	-1.351
UPDRS3.4R.c3	-1.909	0.613	-3.110	-0.708
UPDRS3.4R.c4	0.836	0.940	-1.006	2.679
UPDRS3.4L.c1	-2.245	0.630	-3.480	-1.009
UPDRS3.4L.c2	-3.460	0.616	-4.666	-2.253
UPDRS3.4L.c3	-2.818	0.647	-4.087	-1.549
UPDRS3.4L.c4	-0.375	0.837	-2.015	1.265
UPDRS3.5R.c1	-1.809	0.423	-2.639	-0.980
UPDRS3.5R.c2	-1.640	0.461	-2.543	-0.737
UPDRS3.5R.c3	-0.074	0.587	-1.225	1.077
UPDRS3.5L.c1	-1.486	0.443	-2.353	-0.618
UPDRS3.5L.c2	-1.958	0.454	-2.848	-1.068
UPDRS3.5L.c3	-0.703	0.543	-1.767	0.361
UPDRS3.5L.c4	2.569	1.176	0.265	4.873
UPDRS3.6R.c1	-2.035	0.446	-2.909	-1.161
UPDRS3.6R.c2	-1.855	0.483	-2.802	-0.909

UPDRS3.6R.c3	-0.756	0.561	-1.855	0.343
UPDRS3.6L.c1	-1.573	0.420	-2.396	-0.751
UPDRS3.6L.c2	-1.688	0.451	-2.573	-0.804
UPDRS3.6L.c3	-1.308	0.485	-2.258	-0.357
UPDRS3.6L.c4	1.548	0.867	-0.152	3.248
UPDRS3.7R.c1	-1.318	0.375	-2.052	-0.584
UPDRS3.7R.c2	-0.967	0.433	-1.816	-0.119
UPDRS3.7R.c3	-0.346	0.491	-1.308	0.616
UPDRS3.7R.c4	0.830	0.602	-0.350	2.009
UPDRS3.7L.c1	-1.346	0.437	-2.202	-0.490
UPDRS3.7L.c2	-1.953	0.453	-2.841	-1.064
UPDRS3.7L.c3	-1.298	0.509	-2.297	-0.300
UPDRS3.7L.c4	-0.109	0.603	-1.292	1.074
UPDRS3.8R.c1	-1.231	0.389	-1.994	-0.468
UPDRS3.8R.c2	-1.101	0.432	-1.947	-0.255
UPDRS3.8R.c3	-0.415	0.488	-1.372	0.542
UPDRS3.8R.c4	2.214	0.866	0.516	3.911
UPDRS3.8L.c1	-1.137	0.408	-1.936	-0.338
UPDRS3.8L.c2	-1.348	0.432	-2.194	-0.501
UPDRS3.8L.c3	-0.840	0.475	-1.771	0.092
UPDRS3.8L.c4	1.182	0.658	-0.107	2.471
UPDRS3.9.c1	0.365	0.324	-0.270	0.999
UPDRS3.9.c2	1.506	0.480	0.565	2.447
UPDRS3.9.c3	1.839	0.494	0.871	2.806
UPDRS3.9.c4	3.266	0.666	1.960	4.571
UPDRS3.10.c1	-1.545	0.374	-2.279	-0.811
UPDRS3.10.c2	-0.825	0.444	-1.695	0.044
UPDRS3.10.c3	0.428	0.565	-0.679	1.535
UPDRS3.10.c4	1.740	0.672	0.423	3.057
UPDRS3.11.c1	1.012	0.342	0.341	1.682
UPDRS3.11.c2	3.794	1.075	1.686	5.901
UPDRS3.11.c3	3.100	0.659	1.808	4.392
UPDRS3.11.c4	3.017	0.573	1.894	4.140
UPDRS3.13.c1	-0.791	0.391	-1.558	-0.024
UPDRS3.13.c2	-0.906	0.422	-1.732	-0.080
UPDRS3.13.c3	-0.562	0.457	-1.458	0.334
UPDRS3.13.c4	1.759	0.737	0.314	3.204
UPDRS3.14.c1	-1.677	0.448	-2.555	-0.799
UPDRS3.14.c2	-1.767	0.481	-2.710	-0.824
UPDRS3.14.c3	-1.538	0.504	-2.526	-0.550
UPDRS3.14.c4	0.963	0.744	-0.496	2.422
UPDRS3.15R.c1	-0.874	0.305	-1.472	-0.276
UPDRS3.15R.c2	0.985	0.484	0.038	1.933
UPDRS3.15R.c3	3.889	1.147	1.641	6.138
UPDRS3.15L.c1	-1.309	0.316	-1.927	-0.690
UPDRS3.15L.c2	0.893	0.521	-0.128	1.913
UPDRS3.16R.c1	-0.406	0.282	-0.959	0.146
UPDRS3.16R.c2	2.932	0.783	1.398	4.466
UPDRS3.16L.c1	-1.190	0.307	-1.791	-0.589
UPDRS3.16L.c2	1.561	0.594	0.396	2.726
UPDRS3.17lipJaw.c1	1.718	0.365	1.003	2.433
UPDRS3.17lipJaw.c2	4.901	1.109	2.727	7.074
UPDRS3.17RUE.c1	0.224	0.288	-0.341	0.788
UPDRS3.17RUE.c2	2.375	0.564	1.270	3.480
UPDRS3.17RUE.c3	3.432	0.712	2.037	4.827
UPDRS3.17LUE.c1	0.766	0.303	0.171	1.361
UPDRS3.17LUE.c2	2.509	0.526	1.478	3.540
UPDRS3.17RLE.c1	2.460	0.453	1.573	3.348

UPDRS3.17LLE.c1	4.638	1.081	2.519	6.757
UPDRS3.18.c1	-0.361	0.298	-0.944	0.223
UPDRS3.18.c2	2.397	0.766	0.896	3.898
UPDRS3.18.c3	1.793	0.534	0.746	2.840
UPDRS3.18.c4	1.835	0.497	0.861	2.809

Parâmetros da capacidade do sujeito

Estimation of Ability Parameters

Subject NA Group: 1

NA pattern: x

Collapsed log-likelihood: -1158.145

Number of iterations: 12

Number of parameters: 37

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 1	2.47073220	0.2877748	1.90670386	3.03476053
theta 2	0.98310942	0.2057726	0.57980253	1.38641632
theta 3	-1.24515202	0.2777034	-1.78944077	-0.70086327
theta 4	-1.24515202	0.2777034	-1.78944077	-0.70086327
theta 5	-0.56936199	0.2452905	-1.05012254	-0.08860145
theta 6	-0.75551146	0.2530584	-1.25149682	-0.25952609
theta 7	0.44276352	0.2072490	0.03656295	0.84896409
theta 8	0.26629543	0.2130913	-0.15135587	0.68394673
theta 9	1.20245032	0.2138512	0.78330965	1.62159099
theta 10	-0.33763856	0.2362357	-0.80065208	0.12537496
theta 11	0.56959568	0.2041965	0.16937785	0.96981352
theta 12	0.81664626	0.2027680	0.41922837	1.21406414
theta 13	0.07864999	0.2202132	-0.35295987	0.51025985
theta 16	0.69356124	0.2026295	0.29641466	1.09070782
theta 17	0.31135916	0.2114795	-0.10313311	0.72585143
theta 18	-0.07066359	0.2260017	-0.51361876	0.37229159
theta 19	-0.75551146	0.2530584	-1.25149682	-0.25952609
theta 20	-0.75551146	0.2530584	-1.25149682	-0.25952609
theta 21	-1.48914119	0.2933097	-2.06401763	-0.91426476
theta 22	0.02973650	0.2221117	-0.40559442	0.46506743
theta 24	-0.45132454	0.2406240	-0.92293883	0.02028974
theta 25	0.07864999	0.2202132	-0.35295987	0.51025985
theta 26	0.31135916	0.2114795	-0.10313311	0.72585143
theta 27	-3.62607318	0.6044343	-4.81074272	-2.44140364
theta 28	-0.75551146	0.2530584	-1.25149682	-0.25952609
theta 30	-0.12219114	0.2279874	-0.56903824	0.32465596
theta 31	-2.08952852	0.3448358	-2.76539435	-1.41366269
theta 32	0.98310942	0.2057726	0.57980253	1.38641632
theta 33	0.31135916	0.2114795	-0.10313311	0.72585143
theta 34	-0.95443349	0.2621864	-1.46830940	-0.44055758
theta 35	0.26629543	0.2130913	-0.15135587	0.68394673
theta 36	0.17401839	0.2165397	-0.25039161	0.59842839
theta 37	-0.17463037	0.2300004	-0.62542283	0.27616209
theta 38	0.26629543	0.2130913	-0.15135587	0.68394673
theta 39	0.56959568	0.2041965	0.16937785	0.96981352
theta 40	-0.28232970	0.2341191	-0.74119474	0.17653535
theta 42	0.35576025	0.2099595	-0.05575286	0.76727336
theta 43	-0.39395868	0.2384006	-0.86121534	0.07329798

```
theta 44 0.12673546 0.2183533 -0.30122905 0.55469998
theta 45 -1.32357532 0.2824350 -1.87713784 -0.77001281
theta 46 1.39314143 0.2231706 0.95573513 1.83054772
theta 47 0.48547140 0.2060837 0.08155467 0.88938812
theta 48 -1.57700276 0.2996093 -2.16422629 -0.98977923
theta 49 -0.17463037 0.2300004 -0.62542283 0.27616209
theta 50 0.61114268 0.2034961 0.21229774 1.00998762
theta 51 -0.28232970 0.2341191 -0.74119474 0.17653535
theta 52 -0.75551146 0.2530584 -1.25149682 -0.25952609
theta 53 -0.12219114 0.2279874 -0.56903824 0.32465596
theta 54 -0.07066359 0.2260017 -0.51361876 0.37229159
theta 55 -1.32357532 0.2824350 -1.87713784 -0.77001281
theta 56 -0.12219114 0.2279874 -0.56903824 0.32465596
theta 57 -1.48914119 0.2933097 -2.06401763 -0.91426476
theta 58 1.39314143 0.2231706 0.95573513 1.83054772
theta 60 -1.09562312 0.2693540 -1.62354719 -0.56769904
theta 61 0.65243977 0.2029709 0.25462413 1.05025541
theta 62 -0.02002769 0.2240422 -0.45914228 0.41908690
theta 64 1.06884681 0.2084620 0.66026887 1.47742476
theta 66 -1.57700276 0.2996093 -2.16422629 -0.98977923
```

Subject NA Group: 2

```
NA pattern: x x x x x x x x x x x x x x x x x x x x x x NA x x x x x x x x x
Collapsed log-likelihood: -26.05807
Number of iterations: 5
Number of parameters: 1
```

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 14	-0.3532403	0.2391530	-0.8219715	0.1154909

Subject NA Group: 3

```
NA pattern: x x x x x x x x x x x x x x x x x x NA x x x x x x x x x x x x
Collapsed log-likelihood: -24.75662
Number of iterations: 5
Number of parameters: 1
```

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 41	1.412079	0.2320290	0.9573103	1.866847

Subject NA Group: 4

```
NA pattern: x x x x x x x x NA NA x x x x x x x x x x x x x x x x x x x x
Collapsed log-likelihood: -33.34938
Number of iterations: 4
Number of parameters: 1
```

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 23	0.6507386	0.2080371	0.2429933	1.058484

Subject NA Group: 5

```
NA pattern: x x NA NA NA NA NA x x x x x x x x x x x x x x x x x x x x x x
Collapsed log-likelihood: -61.25619
```


Number of iterations: 5
 Number of parameters: 2

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 15	0.5160207	0.2198866	0.08505087	0.9469905
theta 29	0.4673270	0.2214801	0.03323390	0.9014200

A.1.4 Resultados para a UPDRS4 (complicações do tratamento)

Parâmetros de dificuldade do item-categoria

Results of PCM estimation:

Call: PCM(X = updrs4.full)

Conditional log-likelihood: -212.5461

Number of iterations: 148

Number of parameters: 23

Item (Category) Difficulty Parameters (eta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
UPDRS4.1.c2	-0.173	13.828	-27.276	26.930
UPDRS4.1.c3	-0.657	20.729	-41.285	39.970
UPDRS4.1.c4	-2.031	27.630	-56.184	52.123
UPDRS4.2.c1	-0.403	6.917	-13.960	13.154
UPDRS4.2.c2	-0.616	13.820	-27.702	26.471
UPDRS4.2.c3	-0.892	20.725	-41.512	39.728
UPDRS4.2.c4	-0.435	27.634	-54.597	53.727
UPDRS4.3.c1	-2.171	6.919	-15.731	11.390
UPDRS4.3.c2	-1.230	13.823	-28.323	25.863
UPDRS4.3.c3	0.086	20.743	-40.569	40.742
UPDRS4.3.c4	-0.313	27.644	-54.495	53.869
UPDRS4.4.c1	-1.334	6.920	-14.897	12.228
UPDRS4.4.c2	-1.069	13.824	-28.163	26.025
UPDRS4.4.c3	-2.649	20.724	-43.267	37.969
UPDRS4.4.c4	-2.047	27.630	-56.200	52.106
UPDRS4.5.c1	-1.911	6.917	-15.468	11.646
UPDRS4.5.c2	-0.945	13.823	-28.037	26.147
UPDRS4.5.c3	0.383	20.742	-40.270	41.036
UPDRS4.5.c4	-1.397	27.631	-55.552	52.759
UPDRS4.6.c1	0.290	6.914	-13.261	13.842
UPDRS4.6.c2	2.359	13.846	-24.779	29.496
UPDRS4.6.c3	17.110	393.650	-754.431	788.650
UPDRS4.6.c4	1.551	27.638	-52.618	55.720

Parâmetros da capacidade do sujeito

Estimation of Ability Parameters

Subject NA Group: 1

NA pattern: x x x x x x

Collapsed log-likelihood: -104.6945

Number of iterations: 18

Number of parameters: 15

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 1	-0.12388353	0.3368632	-0.7841233	0.53635627
theta 2	0.20172630	0.3251100	-0.4354776	0.83893025
theta 3	-1.02781955	0.4811941	-1.9709426	-0.08469651
theta 6	-1.65507305	0.6570064	-2.9427819	-0.36736420
theta 7	-1.65507305	0.6570064	-2.9427819	-0.36736420
theta 11	-1.29274172	0.5520166	-2.3746745	-0.21080899
theta 14	0.09552783	0.3269894	-0.5453597	0.73641536
theta 15	-1.65507305	0.6570064	-2.9427819	-0.36736420
theta 16	-0.82001701	0.4330217	-1.6687240	0.02869001
theta 18	-1.02781955	0.4811941	-1.9709426	-0.08469651
theta 20	-0.82001701	0.4330217	-1.6687240	0.02869001
theta 21	-0.64749439	0.3993672	-1.4302397	0.13525094
theta 22	-2.18809285	0.8105003	-3.7766443	-0.59954138
theta 23	-0.01254938	0.3308294	-0.6609631	0.63586434
theta 24	-1.65507305	0.6570064	-2.9427819	-0.36736420
theta 25	-0.36374312	0.3580578	-1.0655236	0.33803733
theta 26	-0.36374312	0.3580578	-1.0655236	0.33803733
theta 27	-3.06076825	1.0935690	-5.2041241	-0.91741242
theta 28	-2.18809285	0.8105003	-3.7766443	-0.59954138
theta 29	-1.29274172	0.5520166	-2.3746745	-0.21080899
theta 30	-0.64749439	0.3993672	-1.4302397	0.13525094
theta 32	-1.29274172	0.5520166	-2.3746745	-0.21080899
theta 33	-0.12388353	0.3368632	-0.7841233	0.53635627
theta 35	-0.36374312	0.3580578	-1.0655236	0.33803733
theta 36	-1.29274172	0.5520166	-2.3746745	-0.21080899
theta 39	-1.02781955	0.4811941	-1.9709426	-0.08469651
theta 40	-0.49789577	0.3753484	-1.2335651	0.23777354
theta 41	0.30736850	0.3253175	-0.3302420	0.94497903
theta 42	-0.24015896	0.3456467	-0.9176141	0.43729616
theta 43	0.09552783	0.3269894	-0.5453597	0.73641536
theta 45	-0.49789577	0.3753484	-1.2335651	0.23777354
theta 48	-1.02781955	0.4811941	-1.9709426	-0.08469651
theta 50	-0.82001701	0.4330217	-1.6687240	0.02869001
theta 54	-0.82001701	0.4330217	-1.6687240	0.02869001
theta 55	-0.36374312	0.3580578	-1.0655236	0.33803733
theta 56	-0.49789577	0.3753484	-1.2335651	0.23777354
theta 57	-1.02781955	0.4811941	-1.9709426	-0.08469651
theta 61	-3.06076825	1.0935690	-5.2041241	-0.91741242
theta 62	-0.36374312	0.3580578	-1.0655236	0.33803733
theta 64	-1.65507305	0.6570064	-2.9427819	-0.36736420
theta 65	-0.24015896	0.3456467	-0.9176141	0.43729616
theta 66	0.09552783	0.3269894	-0.5453597	0.73641536

Subject NA Group: 2

NA pattern: x x x x NA

Collapsed log-likelihood: -3.903920

Number of iterations: 6

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
--	----------	-----------	-------	--------

theta 49 -0.5886905 0.4030058 -1.378567 0.2011864

Subject NA Group: 3
 NA pattern: x x x NA x x
 Collapsed log-likelihood: -7.688785
 Number of iterations: 6
 Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 47	-0.4775032	0.4211269	-1.302897	0.3478905

Subject NA Group: 4
 NA pattern: x x NA x x x
 Collapsed log-likelihood: -16.88473
 Number of iterations: 9
 Number of parameters: 3

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 4	0.05210136	0.3494696	-0.6328464	0.7370491
theta 12	-0.33074878	0.3708719	-1.0576443	0.3961467
theta 44	-1.35787856	0.6050368	-2.5437289	-0.1720282

Subject NA Group: 5
 NA pattern: x NA x x x x
 Collapsed log-likelihood: -4.315623
 Number of iterations: 5
 Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 38	-0.4242392	0.4121342	-1.232007	0.3835289

A.1.5 Resultados para a Escala de London

Parâmetros de dificuldade do item-categoria

Results of PCM estimation:

Call: PCM(X = london.full)

Conditional log-likelihood: -299.1714

Number of iterations: 60

Number of parameters: 27

Item (Category) Difficulty Parameters (eta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
LondonMobility.c2	-1.744	0.509	-2.742	-0.746
LondonMobility.c3	0.042	0.642	-1.216	1.299
LondonMobility.c4	1.183	0.748	-0.284	2.650
LondonMobility.c5	3.096	0.992	1.152	5.040

LondonPhysicalIndependance.c1	-0.962	0.458	-1.859	-0.065
LondonPhysicalIndependance.c2	-0.878	0.536	-1.927	0.172
LondonPhysicalIndependance.c3	-0.519	0.590	-1.676	0.637
LondonPhysicalIndependance.c4	0.954	0.727	-0.470	2.378
LondonPhysicalIndependance.c5	2.065	0.843	0.412	3.717
LondonOccupation.c1	-3.656	0.788	-5.201	-2.111
LondonOccupation.c2	-4.101	0.834	-5.735	-2.467
LondonOccupation.c3	-4.116	0.817	-5.717	-2.515
LondonOccupation.c4	-2.631	0.861	-4.318	-0.944
LondonOccupation.c5	-1.783	0.885	-3.518	-0.048
LondonSocialIntegration.c1	-1.228	0.419	-2.049	-0.406
LondonSocialIntegration.c2	-0.890	0.487	-1.846	0.065
LondonSocialIntegration.c3	0.641	0.616	-0.567	1.849
LondonSocialIntegration.c4	3.615	1.132	1.395	5.834
LondonOrientation.c1	-0.121	0.344	-0.796	0.553
LondonOrientation.c2	1.217	0.483	0.271	2.163
LondonOrientation.c3	4.351	1.081	2.232	6.471
LondonOrientation.c4	4.698	0.951	2.835	6.562
LondonEconomicIndependance.c1	-2.636	0.565	-3.743	-1.529
LondonEconomicIndependance.c2	-2.628	0.580	-3.765	-1.491
LondonEconomicIndependance.c3	-1.122	0.650	-2.396	0.152
LondonEconomicIndependance.c4	5.747	6.318	-6.636	18.131
LondonEconomicIndependance.c5	2.354	0.961	0.470	4.237

Parâmetros da capacidade do sujeito

Estimation of Ability Parameters

Subject NA Group: 1

NA pattern: x x x x x x

Collapsed log-likelihood: -139.1352

Number of iterations: 16

Number of parameters: 20

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 1	1.53072484	0.3724520	0.80073229	2.26071739
theta 2	1.39188595	0.3732830	0.66026478	2.12350711
theta 3	-2.69001795	0.8962863	-4.44670687	-0.93332903
theta 5	0.95807910	0.3901637	0.19337234	1.72278586
theta 6	0.08882843	0.4483959	-0.79001129	0.96766815
theta 7	0.46631779	0.4212857	-0.35938704	1.29202263
theta 10	-0.34394385	0.4836601	-1.29190031	0.60401262
theta 11	-0.86174550	0.5384457	-1.91707977	0.19358876
theta 13	-0.11940308	0.4645738	-1.02995101	0.79114486
theta 14	0.08882843	0.4483959	-0.79001129	0.96766815
theta 15	-0.11940308	0.4645738	-1.02995101	0.79114486
theta 16	1.39188595	0.3732830	0.66026478	2.12350711
theta 18	0.63886579	0.4096507	-0.16403481	1.44176639
theta 20	-0.11940308	0.4645738	-1.02995101	0.79114486
theta 21	-2.02601167	0.7429094	-3.48208726	-0.56993608
theta 22	-0.86174550	0.5384457	-1.91707977	0.19358876
theta 23	0.28345105	0.4341376	-0.56744301	1.13434512
theta 25	0.63886579	0.4096507	-0.16403481	1.44176639
theta 26	-0.11940308	0.4645738	-1.02995101	0.79114486
theta 27	-0.86174550	0.5384457	-1.91707977	0.19358876

```

theta 28 -0.11940308 0.4645738 -1.02995101 0.79114486
theta 30 0.95807910 0.3901637 0.19337234 1.72278586
theta 31 -3.73243296 1.1783659 -6.04198769 -1.42287823
theta 32 -0.58904765 0.5073442 -1.58342405 0.40532876
theta 33 0.28345105 0.4341376 -0.56744301 1.13434512
theta 34 -2.69001795 0.8962863 -4.44670687 -0.93332903
theta 35 0.46631779 0.4212857 -0.35938704 1.29202263
theta 37 -2.69001795 0.8962863 -4.44670687 -0.93332903
theta 38 1.67020660 0.3751659 0.93489493 2.40551827
theta 39 0.95807910 0.3901637 0.19337234 1.72278586
theta 41 1.39188595 0.3732830 0.66026478 2.12350711
theta 42 0.63886579 0.4096507 -0.16403481 1.44176639
theta 46 1.67020660 0.3751659 0.93489493 2.40551827
theta 47 0.95807910 0.3901637 0.19337234 1.72278586
theta 48 -1.17417305 0.5817340 -2.31435070 -0.03399541
theta 49 0.46631779 0.4212857 -0.35938704 1.29202263
theta 50 0.63886579 0.4096507 -0.16403481 1.44176639
theta 51 0.63886579 0.4096507 -0.16403481 1.44176639
theta 54 0.08882843 0.4483959 -0.79001129 0.96766815
theta 55 0.08882843 0.4483959 -0.79001129 0.96766815
theta 56 0.46631779 0.4212857 -0.35938704 1.29202263
theta 57 0.46631779 0.4212857 -0.35938704 1.29202263
theta 58 1.25135403 0.3768560 0.51272991 1.98997815
theta 59 -1.17417305 0.5817340 -2.31435070 -0.03399541
theta 60 -2.02601167 0.7429094 -3.48208726 -0.56993608
theta 61 0.63886579 0.4096507 -0.16403481 1.44176639
theta 62 0.08882843 0.4483959 -0.79001129 0.96766815
theta 65 0.80237066 0.3992390 0.01987669 1.58486462
theta 66 0.63886579 0.4096507 -0.16403481 1.44176639
theta 68 -0.34394385 0.4836601 -1.29190031 0.60401262
theta 70 1.25135403 0.3768560 0.51272991 1.98997815
theta 72 -0.11940308 0.4645738 -1.02995101 0.79114486
theta 73 0.95807910 0.3901637 0.19337234 1.72278586
theta 74 -1.54833005 0.6453604 -2.81321321 -0.28344689
theta 75 -2.02601167 0.7429094 -3.48208726 -0.56993608
theta 77 1.81341535 0.3826086 1.06351630 2.56331441
theta 78 -1.54833005 0.6453604 -2.81321321 -0.28344689

```

Subject NA Group: 2

NA pattern: x x x x NA

Collapsed log-likelihood: -3.158255

Number of iterations: 9

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
theta 17	-2.201813	0.8804553	-3.927474	-0.4761528

Subject NA Group: 3

NA pattern: x x x x NA x

Collapsed log-likelihood: -7.625566

Number of iterations: 4

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
--	----------	-----------	-------	--------

theta 36 0.6645837 0.4375311 -0.1929615 1.522129

Subject NA Group: 4

NA pattern: x x x NA NA

Collapsed log-likelihood: -6.37305

Number of iterations: 3

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
--	----------	-----------	-------	--------

theta 71	0.1988637	0.4990733	-0.779302	1.177029
----------	-----------	-----------	-----------	----------

Subject NA Group: 5

NA pattern: x x x NA NA NA

Collapsed log-likelihood: -4.656901

Number of iterations: 5

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
--	----------	-----------	-------	--------

theta 40	-1.002976	0.6203707	-2.218881	0.2129278
----------	-----------	-----------	-----------	-----------

Subject NA Group: 6

NA pattern: x x NA x x x

Collapsed log-likelihood: -2.230267

Number of iterations: 9

Number of parameters: 1

ML estimated ability parameters (without spline interpolated values):

	Estimate	Std. Err.	2.5 %	97.5 %
--	----------	-----------	-------	--------

theta 24	-3.636596	1.207959	-6.004151	-1.269041
----------	-----------	----------	-----------	-----------

BIBLIOGRAFIA

- Beck, A., Ward, C., Mendelson, M., Mock, J. and Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.
- Becker, R., Chambers, J. and Wilks, A. (1988). *The New S Language: A Programming Environment for Data Analysis and Graphics* (1 ed.). Pacific Grove (EUA): Wadsworth Brooks & Cole.
- Benabid, A., Pollak, P., Louveau, A., S., H. and Rougemont, J. (1987). Combined (thalamotomy and stimulation) stereotactic surgery of the vim thalamic nucleus for bilateral Parkinson's disease. *Appl. Neurophysiol* **50**, 344–346.
- Chambers, J. (1992). Linear models. In *Statistical Models in S*, Chapter 4. Wadsworth & Brooks/Cole.
- de Rijk, M., Launer, L., Breteler, M., Dartigues, J. and Baldereschi, M. (2000). Prevalence of Parkinson's disease in Europe: a collaborative study of population based cohorts. *Neurology* **54**, 21–23.
- Fischer, G. and Molenaar, I. (1994). *Rasch Models: foundations, recent developments, and applications*. New York: Springer-Verlag.
- Folstein, M., Folstein, S. and McHugh, P. (1975). Mini-mental state. a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research* **12**, 189–198.
- Hosmer, D. and Lemeshow, S. (2000). *Applied logistic regression*. New Jersey (EUA): John Wiley and Sons.

- Jolliffe, I. (2002). *Principal Components Analysis* (2 ed.). New York: Springer-Verlag.
- Keats, J. (1990). Measurement in educational research. In *The International Encyclopedia of Educational Evaluation*, pp. 237–244. Pergamon.
- Kelderman, H. and Rijkes, C. (1994). Loglinear multidimensional irt models for polytomously scored items. *Psychometrika* **59**, 149–176.
- Kumar, A. and Calne, D. (2004). Approach to the patient with a movement disorder and overview of movement disorders. In *Movement disorders: neurologic principles and practice*, pp. 3–15. McGraw-Hill.
- Lau, R., Hartman, K. and Ware, J. (1986). Health as a value: methodological and theoretical considerations. *Health Psychol* **5**, 25–43.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology* **22**, 1–55.
- Mair, P. and Hatzinger (2007). Extended rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software* **20**, 1–20.
- Mardia, K., Kent, J. and Bibby, J. (1979). *Multivariate Analysis* (1 ed.). London: Academic Press.
- McLachlan, G. (2004). *Discriminant Analysis and Statistical Pattern Recognition*. New York (EUA): Wiley-Interscience.
- McNeil, D. (1977). *Interactive Data Analysis* (1 ed.). New York: Wiley.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: B.D. Wright.
- Rousseauw, J., du Plessis, J., A., B., Jordaan, P., Kotze, J. and Ferreira, J. (1983). Coronary risk factor screening in three rural communities. *South African Medical Journal* **64**, 430–436.
- Serneels, S. and Verdonck, T. (2008). Principal component analysis for data containing outliers and missing elements. *Computational Statistics and Data Analysis* **52**, 1712–1727.
- Team, R. D. C. (2009). *R: A Language and Environment for Statistical Computing*. Vienna (Austria): R Foundation for Statistical Computing.
- Todorov, V. and Filzmoser, P. (2009). An object oriented framework for robust multivariate analysis. *Journal of Statistical Software* **32**, 1–47.

Virgilio, G.D. e Gonçalves, J. (2003). Escalas clínicas. In *Doença de Parkinson - Manual Prático*, pp. 181–186. LIDEL.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the VI International Congress of Genetics* **1**, 356–366.

