



Universidade de Aveiro  
Ano 2021

**TIAGO FILIPE  
RODRIGUES COSTA**

**RELATÓRIO DE ESTÁGIO REALIZADO NO  
INSTITUTO DE SAÚDE PÚBLICA DA UNIVERSIDADE  
DO PORTO (ISPUP)**





Universidade de Aveiro  
Ano 2021

**TIAGO FILIPE  
RODRIGUES COSTA**

**RELATÓRIO DE ESTÁGIO REALIZADO NO  
INSTITUTO DE SAÚDE PÚBLICA DA UNIVERSIDADE  
DO PORTO (ISPUP)**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre Estatística Médica, realizada sob a orientação científica da Doutora Margarida Sâncio da Cruz Fardilha, Professora Auxiliar com Agregação do Departamento de Ciências Médicas da Universidade de Aveiro, e supervisão do Doutor Milton Severo, Investigador Doutoramento Integrado do Instituto de Saúde Pública da Universidade do Porto (ISPUP).

Apoio financeiro da FCT e do FSE no âmbito do III Quadro Comunitário de Apoio.



## **o júri**

presidente

**Prof. Doutora Vera Mónica Almeida Afreixo**  
Professora auxiliar da Universidade de Aveiro

**Prof. Doutora Joana Antunes de Lima Bastos**  
Professora auxiliar convidada da Faculdade de Medicina da Universidade de Coimbra

**Prof. Doutora Margarida Sâncio da Cruz Fardilha**  
Professora auxiliar com agregação da Universidade de Aveiro

## **agradecimentos**

Em primeiro lugar, gostaria de agradecer às pessoas que me acompanharam ao longo deste trabalho, mais especificamente ao meu supervisor, o Prof. Doutor Milton Severo, ao Doutor Luís Ruano, à Prof. Doutora Vera Afreixo e à minha orientadora, Prof. Doutora Margarida Fardilha.

Em segundo lugar, mas não menos importante, quero deixar um enorme obrigado a toda a minha família, pelo seu apoio incondicional e carinho, em especial aos meus pais, Pedro e Paula, à minha irmã Bárbara, e à minha grande companheira, a minha avó Lurdes, porque nada seria possível sem eles.

Quero ainda agradecer a todos os meus amigos, os que me acompanharam desde sempre e os que fui acumulando, desde a minha aldeia, Laceiras, até à cidade que me acolheu, Aveiro.

Quero deixar um reconhecimento especial à Clara, ao Gil, à Sara, à Mondragão, ao Jorge e à Soares, por todo o apoio, confiança, amizade e motivação ao longo dos últimos anos.

Um obrigado sentido à Maria, à Filipa, à Dalila e à Vigo por toda a companhia e partilha.

Uma palavra de gratidão a todos os meus sobrinhos, em especial ao Miguel, pelo seu apoio e amizade incansável.

Quero também deixar um obrigado às pessoas que sem saberem foram as minhas referências, em especial aos 3 Mestres pelo seu exemplo e partilha de experiências.

Finalmente, agradecer a toda e qualquer pessoa que interagiu comigo, incluindo, indubitavelmente, todos os Professores que tive até agora, por, de alguma forma, me terem ajudado a tornar naquilo que sou hoje.

E, por último, deixar um obrigado e a promessa de regresso à Academia, à Universidade de Aveiro e a Aveiro.

**palavras-chave**

Esclerose múltipla, Défice cognitivo, ISPUP, Modelos lineares de efeitos mistos, Regressão logística.

**resumo**

Este relatório tem por objetivo descrever o estágio no Instituto de Saúde Pública da Universidade do Porto (ISPUP), realizado no âmbito do segundo ano curricular do Mestrado em Estatística Médica da Universidade de Aveiro (UA).

Este estágio de 6 meses no ISPUP incluiu diferentes atividades, todas elas incluídas no projeto “Tracking cognitive decline in Multiple Sclerosis”, onde se pretendeu, fundamentalmente, avaliar a capacidade de uma ferramenta de avaliação da performance cognitiva (“Brain on Track” – BoT) distinguir pessoas com esclerose múltipla consoante o seu défice cognitivo. Neste sentido, foram aplicados conhecimentos adquiridos ao longo do Mestrado e no decorrer do presente estágio, tendo sido feita uma análise descritiva da amostra em estudo, construídos modelos lineares de efeitos mistos e utilizado a regressão logística para avaliar a capacidade de discriminação dos modelos.

As pontuações totais do BoT apresentaram uma trajetória quadrática e a ferramenta apresentou capacidade de discriminar pacientes com esclerose múltipla e défice cognitivo de pacientes com a performance cognitiva preservada, tendo uma das principais limitações sido o tamanho amostral. A ferramenta BoT apresenta-se como verdadeiramente promissora uma vez que mantém a sua capacidade de discriminação e vai de encontro às recomendações de alguns profissionais de ser necessário efetuar um diagnóstico de forma longitudinal.

**keywords**

Multiple sclerosis, cognitive impairment, ISPUP, Linear mixed-effects models, Logistic regression.

**abstract**

This report aims to describe the traineeship held at Instituto de Saúde Pública da Universidade do Porto (ISPUP), carried out within the second curricular year of the Master's Degree in Medical Statistics of the University of Aveiro (UA).

This 6-month traineeship at ISPUP included different activities, all within the scope of the project "Tracking cognitive decline in Multiple Sclerosis", where the main objective was to assess the ability of a tool for cognitive assessment ("Brain on Track" – BoT) to distinguish people with multiple sclerosis according to their cognitive impairment. For this, the knowledge acquired throughout the Master's degree and during the present traineeship was applied, a descriptive analysis of the sample under study was carried out, linear mixed-effects models were built and a logistic regression was also used to assess the discrimination ability of the models.

BoT total scores showed a quadratic trajectory and the tool was able to discriminate patients with multiple sclerosis and cognitive impairment from patients with preserved cognitive performance, despite one of the main limitations identified being the sample size. The BoT tool presents itself as truly promising as it maintains its ability to discriminate patients and meets the recommendations of some professionals that it is necessary to carry out a longitudinal diagnosis.





## Índice

Capítulo 1 .....	1
Introdução.....	1
1.1 Apresentação da entidade acolhedora.....	3
1.2 Descrição da ferramenta utilizada.....	3
Capítulo 2 .....	5
2.1 Performance cognitiva .....	5
2.1.1 Introdução.....	5
2.1.2 Estado da arte .....	6
2.2 Brain on Track – BoT .....	7
2.2.1 Desenvolvimento .....	7
a) Refinamento da ferramenta .....	8
b) Avaliação da fiabilidade do teste à distância.....	9
2.2.2 Aplicação .....	9
2.2.3 Vantagens e desvantagens.....	11
2.3 Modelos .....	12
2.3.1 Modelos lineares de efeitos mistos .....	12
a) Introdução .....	12
b) Aplicação .....	13
c) Exemplo .....	15
2.3.2 Regressão logística .....	19
a) Aplicação .....	19
b) Exemplo .....	21
Capítulo 3 .....	25
3.1 Atividades realizadas em estágio .....	25
3.2 Análise estatística .....	26
3.2.1 Enquadramento .....	26
3.2.2 Resultados.....	28
Capítulo 4 .....	38
4.1 Considerações finais.....	38
4.2 Conclusão.....	40
Bibliografia .....	42
Anexos .....	45

## Índice de Tabelas

Tabela 3.1 – Características da amostra e descrição por caso .....	28
Tabela 3.2 – Pontuações, por subteste, para as primeiras 4 repetições do teste de pacientes com esclerose múltipla .....	29
Tabela 3.3 – Análise de componentes principais para a primeira repetição de teste .....	30
Tabela 3.4 – Área sob a curva ROC (AUC), consoante a classificação escolhida e o modelo utilizado .....	33
Tabela 3.5 – Estimativas dos efeitos dos modelos lineares de efeitos mistos não ajustado e ajustado apenas para pacientes (sem controlos), de acordo com o GD 2 .....	35
Tabela 3.6 – Estimativas dos efeitos dos modelos lineares de efeitos mistos não ajustado e ajustado apenas para pacientes (sem controlos), de acordo com o GD 3 .....	36

## Índice de Figuras

Figura 2.1 – Linha de comando <i>lme</i> .....	16
Figura 2.2 – Área sob a curva ROC (AUC), para o modelo logístico dos pacientes de acordo com a presença de défice cognitivo dado GD 2 .....	24
Figura 3.1 – T scores de cada um dos pacientes com pelo menos 3 repetições de testes .....	31
Figura 3.2 – Curvas do modelo linear de efeitos mistos não ajustado, para pacientes de acordo com GD 2 .....	37
Figura 3.3 – Curvas do modelo linear de efeitos mistos não ajustado, para pacientes de acordo com GD 3 .....	37

## **Índice de Anexos**

Anexo A – Dicionário de variáveis da base de dados da bateria de exames neuropsicológicos inicial .....	46
Anexo B – Dicionário de variáveis da base de dados da bateria de exames neuropsicológicos final .....	51
Anexo C – Dicionário de variáveis da base de dados dos resultados do Brain on Track (BoT).....	56

## **Lista de siglas e abreviaturas**

ISPUP – Instituto de Saúde Pública da Universidade do Porto

UA – Universidade de Aveiro

BoT – Brain on Track

ROC – Receiver Operating Characteristic

FCT – Fundação para a Ciência e a Tecnologia

MMSE - Mini Mental State Examination

MoCA - Montreal Cognitive Assessment

AIC - Critério de Informação de Akaike

AUC - Área sob a Curva ROC

GD – Grau de Defeito

CI – Déficit Cognitivo

EDSS - Expanded Disability Status Scale

NA - Não Aplicável

## Capítulo 1

### Introdução

O presente relatório de estágio foi realizado no âmbito do 2º ano curricular do Mestrado em Estatística Médica da Universidade de Aveiro (UA), entre janeiro e junho de 2021. O objetivo fundamental do estágio passava pela aplicação de conhecimentos adquiridos no decurso do percurso académico, mas também adquirir novos conhecimentos bem como experiência profissional em contexto de um instituto de investigação. O estágio foi realizado no Instituto de Saúde Pública da Universidade do Porto (ISPUP), sob supervisão do Doutor Milton Severo.

O ISPUP é um instituto sediado no Porto que atua em diferentes domínios científicos no campo da saúde pública, nomeadamente, Epidemiologia e Bioestatística, Ambiente, Ciências do Comportamento e Promoção da Saúde e ainda das Ciências Sociais. Durante o período de estágio, as atividades que desenvolvi foram no âmbito do domínio da Epidemiologia e Bioestatística e todas relacionadas com o projeto “Tracking cognitive decline in Multiple Sclerosis”.

O objetivo geral das atividades desenvolvidas foi avaliar a capacidade de uma ferramenta de avaliação cognitiva (“Brain on Track” – BoT) diferenciar pessoas com esclerose múltipla consoante o seu défice cognitivo. Neste sentido, depois de recolhidos os dados, procedeu-se à análise estatística e à escrita de um artigo científico.

Este relatório está dividido em quatro capítulos. Neste primeiro capítulo, para além de uma descrição breve de cada capítulo, é ainda brevemente descrito o local do estágio curricular, ISPUP, bem como o *software* de análise estatística utilizado.

No segundo capítulo é apresentado um enquadramento referente ao projeto no âmbito do qual foram conduzidas as atividades de estágio, apresentando alguns conceitos, uma breve descrição do estado da arte de monitorização da performance cognitiva e a ferramenta de avaliação cognitiva em estudo, BoT. Na parte final deste capítulo são ainda descritos alguns conteúdos referentes aos modelos lineares de efeitos mistos e à regressão logística, com referência à Curva ROC (Receiver Operating Characteristic).

No terceiro capítulo, para além de descritas as atividades desenvolvidas no decorrer do estágio e a metodologia usada, são também apresentados e discutidos os resultados obtidos pela análise estatística levada a cabo.

Finalmente, no quarto e último capítulo, são apresentadas algumas considerações relacionadas com o estágio, incluindo o que aprendi, as principais dificuldades sentidas e como as ultrapassei, e as conclusões retiradas do trabalho desenvolvido.



## 1.1 Apresentação da entidade acolhedora

O ISPUP é o Instituto de Saúde Pública da Universidade do Porto, sediado na mesma cidade, criado no ano de 2007. O objetivo deste instituto passa por criar e divulgar conhecimentos, estimulando ainda a aplicação de tecnologias no âmbito da saúde pública, para que se promova e proteja a saúde das populações, e incorporando programas de ensino, investigação e serviços de modo a responder às necessidades sentidas (localmente, mas também a nível nacional e internacional).<sup>1,2</sup>

O ISPUP é constituído por uma equipa multidisciplinar e vasta, composta por estudantes de mestrado e de doutoramento, bolsiros de investigação e investigadores doutorados experientes em diferentes áreas, como a Bioestatística, as Ciências Ambientais, Ciências da Nutrição, entre outras.<sup>1,2</sup>

Sendo um instituto que prima pela excelência académica e pelo rigor científico, tem sob a sua alçada inúmeros projetos nacionais e internacionais financiados, o que, aliado ao facto de produzirem publicações científicas em revistas indexadas e de renome internacional, o torna numa das instituições portuguesas mais destacadas na investigação em saúde.<sup>1,2</sup>

A avaliação de Unidades de Investigação pela FCT, efetuada em 2017/2018 e conhecida em 2020, atribuiu a classificação de “Excelente” ao ISPUP, mais especificamente à sua Unidade de Investigação em Epidemiologia (EPIUnit).<sup>3</sup>

## 1.2 Descrição da ferramenta utilizada

Brevemente é aqui apresentada uma descrição do *software* de análise estatística e da sua interface utilizados durante o decorrer do estágio, *R* e *RStudio*.

### ***R* e *RStudio***

*R* e *Python* são as linguagens de programação estatísticas mais usadas, apesar de existirem outras opções. O *software R* (versão 3.6.1) consiste numa linguagem de programação utilizada para gráficos e execução de cálculos estatísticos. Esta linguagem providencia ao utilizador diferentes tipos de técnicas estatísticas e de manipulação de dados, produzindo gráficos e tabelas de elevada qualidade, para além de ser uma ferramenta gratuita compatível com diferentes plataformas e sistemas semelhantes a este mesmo.<sup>4,5</sup>

Como interface do *R*, utilizou-se o *software RStudio*, distinguem-se ainda duas edições: a aplicação do sistema operacional (*desktop*) e a que pode ser acedida remotamente por meio de uma página web, o *RStudio Server*.<sup>6</sup>

## Capítulo 2

### 2.1 Performance cognitiva

#### 2.1.1 Introdução

Diferentes estudos avaliaram já a relação entre alterações cognitivas e o normal envelhecimento, demonstrando que este pode prejudicar o acesso a memórias e que a velocidade de processamento diminui frequentemente, com possíveis implicações para diferentes domínios cognitivos – não independentes e com organização hierárquica, na qual os da base são de natureza sensorial básica e os do topo relacionados com funcionamento executivo e de controlo cognitivo.<sup>7-9</sup>

Para além da idade, outras características demográficas estão também relacionadas com a performance cognitiva, nomeadamente: a escolaridade - na medida em que maior escolaridade e sucesso escolar estão associados a um menor risco de demência e criação de reservas cognitivas, respetivamente, podendo, acumularem-se diferentes alterações antes de se manifestarem sintomas cognitivos; e o género - mais intimamente relacionado com condições de vida e oportunidades de educação durante os seus anos de formação.<sup>7,10,11</sup>

Assim, a identificação atempada de défices cognitivos é crucial para determinar a intervenção terapêutica, o treino cognitivo e reabilitação funcional de pacientes com doenças neurodegenerativas, demência e pacientes jovens com doenças do sistema nervoso central, como é o caso da esclerose múltipla.<sup>12</sup>

Para além da evidência já destacada, a investigação associada à performance cognitiva encontra-se ainda limitada quanto ao uso de medições repetidas que permitam identificar tendências longitudinais, quer como formas de diagnóstico para identificação de indivíduos em declínio, quer para melhor compreender padrões de alterações cognitivas. Não obstante às vantagens já enumeradas, e sendo que o diagnóstico de algumas doenças do sistema nervoso central implica o registo de diferenças significativas com o tempo, alguns autores sugerem que, em vez de ser feita de um modo pontual, a avaliação da performance cognitiva deva ser feita de forma longitudinal, recomendada também para aumentar a própria fiabilidade do diagnóstico desses pacientes.<sup>12-14</sup>

### 2.1.2 Estado da arte

As primeiras ferramentas disponibilizadas para avaliação da performance cognitiva providenciam uma avaliação neuropsicológica extensiva e pontual, sendo conduzidas por um profissional treinado (ex. *Mini Mental State Examination (MMSE)*). Estas mesmas metodologias possuem elevada sensibilidade e especificidade para detecção de demência, sendo, no entanto, a sua aplicação dispendiosa, quer a nível de tempo, quer a nível de recursos. Estas características tornam-nas inviáveis para aplicação na população em geral como metodologia de monitorização da função cognitiva, pelo que surgem alternativas menos dispendiosas, mas com o senão de perderem capacidade discriminativa para prever a progressão de demência ou de identificação de défice cognitivo ligeiro, para além de continuarem a requerer um avaliador externo e de não estarem validados para a prática de testes repetidos.<sup>12,13</sup>

Testes computadorizados com potencial para ultrapassar estas barreiras estão já disponíveis. Para além da redução de custos, permitem ainda uma medição exata com armazenamento de respostas, a minimização da subjetividade do examinador externo e a criação de diferentes versões e sequências para se mitigar o efeito de aprendizagem. Apesar de requererem experiência prévia com computadores para a realização destes tipos de testes e de ser, por vezes, apontada falta de adequabilidade das medições psicométricas, alguns dos testes demonstraram já fiabilidade e capacidade discriminatória relativamente aos testes tradicionais (ex. *Cognitive Stability Index*).<sup>12,13,15,16</sup>

Sendo, no entanto, concebidos de modo a replicar os testes tradicionais, a maioria destes não permite o diagnóstico ou monitorização de défice cognitivo. De modo a ultrapassar este último inconveniente, existem também outras ferramentas alternativas que têm já essa funcionalidade, continuando a necessitar igualmente de um profissional de saúde treinado para que o teste possa ser iniciado e a não existir nenhuma ferramenta destinada a utilização contínua/ longitudinal (ex. *MicroCog*).<sup>12,13,17</sup>

Assim, os testes autónomos via web surgem como uma alternativa vantajosa que permitiria a sua repetição periódica com menores custos a partir de casa, a diminuição do efeito de aprendizagem (por introdução de elementos aleatórios), a adaptação da dificuldade do teste à performance cognitiva do paciente e, não menos importante, o diagnóstico de pacientes e a monitorização da sua performance cognitiva.<sup>12,13</sup>

## 2.2 Brain on Track – BoT

Brain on Track (BoT) é um teste cognitivo computadorizado via web desenvolvido não para replicar os testes tradicionais, mas antes para ser executado autonomamente, a partir de casa do paciente, e permitir o diagnóstico e monitorização de doenças cognitivas através de medições repetidas. Através de diferentes estudos, este teste demonstrou já boa reprodutibilidade, correlação significativa com os testes existentes, aptidão para identificar diferenças clinicamente relevantes no diagnóstico de demência e de défice cognitivo ligeiro e alta confiabilidade nos testes feitos a partir de casa.<sup>12,13</sup>

### 2.2.1 Desenvolvimento

Descrevendo sinteticamente o trabalho realizado por Ruano et al. (2016), como ponto de partida para a construção desta ferramenta, recorreram a uma plataforma online já existente – *CogWeb* - que disponibiliza mais de 60 exercícios de treino cognitivo com foco em diferentes domínios (e que demonstraram já aplicabilidade para diferentes idades e doenças com possibilidade de uso repetido e independente a partir dos computadores dos pacientes), permitindo a participação em programas de treino remotos a partir da casa do paciente<sup>12,13</sup>

Sendo a ferramenta BoT destinada a um uso repetido, foram introduzidos elementos aleatórios e sequências de modo a minimizar o efeito de aprendizagem que pudesse advir da repetição dos diferentes subtestes. Assim, para além de serem criadas diferentes versões de subtestes, criaram-se ainda exercícios com diferentes graus de dificuldade para que se pudesse realizar uma avaliação ajustada à performance cognitiva de cada paciente.<sup>12,13</sup>

Importa ainda realçar que cada um dos subtestes está associado a um conjunto de instruções (exibidas no ecrã e simultaneamente lidas por uma voz gravada) e tem uma duração de 2 minutos, incluindo-se neste intervalo de tempo a leitura das instruções. O resultado do subteste é dado pelo número de tarefas executadas de forma correta, podendo variar entre 0 e o máximo de tarefas que cada participante pode fazer corretamente dentro do tempo limite.<sup>12,13</sup>

### **a) Refinamento da ferramenta**

A ferramenta foi aprimorada em duas etapas distintas.

Numa primeira etapa (A), foram selecionados por um grupo de neuropsicólogos e neurologistas 9 subtestes, tendo alguns destes sofrido adaptações no que toca à utilização de teclado de modo a minimizar erros – passando então a depender unicamente da interface “rato” ou da seleção de teclas únicas. Para determinar se a ferramenta tinha capacidade de detetar alterações do estadió saudável para estadios iniciais de doenças cognitivas, a ferramenta foi então aplicada em dois grupos distintos: um grupo de participantes doentes – referenciados a uma clínica de memória e com diagnóstico de défice cognitivo ligeiro ou demência ligeira – e participantes controlo – recrutados a partir de centros de aprendizagem para adultos da região do hospital do estudo, familiares dos pacientes, etc. Na sua generalidade, todos os participantes tinham idade igual ou superior a 18 anos e nenhuma deficiência que impossibilitasse ou condicionasse o uso do computador para fins deste teste. Para além do diagnóstico inicial, um neurologista procedeu à confirmação do diagnóstico num follow-up de pelo menos 6 meses, incluindo análises sanguíneas para causas tratáveis de demência, estudos de imagiologia e avaliação neuropsicológica completa. Finalmente, esta etapa A foi conduzida em ambiente de clínica hospitalar e sob observação de um membro da equipa de investigação, tendo sido registadas quaisquer dificuldades sentidas em perceber ou executar tarefas por parte dos participantes.<sup>12,13</sup>

Numa segunda etapa (B), e depois de analisados os resultados da etapa anterior (A), mantiveram-se 7 dos subtestes inicialmente selecionados e adicionaram-se 5 novos subtestes. A etapa B foi conduzida segundo o mesmo protocolo de estudo, configurações e critérios de inclusão da primeira etapa.<sup>12,13</sup>

Constatou-se então que entre os dois grupos de participantes não existiam diferenças significativas na idade e escolaridade e que, a partir da análise de componentes principais, foi possível identificar uma única componente. Tendo sido predefinidos os valores de fator loading dos subtestes a reter - valor superior a 0.50 -, foram excluídos 3 (entre as duas etapas) e, por avaliação da sua consistência interna, 1 subteste foi também descartado (na etapa B). Foi ainda comprovado por avaliação da correlação dos resultados do BoT com resultados do Montreal Cognitive Assessment (MoCA) e do Mini Mental State Examination (MMSE) que estes tinham uma relação positiva estatisticamente significativa. Por análise de regressão linear foi possível identificar uma associação significativa entre os resultados do BoT e os fatores idade e escolaridade e ainda uma interação entre o grupo de teste e o fator escolaridade, na primeira etapa –

diferenças de resultados maiores entre os grupos de participantes de maior escolaridade. Finalmente, avaliou-se ainda a exatidão do BoT em ambas as etapas e verificou-se que existiam subtestes mais difíceis de executar no primeiro momento, tendo sido substituídos por alternativas mais simples ou mesmo excluídos. <sup>12,13</sup>

### **b) Avaliação da fiabilidade do teste à distância**

Depois de aprimorada a ferramenta de teste BoT, finalmente com 8 subtestes, a mesma foi utilizada para avaliar a fiabilidade dos resultados quando o teste é repetido a partir de casa, tendo sido selecionada para o efeito uma subamostra de participantes da coorte EpiPorto – reunida entre 1999 e 2003, como representativa da população adulta de moradores do Porto. Os critérios de inclusão consistiam num resultado de teste MoCA, ajustado para a idade e escolaridade da população portuguesa (de modo a excluir pacientes com défices cognitivos), acesso a um computador no domicílio e capacidade de utilização do mesmo e da interface “rato” sem ajuda externa. <sup>12,13</sup>

Inicialmente, o teste BoT foi conduzido sob supervisão de um elemento da equipa de investigação, para ensinar ao participante como fazer login na página web da BoT e garantir que o mesmo compreendia as instruções e mecanismos de cada tarefa, mitigando possíveis efeitos de aprendizagem em respostas posteriores ao teste inicial. O participante repetiu o teste 1 semana depois, a partir de sua casa, e, após novo contacto pela equipa de investigação, voltou a repeti-lo 3 e 6 meses depois. <sup>12,13</sup>

Por análise da consistência dos resultados de cada subteste nas repetições consecutivas, foi possível observar que apenas 1 subteste não produziu resultados consistentes. Por utilização do teste T para amostras emparelhadas foi também possível constatar que 2 subtestes produziam resultados significativamente diferentes entre as duas últimas repetições do teste, podendo evidenciar efeitos de aprendizagem. <sup>12,13</sup>

### **2.2.2 Aplicação**

No seguimento da investigação desenvolvida por Ruano et al. (2019), partiu-se do teste BoT melhorado pelo seu trabalho anterior e, paralelamente à expansão dos domínios avaliados e à criação de níveis de dificuldade dos subtestes, descreveu-se a variação da performance cognitiva, ao longo de 1 ano, em participantes com défice cognitivo ligeiro e participantes saudáveis (de uma subamostra da coorte EpiPorto). Procedeu-se também à

avaliação da capacidade desta versão melhorada discriminar os dois grupos de participantes num único teste e em repetições teste. Foram então adicionados novos subtestes e criados 3 diferentes níveis de dificuldade adaptada à performance cognitiva esperada do participante de acordo com a sua escolaridade, passando o teste BoT a ter uma duração total de 24 minutos.<sup>12,13</sup>

Como já referido, o teste foi conduzido em dois grupos de participantes, devendo estes, em ambos os casos, ter idade igual ou superior a 18 anos, acesso a computador em casa e ser capazes de usar o computador com a interface “rato” sem ajuda de terceiros. Os participantes com défice cognitivo ligeiro foram recrutados a partir de uma clínica de memória e foram submetidos a análises de imagiologia cerebral e testes laboratoriais, tendo os elementos de ambos os grupos sido submetidos a uma avaliação neuropsicológica composta por uma bateria de testes de performance cognitiva validados para a população portuguesa. No caso do grupo de participantes com défice cognitivo ligeiro estes tiveram ainda o seu diagnóstico reavaliado após 1 ano de follow-up. Depois de integrados no estudo, todos os participantes realizaram o teste BoT na clínica ou no laboratório de análises, autonomamente, sob observação de um elemento da equipa de investigação, com propósito semelhante ao do trabalho anterior. Decorrida 1 semana, foi pedido aos participantes que voltassem a repetir o teste a partir de casa e depois a cada 3 meses durante 1 ano.<sup>12,13</sup>

Entre os grupos de participantes foram comparados os resultados obtidos, bem como as variáveis idade e educação. Para permitir a comparação dos resultados ao longo do ano, recorreu-se a modelos lineares mistos, com ajuste pelo critério de máxima verosimilhança restrita, tendo-se incluído, a priori, as variáveis idade, educação e grupo no modelo. Avaliou-se a capacidade discriminatória do teste, com recurso a curvas cúbicas com nó fixo (definido com base no critério de informação bayesiano) e efeitos aleatórios para cada curva e constante (intercept).<sup>12,13</sup>

Desta análise constatou-se que o grupo de participantes com défice cognitivo ligeiro era globalmente mais velho, com menor escolaridade e com pior performance cognitiva. Verificou-se uma associação significativa entre maior idade e menor escolaridade no que respeita a menores médias de resultados de teste e uma tendência global de melhoria da performance com o tempo em ambos os grupos, ainda que com uma queda na performance depois de um aumento inicial (mais acentuada nos grupos de participantes com a patologia). Finalmente, atestou-se que a capacidade discriminatória do teste era maior aquando da sua utilização repetida comparativamente ao caso de utilização única.<sup>12,13</sup>



### 2.2.3 Vantagens e desvantagens

A partir dos estudos conduzidos, a ferramenta de teste BoT obteve bons resultados de consistência, usabilidade, reprodutibilidade, correlação positiva dos resultados com resultados de testes já validados e capacidade de identificar diferenças clinicamente significativas. A ferramenta apresenta ainda um potencial inclusivo, na medida em que foi utilizada no seu desenvolvimento inicial uma amostra com menor escolaridade do que o comumente reportado na literatura, o que, por outro, pode ser indicativo de uma futura limitação, no sentido em que poderá dificultar a sua generalização a populações mais escolarizadas. <sup>12,13</sup>

Das diferentes limitações associadas a testes computadorizados no geral, e, por essa razão, também inerentes à ferramenta BoT, destacam-se a possibilidade de não *compliance* com a ausência de supervisão e limitações relacionadas com características tecnológicas dos diferentes dispositivos dos pacientes (*hardware, software, velocidade de internet, etc.*). <sup>12,13</sup>

Também o efeito de aprendizagem não deve ser menosprezado, havendo a possibilidade deste ter tido impacto nos diferentes estudos, podendo ser minimizado através da familiarização do indivíduo com a tarefa (através de um primeiro teste sob observação), da existência de diferentes versões de teste e de um número e intervalo entre avaliações balanceado. <sup>12,13</sup>

A ferramenta BoT apresenta assim grande potencial, uma vez que, para além do que já foi demonstrado, é também uma ferramenta que permitiria a realização autónoma da avaliação sem supervisão, à distância, a partir da casa do paciente. O cariz inovador desta ferramenta é ainda enaltecido pela possibilidade de permitir avaliações adaptadas à performance cognitiva esperada para escolaridade do paciente. <sup>12,13</sup>

Não obstante, é necessário mais desenvolvimento, melhoria e validação para a utilização repetida da ferramenta BoT. Outros pontos destacados pelos autores dos estudos incidiam sobre o potencial para esta ser uma ferramenta adequada na identificação e monitorização da performance cognitiva em indivíduos idosos bem como ser uma alternativa para melhorar o direcionamento de pacientes para cuidados mais especializados. Adicionalmente, os autores apresentaram ainda a possibilidade de a mesma ferramenta poder ser ainda utilizada para identificar candidatos a tratamentos de Alzheimer pré-sintomáticos, contribuindo, conseqüentemente, para atrasar ou evitar a evolução para demência. <sup>12,13</sup>

Dito isto, como objetivo seguinte estabeleceu-se a necessidade de avaliar a capacidade da ferramenta BoT discriminar indivíduos de acordo com a sua performance cognitiva,

principalmente em concomitância com outras doenças, nomeadamente esclerose múltipla, de forma longitudinal.

## **2.3 Modelos**

No âmbito deste estudo e das atividades descritas no presente relatório, foram utilizados 2 tipos de modelos: modelos lineares de efeitos mistos e modelos de regressão logística. Destes modelos, os primeiros foram aplicados para extrair as trajetórias individuais da pontuação dos testes do Brain on Track (BoT) e os modelos de regressão logística para prever o défice cognitivo tendo como variáveis independentes os efeitos aleatórios das referidas trajetórias individuais. Para avaliar a capacidade de discriminação dos modelos recorreu-se à área sob a curva ROC (AUC).

### **2.3.1 Modelos lineares de efeitos mistos**

#### **a) Introdução**

Os estudos longitudinais – estudos em que a variável resposta (*outcome*) é medida de forma repetida - são um dos diferentes tipos de estudo clínicos. Os estudos clínicos podem ser classificados em Experimentais (randomizados ou não randomizados) ou Observacionais. Estes últimos podem-se dividir em estudos Caso-controlo (com medição retrospectiva de possíveis preditores de uma dada variável dicotómica medida num determinado momento) ou Coorte (retrospectivos, prospetivos ou transversais/ “cross-sectional”).<sup>18</sup>

Estudos experimentais longitudinais têm por objetivo analisar o efeito de uma ou mais intervenções no outcome em estudo, sendo considerados por alguns autores o único desenho epidemiológico capaz de avaliar as relações de causalidade. Por outro lado, estudos longitudinais observacionais focados no desenvolvimento individual dos participantes são influenciados por 3 fatores distintos (idade, período – timing ou momento da medição - e coorte de nascimento – grupo de sujeitos nascidos no mesmo ano). Um problema comum nestes estudos é a distinção do efeito da idade dos efeitos dos restantes fatores - frequentemente relacionados com mudanças na metodologia de medição ao longo do tempo ou condições ambientais num dado momento de medição –

devendo-se ser sempre cuidadosos com a interpretação de resultados relacionados com a idade. Adicionalmente, medições individuais repetidas podem ser também influenciadas por alterações na atitude perante a medição em si – efeito de aprendizagem -, enquanto que a análise destes dados pode ser enviesada por um baixo grau de reprodutibilidade ou por ausência de alguns dados. <sup>18,19</sup>

Alguns métodos de análise mais tradicionais reduzem a abordagem longitudinal à análise de relações entre alterações em diferentes parâmetros, em dois momentos temporais. Outros métodos analisam a relação longitudinal entre diferentes variáveis através de uma única medição, no fim do período longitudinal, do *outcome* em estudo. Neste sentido, uma das limitações apontadas às metodologias anteriormente referidas são a não utilização de toda a informação longitudinal que dispõem. <sup>18,19</sup>

Como alternativa surge a utilização de curvas individuais de regressão com o tempo para estudo da relação entre o *outcome* em estudo e as variáveis preditoras, sendo o pressuposto de relação linear entre *outcome* e tempo a maior desvantagem. Relações longitudinais podem ser bastante complicadas pelo que diferentes análises devem ser utilizadas na avaliação dos seus diferentes aspetos. <sup>18,19</sup>

Uma das metodologias mais utilizadas em análise longitudinal de dados é a dos modelos lineares de efeitos mistos.

## **b) Aplicação**

Como já referido anteriormente, para obter as trajetórias individuais das pontuações dos testes BoT ao longo do tempo recorreu-se aos modelos lineares de efeitos mistos.

Em determinadas condições os modelos lineares de efeitos mistos são também referidos como modelos multinível de coeficientes aleatórios e surgem como modelos de análise que permitem a variação dos coeficientes de regressão entre sujeitos, de modo a corrigir a dependência verificada entre observações longitudinais de um mesmo indivíduo (correlacionadas entre si). <sup>18,19</sup>

Em comparação com as metodologias tradicionais, é permitida nestes modelos a variação da constante (*intercept*) entre sujeitos (fórmula 1), sendo nas metodologias tradicionais o mesmo valor fixo. Também possível são o caso de um *intercept* fixo com variação do desenvolvimento das próprias variáveis com o tempo – declive (*slope*) – (fórmula 2) ou a combinação de *intercept* e *slope* aleatórios (fórmula 3). Nas fórmulas mencionadas anteriormente,  $Y_{it}$  são as observações para o sujeito  $i$  no momento  $t$ ,  $\beta_0$  o *intercept*,  $\beta_{0i}$  o *intercept* aleatório,  $t$  o tempo,  $\beta_1$  o coeficiente de regressão para o tempo,

$\beta_{1i}$  o coeficiente de regressão aleatório para o tempo e  $\varepsilon_{it}$  o erro do sujeito  $i$  no momento  $t$ .<sup>18,19</sup>

$$(1) \quad Y_{it} = \beta_{0i} + \beta_{1t} + \varepsilon_{it}$$

$$(2) \quad Y_{it} = \beta_0 + \beta_{1it} + \varepsilon_{it}$$

$$(3) \quad Y_{it} = \beta_{0i} + \beta_{1it} + \varepsilon_{it}$$

A par deste processo são ainda utilizados procedimentos de estimação (ex.: Máxima Verossimilhança – estimativas que maximizam a função de verossimilhança, consistentes e assintoticamente eficientes, e Máxima Verossimilhança Restrita).

Adicionalmente, a identificação do modelo mais parcimonioso e com melhor ajuste aos dados é feita maioritariamente com base nos critérios de informação com ponderação do número de parâmetros estimados (ex.: Critério de Informação de Akaike - AIC) e/ou utilizando o teste de Razão de Verossimilhanças entre modelos aninhados. A fórmula 4 apresenta a estatística do teste de Razão de Verossimilhança, onde  $l_1$  e  $l_2$  correspondem aos valores a comparar dos modelos com maior número de parâmetros e o reduzido, respetivamente.<sup>18,19</sup>

$$(4) \quad G^2 = 2(l_1 - l_2)$$

Outra alternativa aos modelos lineares de efeitos mistos é a análise GEE (equações de estimação generalizada). Em ambas as metodologias são feitas correções à dependência de observações do mesmo sujeito, sendo que no caso das GEE é especificada uma determinada estrutura de correlação, enquanto nos modelos de efeitos mistos é permitido que os coeficientes de regressão variem entre sujeitos. É também possível adicionar uma determinada estrutura de correlação aos modelos lineares de efeitos mistos, de forma a completar a correção feita pelos coeficientes aleatórios, no entanto isto deve ser feito de modo cauteloso sob perigo de hipercorreção.<sup>18,19</sup>

Quanto a ter em conta a heteroscedasticidade, os modelos lineares de efeitos mistos apresentam uma vantagem, em comparação com a análise GEE, na medida em que consideram, parcialmente, esta questão ao permitirem um declive (*slope*) aleatório, o que pode, no entanto, não ser suficiente. É também possível encontrar exemplos que tomam

a variável tempo como uma variável contínua, o que, aleado à modelação linear da variável *outcome* ao longo do tempo, surge como desvantajoso uma vez que nem sempre isto se verifica. Nesse sentido, e para além das alternativas de modelação como função quadrática ou cúbica, é ainda sugerido que o tempo seja codificado como variável categórica. É de notar que esta categorização é apenas válida quando o número de medições repetidas (para o mesmo sujeito) é baixo em comparação com o tamanho da amostra, para além de que, em caso de intervalos de tempo desiguais para os sujeitos ou de falta de observações (diferentes números de medições repetidas), esta também se torna uma limitação maior. <sup>18,19</sup>

Importa ainda referir que, antes de utilizar um modelo obtido para fazer inferências acerca dos dados, deve-se proceder à validação dos pressupostos, mais comumente feito através da visualização de gráficos dos resíduos, valores ajustados e estimativas dos efeitos aleatórios ou por aplicação de testes de hipóteses. Recuperando o que já foi dito anteriormente, deve-se avaliar o pressuposto de que os erros intra-grupo são independentes e normalmente distribuídos, com média 0 e variância  $\sigma^2$ , e independentes dos efeitos aleatórios. Adicionalmente, deve-se verificar também se é cumprido o pressuposto de que os efeitos aleatórios são normalmente distribuídos, com média 0 e matriz de covariância  $\Psi$  (independente do grupo), e que estes são independentes para grupos diferentes. <sup>18,19</sup>

### c) Exemplo

Com recurso à biblioteca “nlme” do *software R*, apresenta-se de seguida um exemplo da construção de um modelo linear misto com recurso ao comando *lme*, que por defeito utiliza a estimação de Máxima Verosimilhança Restrita (uma variante ao método de Máxima Verosimilhança), apresentando-se tipicamente como na Figura 2.1. No primeiro argumento (“fixed”) são especificados os efeitos fixos do modelo por meio de uma fórmula linear tipicamente de 2 lados e a variável *outcome*, no segundo (“data”) é especificada a data frame a que o modelo será relativo e, finalmente, no terceiro os efeitos aleatórios através de uma fórmula de, normalmente, 1 lado – para o qual, se omitido, são assumidos, por defeito, as variáveis especificadas no lado direito do argumento “fixed”. O exemplo apresentado nesta secção é um dos modelos construídos no seguimento das atividades desenvolvidas no estágio.

```
lme( fixed, data, random )
```

**Figura 2.1** Linha de comando lme.

**Fonte:** Pinheiro & Douglas, 2000

Para o efeito, modelaram-se inicialmente as pontuações totais dos pacientes (removendo-se os controlos, “m\_final2[grupo2!=0,]”) através de um modelo não ajustado. É possível observar as estimativas dos resultados do comando *lme*, nos quais se apresentam as estimativas do desvio-padrão e as correlações dos efeitos aleatórios especificados, o erro padrão dentro dos dados de grupo (“Utilizadores”) e os efeitos fixos.

```
mmisto1<-lme(total_tscores~tempo*factor(grupo2)+I(tempo^2)+
I(tempo^2)*factor(grupo2),random=~1+tempo+I(tempo^2)|Utilizadores,
m_final2[grupo2!=0,])
mmisto1

##
##Linear mixed-effects model fit by REML
## Data: m_final2[grupo2 != 0, ]
## Log-restricted-likelihood: -3692.252
## Fixed: total_tscores ~ tempo * factor(grupo2) + I(tempo^2) +
##I(tempo^2) *      factor(grupo2)
##
##      (Intercept)          tempo      factor(grupo2)2
## 49.702394157          0.777773855          -7.017393345
##
##      I(tempo^2)      tempo:factor(grupo2)2  factor(grupo2)2:I(tempo^2)
## -0.024302018          -0.174360818          0.008284573
##
##Random effects:
## Formula: ~1 + tempo + I(tempo^2) | Utilizadores
## Structure: General positive-definite, Log-Cholesky parametrization
##           StdDev   Corr
## (Intercept) 8.2832048 (Intr) tempo
## tempo       0.7394950 -0.220
## I(tempo^2)  0.0263698  0.263 -0.972
##Residual    3.8623517
##
##Number of Observations: 1245
##Number of Groups: 75
```

Caso se pretenda adicionar alguma variável ou mudar o método de estimação pode-se recorrer ao comando *update*, tendo no presente exemplo sido utilizado para ajustar as pontuações às três variáveis adicionadas, sendo que para se obter mais detalhes se utiliza a função *summary*.

```
mmisto2<-update(mmisto1,~tempo*factor(grupo2)+I(tempo^2)*factor(grupo2)
+ idade_ini+Género.x+escol.cut)
summary(mmisto2)
```

```
##
##Linear mixed-effects model fit by REML
## Data: m_final2[grupo2 != 0, ]
##      AIC      BIC    logLik
## 7385.236 7472.256 -3675.618
##
##Random effects:
## Formula: ~1 + tempo + I(tempo^2) | Utilizadores
## Structure: General positive-definite, Log-Cholesky parametrization
##              StdDev      Corr
## (Intercept) 6.95745226 (Intr) tempo
##tempo        0.72167885 -0.096
##I(tempo^2)   0.02517329  0.121 -0.970
##Residual    3.86620286
##
##Fixed effects: total_tscores ~ tempo + factor(grupo2) + I(tempo^2) +
##idade_ini + Género.x + escol.cut + tempo:factor(grupo2) +
##factor(grupo2):I(tempo^2)
##              Value Std.Error  DF  t-value p-value
## (Intercept)  47.12603  8.742646 1166  5.390363  0.0000
##tempo        0.77019  0.138025 1166  5.580049  0.0000
##factor(grupo2)2 -8.05928  1.700570  69 -4.739163  0.0000
##I(tempo^2)    -0.02385  0.005182 1166 -4.602369  0.0000
##idade_ini    -0.36916  0.092975  69 -3.970556  0.0002
##Género.xMasculino  4.22537  1.714960  69  2.463828  0.0162
##escol.cut(5,10] 14.29503  7.454032  69  1.917758  0.0593
##escol.cut(10,30] 16.09030  7.452388  69  2.159080  0.0343
##tempo:factor(grupo2)2 -0.16004  0.198257 1166 -0.807223  0.4197
##factor(grupo2)2:I(tempo^2) 0.00750  0.007518 1166  0.997897  0.3185
## Correlation:
##              (Intr) tempo  fc(2)2 I(t^2) idad_n GnrxMs
##e.(5,1 e.(10, t:(2)2
##tempo        -0.026
##factor(grupo2)2  0.016  0.135
##I(tempo^2)      0.024 -0.945 -0.134
##idade_ini     -0.548  0.001  0.007  0.002
##Género.xMasculino -0.156  0.000 -0.117  0.000 -0.072
##escol.cut(5,10] -0.874  0.000 -0.121  0.001  0.139  0.106
##escol.cut(10,30] -0.909 -0.001 -0.129  0.003  0.188  0.144
##0.967
##tempo:factor(grupo2)2  0.018 -0.696 -0.196  0.658 -0.001  0.003
##0.000  0.001
##factor(grupo2)2:I(tempo^2) -0.017  0.651  0.194 -0.689  0.002 -0.003 -
##0.001 -0.002 -0.944
##
##
##
```

```
##Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
##-9.80082300 -0.42398954  0.06293804  0.49210190  5.88530024
##
##Number of Observations: 1245
##Number of Groups: 75
```

Observa-se então que, tendo por referência um valor de alfa ( $\alpha$ ) 0,05, a interação com a variável tempo (“tempo:factor(grupo2)” e “tempo:factor(grupo2)2”) não é significativa (valores p superiores a 0,05), assinalando-se a existência de diferenças significativas mediante o género (“Género.xMasculino”), a idade à entrada no estudo (“idade\_ini”) e a escolaridade categorizada dos indivíduos (“escol.cut(10,30]”) no que toca às pontuações totais das repetições do teste BoT. Os resíduos de resposta podem ser extraídos com recurso ao comando *resid*, sendo necessário especificar-se no argumento *type* do mesmo comando outro tipo de resíduos que se pretenda obter, como, por exemplo, os de Pearson (introduzindo-se “pearson”).

Adicionalmente, é também possível comparar diferentes objetos *lme* por meio do comando *anova*. Como no presente exemplo os 2 modelos apenas diferem nos efeitos fixos deve ser utilizado o método de estimação de Máxima Verosimilhança (“method=“ML””). Por observação de um valor p inferior a 0,05 para o teste da Razão de Verosimilhanças, constata-se que os modelos são significativamente diferentes e que o modelo 2 se ajusta melhor aos dados do que o modelo 1 (tendo em conta os valores de critérios de informação menores).

```
mmisto1ml<-update(mmisto1, method="ML")
mmisto2ml<-update(mmisto2, method="ML")
anova.lme(mmisto1ml,mmisto2ml)

##
##      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
##mmisto1ml   1 13 7389.356 7456.006 -3681.678
##mmisto2ml   2 17 7372.036 7459.193 -3669.018 1 vs 2 25.32082 <.0001
```



### 2.3.2 Regressão logística

#### a) Aplicação

No seguimento da análise estatística realizada, utilizou-se também um modelo de regressão logística no qual a variável dependente era dicotómica (assumindo o valor 0 ou 1) e correspondia a uma de duas categorias que distinguem os elementos da amostra consoante o seu défice cognitivo (se presente ou não). Este modelo de regressão permite estimar o valor esperado da variável dependente condicionado pelos valores registados das variáveis independentes definidas, sendo o modelo definido por

$$(5) \quad y = \pi(x) + \varepsilon$$

no qual  $\varepsilon$  representa o erro, com distribuição de média  $E[\varepsilon] = 0$  e variância  $V[\varepsilon] = \pi(x)[1 - \pi(x)]$ , correspondendo  $\pi(x)$  à probabilidade de ser um indivíduo codificado com o valor 1 da variável dicotómica. A estimativa dos coeficientes de regressão é feita com recurso ao método de máxima verosimilhança, podendo ser aplicados outros métodos como o processo iterativo dos mínimos quadrados pesados.<sup>20</sup>

Neste tipo de regressão é utilizada também uma função *link*, mais frequentemente a transformação *logit*, podendo ser utilizadas outras como a função *probit* ou a *cauchit*, que permitem que a relação entre regressores e a variável resposta seja modelada por uma regressão linear.<sup>20</sup>

A transformação *logit*, em particular, aplica o logaritmo ao Rácio de Verosimilhança (razão entre probabilidade de sucesso e de insucesso,  $\ln\left(\frac{\pi(x)}{1-\pi(x)}\right)$ ), permitindo obter as estimativas dos coeficientes dos regressores

$$(6) \quad \text{logit}(\pi(x)) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

em que  $p$  corresponde ao número de regressores incluídos.<sup>20</sup>

A regressão logística permite ainda uma fácil interpretação das estimativas obtidas ao recorrer à medida de associação *Odds Ratio (OR)*, calculadas de acordo com  $OR = e^{\beta_i}$ , onde  $\beta_i$  corresponde à estimativa do coeficiente da variável dependente  $i$ .<sup>20</sup>

Depois de obtidas as estimativas dos coeficientes de regressão, estas são avaliadas quanto à sua significância, podendo-se aplicar o Teste de Wald, para avaliação de um

único coeficiente, e o do Teste de Razão de Verossimilhanças, para avaliação da significância de todo o modelo comparativamente a um modelo reduzido (apenas com  $\beta_0$ ).<sup>20</sup>

Adicionalmente, para avaliação da qualidade do ajuste do modelo aos dados, pode-se aplicar o Teste de Ajustamento do Modelo que testa as hipóteses do modelo se ajustar (H0) ou não (H1) aos dados.<sup>20</sup>

Não obstante, outras formas de avaliar a qualidade do modelo passam pela avaliação do efeito do modelo com recurso ao pseudo- $R^2$ , podendo-se calcular o  $R^2$  de Cox e Snell – que assume sempre valores inferiores a 1 -, o  $R^2$  de Nagelkerke – uma versão ajustada do anterior, variando no intervalo [0,1] -, ou o  $R^2$  de McFadden. Finalmente, pode-se ainda recorrer à Curva ROC (Receiver Operating Characteristic), cuja área sob a mesma quantifica a capacidade do modelo discriminar os indivíduos de acordo com um dada característica de interesse, variando este valor entre 0 e 1 (capacidade discriminante total). Quanto mais o valor da área sob a curva ROC (AUC) se aproxima de 1, maior a sua capacidade discriminante, sendo que, segundo Hosmer & Lemeshow (2000):

- se  $AUC = 0.5$ , sugere não ter capacidade de discriminação;
- se  $0.7 \leq AUC < 0.8$ , considera-se ter uma capacidade de discriminação aceitável;
- se  $0.8 \leq AUC < 0.9$ , considera-se ter uma capacidade de discriminação excelente; e
- se  $0.9 \leq AUC$ , considera-se ter uma capacidade de discriminação extraordinária.<sup>20</sup>

## b) Exemplo

Apresenta-se de seguida, como exemplo de aplicação de um modelo de regressão logística, um dos modelos construídos no seguimento das atividades desenvolvidas no estágio. Para o efeito, depois de modeladas as pontuações totais dos pacientes (removendo-se os controlos, "m\_final2[grupo2!=0,]") por meio de um modelo linear de efeitos mistos, foram extraídas as estimativas dos efeitos aleatórios para cada indivíduo, recorrendo ao comando *ranef*, e construídas novas *data frames*, como visível nas linhas de código seguintes.

```
mmisto<-lme(total_tscores~tempo+I(tempo^2),random=
~1+tempo+I(tempo^2)|Utilizadores,m_final2[grupo2!=0,])
dt1<-data.frame(ranef(mmisto))
dt1$Utilizadores<-row.names(dt1)
b4f<-merge(b4,dt1, by="Utilizadores")
```

De seguida, recorrendo ao comando *glm*, define-se o modelo logístico, especificando-se a fórmula - com a variável resposta, que deverá assumir valores de 0 ou 1 ("I(b4f\$grupo2-1)"), e as variáveis independentes no segundo termo -, a distribuição do erro e a função *link* a ser usada (por exemplo, "family="binomial" (link="probit)") e a *data frame* em que se encontram os dados a ser modelados.

```
mlog<-glm(I(b4f$grupo2-1) ~ X.Intercept.+tempo+I.tempo.2.,
data=b4f[b4f$grupo2!=0,], family = "binomial")
summary(mlog)

##
## Call:
## glm(formula = I(b4f$grupo2 - 1) ~ X.Intercept. + tempo + I.tempo.2.,
##      family = "binomial", data = b4f[b4f$grupo2 != 0, ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1077  -0.9415   0.2648   1.0386   1.6490
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.05544    0.25851   0.214  0.83019
## X.Intercept.  -0.13040    0.03942  -3.308  0.00094 ***
## tempo          2.90416    3.24562   0.895  0.37090
## I.tempo.2.    107.60200   99.21145   1.085  0.27811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 103.959 on 74 degrees of freedom
## Residual deviance: 87.662 on 71 degrees of freedom
## AIC: 95.662
##
## Number of Fisher Scoring iterations: 4
```

O comando *summary* permite obtenção de informação mais detalhada. Sendo possível obter métricas referentes aos resíduos, estimativas dos coeficientes com as respectivas estatísticas de teste (no presente exemplo, apenas os efeitos aleatórios da constante apresentam significância estatística – para “X.Intercept.”, valor  $p < 0,001$ ), valores de desviância (utilizados no Teste de Ajustamento do Modelo) e uma versão do AIC. O Teste de Ajustamento do Modelo pode ser realizado da forma apresentada nas linhas de código abaixo, concluindo-se que, neste caso, dado que valor  $p > 0,05$ , o modelo ajusta-se aos dados.

```
p_value_D<-1-pchisq(mlog$deviance,mlog$df.residual)
p_value_D

##
## 0.08743378
```

Quanto ao teste de Razão de Verossimilhanças, que permite avaliar a significância de todo o modelo, este pode ser realizado utilizando o comando *lrtest*, no qual se inserem os objetos dos modelos a comparar. Para esse efeito é necessário que se crie um modelo reduzido apenas com a constante (“mlog\_Red”). Como seria de esperar depois de se constatar que existe um termo significativo, o modelo em estudo é significativamente diferente do modelo reduzido (valor  $p = 0,001$ ).

```
mlog_Red <- glm( I(grupo2-1) ~ 1, data=b4f[b4f$grupo2!=0,],
family="binomial")
LR_test<-lrtest(mlog_Red,mlog)
LR_test

## Likelihood ratio test
##
## Model 1: I(grupo2 - 1) ~ 1
## Model 2: b4f$grupo21 ~ X.Intercept. + tempo + I.tempo.2.
## #Df LogLik Df Chisq Pr(>Chisq)
## 1 1 -51.979
## 2 4 -43.831 3 16.297 0.0009857 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Recorrendo à biblioteca “DescTools” e concretamente ao comando *PseudoR2* é possível obter estimativas dos pseudo-R<sup>2</sup> definidos no argumento “*which*”.

```
library(DescTools)
PseudoR2(mlog, which = c("McFadden", "CoxSnell", "Nagelkerke"))

##   McFadden   CoxSnell Nagelkerke
## 0.1567615  0.1953032  0.2604197
```

Finalmente, de modo a avaliar a capacidade do modelo discriminar os pacientes de acordo com a sua performance cognitiva (apresentando ou não défice cognitivo), definiu-se a curva ROC recorrendo-se ao comando *roc* especificando-se a variável resposta e os valores preditores (neste caso dados pela resposta do modelo logarítmico definido, “*predict(mlog, type="response")*”). Através dos comandos *auc* e *ci.auc* é possível obter as estimativas pontuais e intervalares, respetivamente, da área sob a curva ROC (AUC), curva essa que poderá ser ainda visualizada através do comando *plot* (Figura 2.2). Dadas as estimativas apresentadas, pode-se afirmar que o modelo possui uma capacidade de discriminação aceitável.

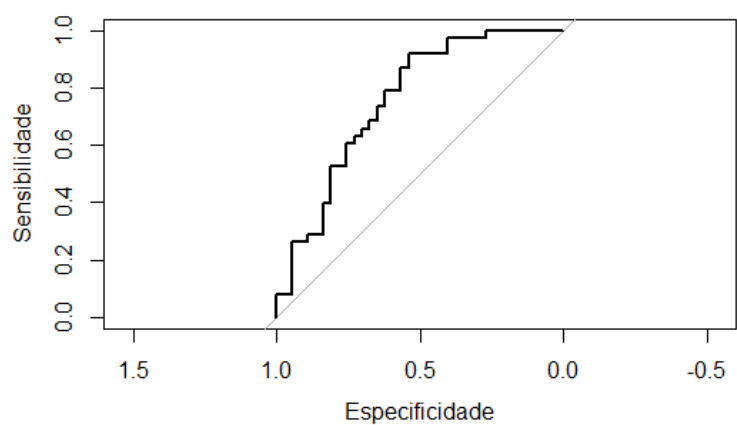
```
roc_2<-roc(b4f[b4f$grupo2!=0,]$grupo2,predict(mlog, type="response"))
auc(roc_2)

## Area under the curve: 0.7553

ci.auc(roc_2)

## 95% CI: 0.6441-0.8665 (DeLong)

plot(roc_2)
```



**Figura 2.2** Área sob a curva ROC (AUC), para o modelo logístico dos pacientes de acordo com a presença de défice cognitivo dado GD 2.

## Capítulo 3

### 3.1 Atividades realizadas em estágio

As atividades realizadas no decorrer do estágio não se centraram unicamente na análise dos dados que foram disponibilizados. Desde o início do estágio até ao seu término, tive oportunidade de, não só consolidar e aplicar conhecimentos já adquiridos, como adquirir novos conhecimentos relacionados com a performance cognitiva de pessoas com esclerose múltipla, o que, apesar de não ser algo relacionado diretamente com conhecimentos estatístico, se revelou estimulante. Adicionalmente, e referente ao *software* de análise estatístico utilizado (*R* versão 3.6.1), tive também oportunidade de desenvolver novos conhecimentos na medida em que recorri a novas bibliotecas do *software* (“*data.table*”, “*stringr*”, “*tree*”, “*lubridate*”, “*rpat*”, “*table1*” e “*rvest*”) que ainda não tinha utilizado e aprofundei conhecimentos acerca das funcionalidades de outras bibliotecas que já tinha utilizado anteriormente (“*ggplot2*”, “*dplyr*”, “*nlme*”, “*pROC*”, “*psych*”, “*lme4*”, “*DescTools*” e “*nFactors*”).

As principais atividades desenvolvidas durante o período do estágio foram

- Organização de três bases de dados distintas com criação dos respetivos dicionários de variáveis (Anexos A, B e C) e posterior
  - Identificação de outliers;
  - Cálculo de pontuações dos testes;
  - Construção de novas variáveis;
- Leitura de bibliografia (artigos e protocolos) e aquisição de conhecimentos referentes ao estudo desenvolvido e à ferramenta BoT;
- Leitura de bibliografia e aquisição de conhecimentos referentes a metodologia estatística e comandos aplicados na análise estatística;
- Reuniões mensais com o médico neurologista envolvido no estudo, Doutor Luís Ruano, e reuniões semanais com o supervisor do estágio, Doutor Milton Severo;
- Construção de tabelas de análises descritivas preliminares e finais;
- Replicação de tabelas de artigos de fases anteriores do estudo e de desenvolvimento da ferramenta;
- Construção de modelos lineares de efeitos mistos para os scores totais da ferramenta BoT de cada indivíduo incluído no estudo;

- Avaliação da capacidade discriminante dos modelos construídos por obtenção da respetiva curva ROC e da sua AUC;
- Escrita de um artigo científico de acordo com os resultados obtidos no estudo.

Para além dos conhecimentos que foram adquiridos, como já referido, a execução destas atividades forneceu-me um primeiro contacto com aquilo que se faz num instituto de investigação permitindo conhecer igualmente uma parte da dinâmica que a sua atividade envolve.

## **3.2 Análise estatística**

### **3.2.1 Enquadramento**

O principal objetivo do estudo passava por avaliar a aplicação da ferramenta da BoT na discriminação de pessoas com esclerose múltipla – pacientes - com e sem declínio cognitivo, com recurso a dados longitudinais.

Para esse efeito, foi constituída uma amostra de 99 pessoas, das quais 21 não tinham esclerose múltipla (controlos) e as restantes 78 eram pessoas com esclerose múltipla recrutadas em serviços hospitalares distintos.

Depois de recrutados, estes indivíduos foram sujeitos a uma bateria inicial de testes neuropsicológicos de modo a caracterizar a sua performance cognitiva, aplicada externamente por um especialista, e tiveram uma primeira sessão de treino com a ferramenta BoT de modo a mitigar eventuais dúvidas e dificuldades. Determinava-se que um indivíduo tinha défice cognitivo mediante apresentasse um grau de defeito (GD) igual ou superior a 2 em, pelo menos, 2 dos testes neuropsicológicos realizados na bateria inicial. Adiante serão ainda mostradas análises de dados em que a determinação de défice cognitivo é idêntica, sendo, no entanto, o valor de GD de referência igual ou superior a 3.

Depois dessa primeira sessão de treino, e ao longo de um período de 12 meses era pedido, periodicamente, ao indivíduo que realizasse uma nova repetição com a ferramenta BoT autonomamente, sendo que o mesmo era livre de fazer o número de repetições que quisesse. Ao fim do período de 12 meses, o indivíduo seria submetido a



uma nova bateria de testes neuropsicológicos semelhante àquela que foi realizada inicialmente.

Importa reiterar que as repetições realizadas com a ferramenta BoT – de agora em diante designadas “repetição de teste” – eram adaptadas à escolaridade do indivíduo mediante este possuísse menos de 5, entre 5 a 9 ou 10 ou mais anos de escolaridade.

### 3.2.2 Resultados

Numa primeira instância, depois de organizadas as bases de dados e revistos os dados, procedeu-se a uma análise descritiva da amostra, tendo-se verificado que, à semelhança do descrito anteriormente por Ruano et al. (2019), não existiam diferenças na idade e educação entre pacientes e controlos e que a proporção de mulheres era significativamente maior nos controlos (90,5%;  $p = 0.015$ ), como visível na Tabela 3.1. A prevalência inicial de défice cognitivo entre pacientes era de 43,6%, passando a 29,5% quando reavaliada ao fim de 12 meses, devendo-se ter em consideração que menos 22 pacientes foram avaliados.

**Tabela 3.1** Características da amostra e descrição por caso.

**Abreviaturas:** MoCA, Montreal Cognitive Assessment; BoT, Brain on Track; CI, Déficit Cognitivo; EDSS, Expanded Disability Status Scale; NA, não aplicável

<sup>a</sup> Média e (desvio-padrão) são indicados;

<sup>b</sup> Frequência absoluta e (percentagem) são indicadas.

	Total (n=99)	Controlos (n=21)	Pacientes (n=78)	P value
<b>Idade, anos</b> <sup>a</sup>	37.9 (8.79)	38.8 (7.78)	37.6 (9.1)	0.576
<b>Sexo</b> <sup>b</sup>				
Feminino	65 (65.7%)	19 (90.5%)	46 (59.0%)	0.015
Masculino	34 (34.3%)	2 (9.5%)	32 (41.0%)	
<b>Educação, anos</b> <sup>a</sup>	12.3 (3.68)	12.4 (3.64)	12.3 (3.7)	0.954
<b>MoCA inicial, score</b> <sup>a</sup>	24.6 (5.00)	NA	24.6 (5.0)	-
Missing <sup>b</sup>			1 (1.3%)	
<b>MoCA final, score</b> <sup>a</sup>	24.9 (3.4)	NA	24.9 (3.4)	-
Missing <sup>b</sup>			26 (33.3%)	
<b>CI inicial, prevalência</b> <sup>b</sup>	34 (34.3%)	NA	34 (43.6%)	-
Missing <sup>b</sup>			1 (1.3%)	
<b>CI final, prevalência</b> <sup>b</sup>	23 (23.2%)	NA	23 (29.5%)	-
Missing <sup>b</sup>			23 (29.5%)	
<b>EDSS inicial, score</b> <sup>a</sup>	1.75 (1.25)	NA	1.75 (1.25)	-
Missing <sup>b</sup>			1 (1.3%)	
<b>EDSS final, score</b> <sup>a</sup>	1.83 (1.26)	NA	1.83 (1.26)	-
Missing <sup>b</sup>			1 (1.3%)	

Como também já foi referido, a ferramenta BoT inclui diferentes exercícios – subtestes –, obtendo-se de cada um a respetiva pontuação. Assim, dada a necessidade de calcular a pontuação total de cada repetição de teste procedeu-se à soma das pontuações padronizadas de cada subteste (Z scores, recorrendo à média e desvio padrão dos scores de cada subteste da totalidade da amostra). Por sua vez, estas pontuações totais foram convertidas noutras por multiplicação por 10 e adição de 50 unidades de modo que a maioria dos valores fosse positivo e variasse entre 0 e 100 (T scores). Quanto a esta padronização, quando relativa aos subtestes “Attention task III” ou “Delayed verbal memory task”, era feita de modo semelhante, mas para cada um dos intervalos de anos de escolaridade a que os indivíduos pertenciam de acordo com o anteriormente estipulado.

A mediana (percentil 25; percentil 75) de repetições de testes completas por toda a amostra foi 13 (9;20) e quanto a repetições de testes incompletas as mesmas métricas foram 0 (0;1).

Na Tabela 3.2, observam-se as pontuações por subteste das primeiras 4 repetições de teste realizadas dos pacientes.

**Tabela 3.2** Pontuações, por subteste, para as primeiras 4 repetições do teste de pacientes com esclerose múltipla. Média e (desvio-padrão) são indicados;

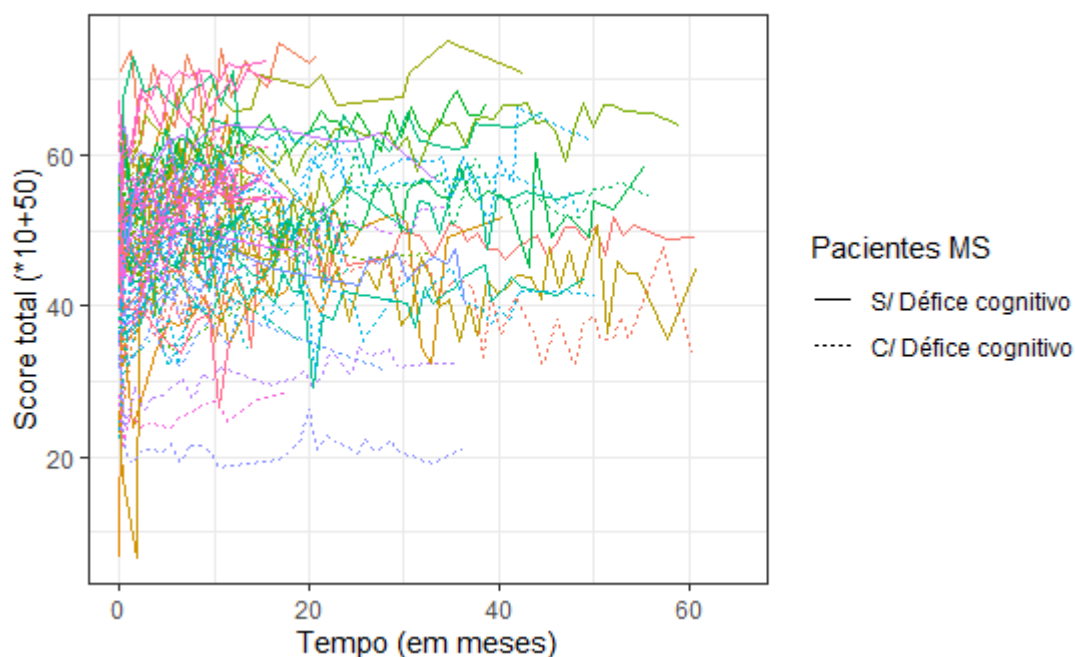
	Repetição 1	Repetição 2	Repetição 3	Repetição 4
<b>Word categories</b>	20.8 (8.2)	20.7 (9.5)	22.4 (9.4)	22.7 (8.9)
<b>Colour interference task</b>	19.5 (7.2)	21.7 (7.2)	21.8 (7.8)	23.1 (8.1)
<b>Opposite task</b>	56.0 (19.6)	59.2 (20.4)	61.9 (18.6)	64.9 (22.7)
<b>Written comprehension</b>	10.8 (5.4)	12.6 (6.5)	12.1 (5.2)	12.0 (5.4)
<b>Visual memory task II</b>	8.9 (4.3)	9.7 (4.1)	10.0 (3.8)	10.0 (4.1)
<b>Attention task III</b>	10.6 (3.5)	12.1 (4.1)	12.4 (3.7)	13.0 (3.9)
<b>Delayed verbal memory task</b>	15.6 (3.8)	16.0 (4.0)	16.5 (4.0)	15.8 (3.8)
<b>Calculus task</b>	13.3 (4.2)	13.9 (4.6)	14.2 (5.0)	13.4 (4.8)
<b>Puzzles</b>	5.8 (2.7)	6.3 (3.1)	6.2 (3.1)	6.0 (3.1)
<b>Verbal memory II</b>	8.8 (2.9)	9.3 (3.2)	9.4 (2.8)	9.7 (3.0)
<b>Sequences</b>	14.8 (7.6)	16.3 (8.4)	17.0 (8.3)	17.4 (9.0)

Na sua globalidade, as pontuações (scores) apresentam, para cada subteste, uma tendência crescente entre as repetições, o que pode ser sugestivo de efeitos de aprendizagem. Não obstante, há também subtestes cujos scores começam a diminuir a partir da segunda repetição, como no caso de “Written comprehension” e “Puzzles”, ou mesmo na quarta repetição apresentando valores semelhantes ao da primeira, como no caso de “Delayed verbal memory task” e “Calculus task”. Na Figura 3.1, estão representados todos os T scores para cada um dos pacientes com pelo menos 3 repetições de teste.

Adicionalmente, foi também realizada uma análise de componentes principais de modo a avaliar a dimensionalidade de cada subteste (Tabela 3.3) e a sua correlação com o score do teste. De modo a avaliar a sua consistência interna, calculou-se um alfa de Cronbach de 0,84 (sugestivo de elevada consistência) que não poderia ser aumentado por remoção de nenhum dos subtestes. A variância explicada pela componente principal foi 58% e todos os subtestes apresentavam factor loadings superiores a 0.60.

**Tabela 3.3** Análise de componentes principais para a primeira repetição de teste  
**Abreviaturas:** SD, Desvio-padrão

	Pontuações, média (SD)		PCA
	MS patients	Controls	
<b>Word categories</b>	20.8(8.2)	21.9(6.6)	0.83
<b>Colour interference task</b>	19.5(7.2)	21.9(6.9)	0.82
<b>Opposite task</b>	56.0(19.6)	53.7(21.8)	0.76
<b>Written comprehension</b>	10.8(5.4)	10.8(5.0)	0.73
<b>Visual memory task II</b>	8.9(4.3)	10.6(3.3)	0.72
<b>Attention task III</b>	10.6(3.5)	12.8(2.6)	0.66
<b>Delayed verbal memory task</b>	15.6(3.8)	16.6(3.9)	0.61
<b>Calculus task</b>	13.3(4.2)	14.8(3.3)	0.75
<b>Puzzles</b>	5.8(2.7)	5.7(2.3)	0.80
<b>Verbal memory II</b>	8.8(2.9)	10.3(2.4)	0.84
<b>Sequences</b>	14.8(7.6)	16.7(7.1)	0.80



**Figura 3.1** T scores de cada um dos pacientes com pelo menos 3 repetições de testes

Passando à modelação dos modelos lineares de efeitos mistos, foram utilizados os T scores dos primeiros 32 meses de estudo de todos os indivíduos da amostra que realizaram pelo menos 3 repetições. A escolha deste limite temporal deveu-se ao facto de o 3º quartil dos intervalos de tempo entre a primeira e a última repetições de testes se situar nos 32,75 meses. Para cada um dos GD 2 e 3 que permitem identificar o défice cognitivo do paciente, os indivíduos foram classificados de formas distintas:

- como tendo ou não défice cognitivo, incluindo os controlos (com controlos) ou não (sem controlos) – no caso dos sem controlos, e utilizando as 4 categorias diferenciadas a seguir, a comparação corresponde à comparação das categorias 0 e 2 versus as categorias 1 e 3;
- numa de 4 categorias, consoante tivessem ou não défice cognitivo dado os dois momentos de avaliação da performance cognitiva:
  - 0, se não apresentava défice cognitivo em nenhum dos momentos;
  - 1, se apresentava défice cognitivo apenas na avaliação inicial;
  - 2, se apresentava défice cognitivo apenas na avaliação final; ou
  - 3, se apresentava défice cognitivo nos dois momentos;

- numa de 3 categorias, consoante tivessem ou não défice dado os dois momentos de avaliação da performance cognitiva:
  - 0, se não apresentava défice cognitivo em nenhum dos momentos;
  - 1, se apresentava défice cognitivo num dos momentos de avaliação;
  - 2, se apresentava défice cognitivo nos dois momentos.

Deste modo, para cada uma das 3 classificações utilizadas na modelação dos modelos lineares e para cada uma das combinações de categorias incluídas nas mesmas classificações, construíram-se dois modelos - um apenas com a constante (*intercept*) e a variável tempo e um outro modelo que incluía, para além destes, o termo quadrático do variável tempo (sendo que em ambos modelos, os termos descritos eram simultaneamente inseridos como termos de efeitos fixos e aleatórios). Este termo quadrático, para todas as combinações de classificações, revelou-se sempre significativo, reiterando que os valores dos T scores das repetições de teste se ajustavam a um modelo quadrático.

Depois de obtido o modelo quadrático e de modo a avaliar a capacidade distinguir os indivíduos de acordo com as classificações apresentadas através dos efeitos aleatórios do *intercept*, procedeu-se ao cálculo da área sob a curva ROC (AUC). De modo análogo, para avaliar a capacidade da ferramenta BoT distinguir os indivíduos de acordo com as mesmas classificações, mas utilizando todos os efeitos aleatórios estimados, procedeu-se à estimação da AUC usando modelos de regressão logística (com a transformação *logit*). A Tabela 3.4 apresenta uma sistematização de todos os valores de AUC estimados de acordo com as estimativas de efeitos aleatórios utilizadas.

**Tabela 3.4** Área sob a curva ROC (AUC), consoante a classificação escolhida e o modelo utilizado. AUC [intervalo de confiança a 95%] são indicados. (n = número de indivíduos incluídos na construção do modelo)

	AUC [95% inf. - sup.]	
	Efeitos aleatórios da constante	Todos efeitos aleatórios
<b>GD 2</b>		
Com controlos (n=95)	0.73 [0.63 - 0.83]	0.74 [0.64 - 0.84]
Sem controlos (n=75) (0+2 vs 1+3)	0.73 [0.62 - 0.84]	0.76 [0.64 - 0.87]
<u>4 categorias</u>		
0 vs 1 (n=32)	0.74 [0.57 - 0.91]	0.75 [0.58 - 0.93]
0 vs 2 (n=25)	0.85 [0.68 - 1.00]	0.83 [0.66 - 1.00]
0 vs 3 (n=39)	0.83 [0.70 - 0.96]	0.83 [0.70 - 0.96]
1 vs 2 (n=15)	0.80 [0.55- 1.00]	0.68 [0.35 - 1.00]
1 vs 3 (n=29)	0.65 [0.44 - 0.86]	0.66 [0.44 - 0.87]
2 vs 3 (n=22)	0.50 [0.20 - 0.80]	0.67 [0.26 - 1.00]
<u>3 categorias</u>		
0 vs 1 (n=36)	0.77 [0.61 - 0.92]	0.79 [0.63 - 0.94]
1 vs 2 (n=33)	0.62 [0.42 - 0.82]	0.60 [0.40 - 0.80]
<b>GD 3</b>		
Com controlos (n=95)	0.92 [0.86 - 0.99]	0.92 [0.85 - 0.99]
Sem controlos (n=75) (0+2 vs 1+3)	0.92 [0.84 - 0.99]	0.91 [0.84 - 0.99]
<u>4 categorias</u>		
0 vs 1 (n=49)	0.93 [0.82 - 1.00]	0.92 [0.78 - 1.00]
0 vs 2 (n=47)	0.56 [0.23 - 0.88]	0.62 [0.26 - 0.99]
0 vs 3 (n=48)	0.98 [0.93 - 1.00]	0.98 [0.93 - 1.00]
1 vs 2 (n=6)	0.88 [0.53 - 1.00]	1.00
1 vs 3 (n=7)	0.67 [0.16 - 1.00]	1.00
2 vs 3 (n=5)	1.00	1.00
<u>3 categorias</u>		
0 vs 1 (n=51)	0.77 [0.53 - 1.00]	0.78 [0.55 - 1.00]
1 vs 2 (n=9)	0.83 [0.48 - 1.00]	1.00

Por observação da tabela anterior, e de acordo com Hosmer & Lemeshow (2000), conclui-se que grande parte das estimativas de AUC obtidas demonstram modelos com capacidades de discriminação aceitável (com de AUC entre 0,7 e 0,8), observando-se também modelos com capacidade discriminativa fraca ou mesmo não significativas (uma vez que o seu intervalo de confiança inclui o valor 0.5). Exemplos de modelos com capacidade de discriminação aceitável são os modelos que incluem todos os indivíduos (com controlos) e todos os pacientes (sem controlos) para o grau de defeito 2 (GD 2), incluindo, no entanto, os respetivos intervalos de confiança valores próximos de 0,6 que traduzem uma capacidade de discriminação mais fraca. Adicionalmente, verifica-se ainda que, para o grau de defeito 3 (GD 3), existem modelos com capacidade discriminativa excelente, como o modelo para todos os pacientes (AUC = 0.91 [0.84 - 0.99]), ou mesmo capacidade discriminativa extraordinária, como no caso dos modelos que comparam as classificações 0 vs 3 (AUC = 0.98 [0.93 - 1.00]) cuja estimativa pontual e intervalar são valores superiores a 0,9. Globalmente, modelos que compararam diferenças do nível 0 para as restantes apresentaram capacidade de discriminação aceitável, ao passo que os modelos das comparações entre pacientes com défice cognitivo nas duas avaliações (3) versus pacientes com défice cognitivo numa delas (1 ou 2) apresentaram dificuldades de discriminação. Os modelos que incluíam todos os efeitos aleatórios apresentavam quase sempre capacidade de discriminação igual ou superior aos que usavam apenas a constante.

Não é de estranhar que, globalmente, os modelos para o GD 2 possuam menor capacidade de distinguir os indivíduos do que os modelos contruídos para o GD 3, na medida em que o grau de defeito 3 é maior, implicando défices cognitivos mais acentuados e, por conseguinte, piores pontuações nas repetições dos testes da ferramenta BoT. Selecionando especificamente as estimativas de AUC dos modelos logísticos (com todos os efeitos aleatórios) para todos os pacientes (sem controlos), que avaliam a capacidade de distinguir uma pessoa com esclerose múltipla e défice cognitivo de um paciente com capacidade cognitiva preservada, temos valores de AUC de 0,76 e 0,91 para os GD 2 e 3, respetivamente. Estes valores não são muito distintos da estimativa apresentada por Ruano et al. (2019) – AUC = 0,91 – reiterando a capacidade da ferramenta BoT distinguir pacientes consoante tenham ou não a sua capacidade cognitiva preservada ou, por outras palavras, que tenham ou não défice cognitivo. Pelas mesmas estimativas poder-se-ia dizer que no caso do modelo para GD 3 a capacidade discriminativa é maior, no entanto, por observação dos intervalos de confiança a 95% estimados, constatamos que estes se intersejam, não sendo de crer que esta diferença



na capacidade de discriminação seja significativa, provavelmente devido à falta de poder amostral. Assim, para além de a ferramenta BoT mostrar ter poder de discriminação que permite distinguir pacientes com esclerose múltipla e défice cognitivo daqueles com capacidade cognitiva preservada, a mesma ferramenta apresenta também capacidade de distinguir pacientes que tenham uma recaída da performance cognitiva face aos que não tiveram (0 vs 2).

Finda esta parte da análise estatística, procedeu-se à construção de novos modelos lineares de efeitos mistos, construídos à semelhança do modelo linear de efeitos mistos utilizado para cálculo da AUC, mas que agora incluíam ainda nos efeitos fixos a interação entre o termo quadrático da variável tempo e a variável da classificação dos indivíduos bem como a interação entre esta última variável e a variável tempo de ordem 1. Depois de identificado o modelo com melhor ajuste aos dados por comparação do AIC, construiu-se um outro modelo linear de efeitos mistos semelhante a este, mas incluindo agora um ajuste à variável sexo, à variável idade à entrada no estudo e à variável anos de escolaridade categorizada de acordo com os intervalos 0 a 4, 5 a 9 e igual ou superior a 10. Procedeu-se depois à comparação destes dois últimos modelos, tendo se verificado, em todos os casos, que o modelo ajustado era significativamente diferente do modelo não ajustado e possuía melhor ajuste aos dados ao apresentar menor AIC. As Tabelas 3.5 e 3.6 apresentam, respetivamente, para o GD 2 e GD 3, um sumário das estimativas dos efeitos dos modelos não ajustados e ajustados.

**Tabela 3.5** Estimativas dos efeitos dos modelos lineares de efeitos mistos não ajustado e ajustado apenas para pacientes, de acordo com o GD 2. Estimativa [intervalo de confiança 95%] são indicados. **Abreviaturas:** PacienteCI, Paciente com défice cognitivo

	Estimativa [95% inf ; sup.]	
	Não ajustado	Ajustado
<b>Efeitos Fixos, coeficiente</b>		
Intercept	49,702 [46,951 ; 52,454]	47,126 [29,973 ; 64,279]
Tempo	0,778 [0,502 ; 1,054]	0,770 [0,499 ; 1,041]
PacienteCI	-7,017 [-10,949 ; -3,086]	-8,059 [-11,452 ; -4,667]
Tempo <sup>2</sup>	-0,024 [-0,035 ; -0,014]	-0,024 [-0,034 ; -0,014]
Interação Tempo:PacienteCI	-0,174 [-0,571 ; 0,222]	-0,160 [-0,549 ; 0,229]
Interação Tempo <sup>2</sup> :PacienteCI	0,008 [-0,007 ; 0,024]	0,008 [-0,007 ; 0,022]
<b>Efeitos Aleatórios, desvio-padrão</b>		
Intercept	8,283 [6,963 ; 9,853]	6,957 [5,796 ; 8,352]
Tempo	0,739 [0,552 ; 0,990]	0,722 [0,542 ; 0,960]
Tempo <sup>2</sup>	0,026 [0,017 ; 0,041]	0,025 [0,016 ; 0,040]
<b>Crítério de Informação Akaike (AIC)</b>	7410,50	7385,24

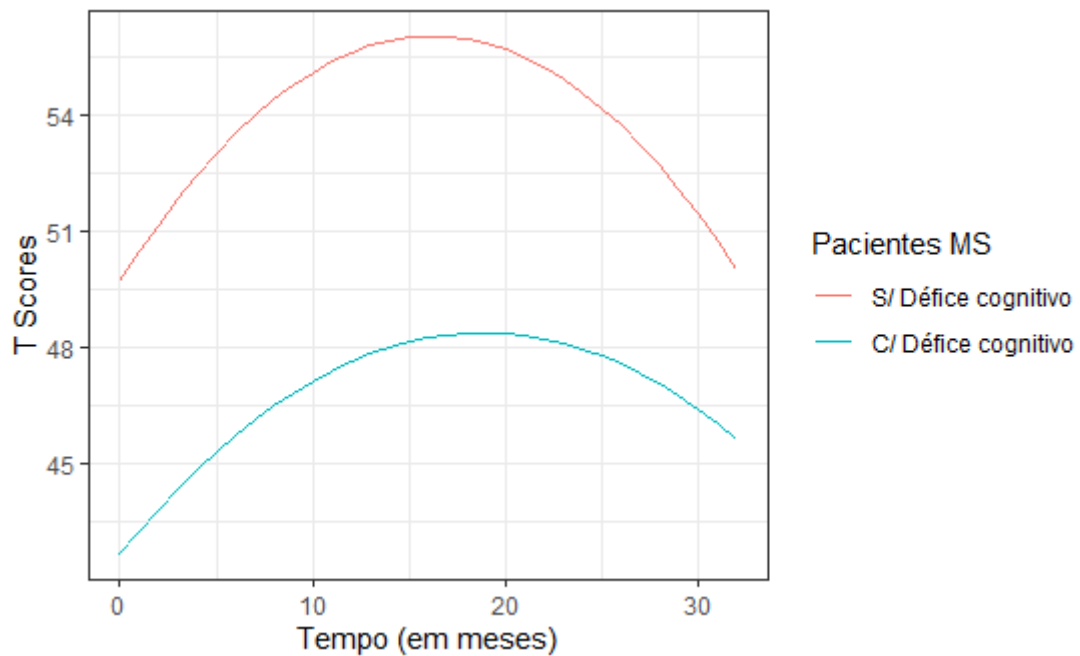
**Tabela 3.6** Estimativas dos efeitos dos modelos lineares de efeitos mistos não ajustado e ajustado apenas para pacientes, de acordo com o GD 3. Estimativa [intervalo de confiança 95%] são indicados. **Abreviaturas:** PacienteCI, Paciente com déficit cognitivo

	Estimativa [95% inf ; sup.]	
	Não ajustado	Ajustado
<b>Efeitos Fixos, coeficiente</b>		
Intercept	48,382 [46,494 ; 50,270]	46,633 [31,155 ; 62,110]
Tempo	0,677 [0,467 ; ,887]	0,656 [0,456 ; 0,856]
PacienteCI	-13,967 [-18,762 ; -9,172]	-13,981 [-18,065 ; -9,898]
Tempo <sup>2</sup>	-0,020 [-0,028 ; -0,012]	-0,018 [-0,025 ; -0,011]
Interação Tempo:PacienteCI	0,099 [-0,440 ; 0,637]	0,122 [-0,393 ; 0,638]
Interação Tempo <sup>2</sup> :PacienteCI	-0,003 [-0,024 ; 0,018]	-0,005 [-0,023 ; 0,014]
<b>Efeitos Aleatórios, desvio-padrão</b>		
Intercept	7,347 [6,160 ; 8,763]	6,106 [5,077 ; 7,345]
Tempo	0,718 [0,538 ; 0,958]	0,683 [0,520 ; 0,897]
Tempo <sup>2</sup>	0,025 [0,016 ; 0,039]	0,022 [0,014 ; 0,033]
<b>Critério de Informação Akaike (AIC)</b>	7395,51	7370,99

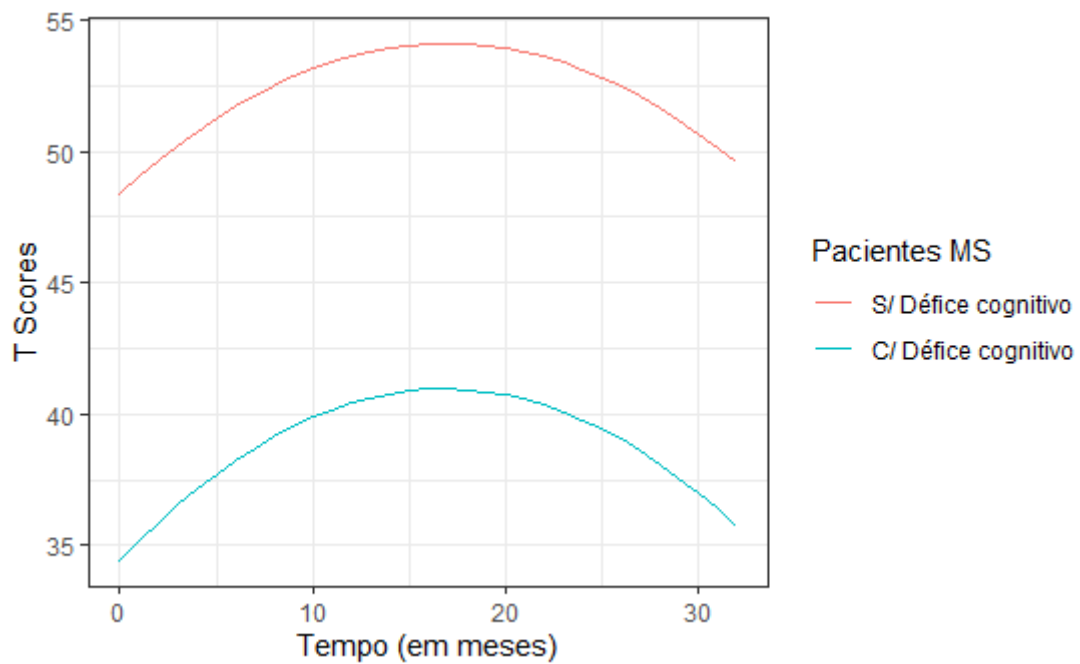
Para melhor compreensão, nas tabelas anteriores o fator “Paciente CI” representa um indivíduo com esclerose múltipla e déficit cognitivo e as duas últimas linhas dos efeitos fixos são relativas às interações entre as variáveis especificadas.

Constata-se então que para ambos os casos, tanto o modelo não ajustado como o ajustado apresentam valores idênticos, havendo uma diminuição global dos valores das estimativas significativas aquando do ajuste. Em ambas, e como já evidenciado anteriormente, as pontuações, nos primeiros 32 meses, mostram seguir um modelo quadrático no qual há um aumento progressivo das pontuações (estimativas do coeficiente da variável “Tempo” positivas), aumento este que vai diminuindo com o decorrer do tempo (estimativas do coeficiente da variável “Tempo<sup>2</sup>” negativas). Ainda que não significativas, as interações apresentavam estimativas simétricas consoante baseadas no grau de defeito 2 (GD 2) – onde traduziam um aumento progressivo de pontuações menor no caso dos pacientes com déficit cognitivo - ou no grau de defeito 3 (GD 3) - onde traduziam um aumento progressivo de pontuações maior no caso dos pacientes com déficit cognitivo. Uma outra diferença ressalta ao analisar-se a estimativa do coeficiente do fator “PacienteCI” que para GD 3 assume valores próximos do dobro de GD 2, o que traduzirá mais uma vez a sua maior capacidade de discriminação.

De acordo com estas estimativas apresentadas nas tabelas referidas, obtiveram-se as representações das curvas que definem os modelos calculados para GD 2 e GD 3, respetivamente (Figuras 3.2 e 3.3)



**Figura 3.2** Curvas do modelo linear de efeitos mistos não ajustado, para pacientes de acordo com GD 2.



**Figura 3.3** Curvas do modelo linear de efeitos mistos não ajustado, para pacientes de acordo com GD 3.

## Capítulo 4

### 4.1 Considerações finais

O tempo inicial investido na leitura de artigos, do protocolo de estudo e de outros documentos relativos ao tema do estudo e às metodologias utilizadas foi crucial para uma melhor compreensão dos processos que me encontrava a levar a cabo. Apesar de ser uma etapa desafiante e de contacto com informação que, em alguns casos, não tinha conhecimento, não foi uma etapa em que surgissem muitos obstáculos na medida em que a minha formação académica contempla não só conhecimentos estatísticos, mas também conhecimentos relativos a patologias e ao funcionamento de sistemas de órgãos, adquiridos ao longo da Licenciatura em Ciências Biomédicas.

No que respeita à organização e revisão dos dados das bases de dados, foi uma etapa que me permitiu concretizar a importância de se ter uma gestão de dados cuidada, não só para facilitar a sua interpretação (como, por exemplo, através da codificação de variáveis), mas também para detetar e eliminar possíveis erros que possam ser inseridos. Apesar de, em parte, monótona, esta etapa permitiu-me melhorar as minhas capacidades técnicas no que respeita a fazer uma utilização mais eficiente de ferramentas de gestão de dados, de modo a conseguir alcançar os objetivos pretendidos em menor tempo.

No que respeita à análise estatística, englobando desde os primeiros passos de análises preliminares até a construção dos modelos finais, foi uma experiência bastante enriquecedora uma vez que ganhei mais experiência na análise de dados bem como aprofundei e adquiri novos conhecimentos acerca de metodologias estatísticas e ferramentas do *software R* e *RStudio*, nomeadamente comandos que ainda não conhecia ou que ainda não tinha tido a oportunidade de utilizar. O facto de terem sido aplicados diferentes métodos, quase na totalidade abordados em diferentes unidades curriculares do Mestrado em Estatística Médica, permitiu-me consolidar esses conhecimentos e ainda esclarecer algumas dúvidas que foram surgindo, quer através da procura autónoma de informação, quer junto do meu supervisor de estágio, o Doutor Milton Severo. Na sua globalidade, estas tarefas de análise estatística permitiram-me melhorar a minha gestão de tempo, havendo ainda espaço para a tornar melhor, e permitiram-me experienciar que, ainda que o resultado final obtido esteja correto, a seleção da metodologia e ferramentas mais adequadas para o alcançar é crucial para melhorar a eficiência de qualquer projeto. No que respeita à escrita do artigo científico, surgiram algumas dificuldades na seleção de informação crucial a incluir no artigo e na apresentação das conclusões, tendo para

esse efeito consultado diferentes fontes que apresentam um conjunto de práticas de modo a colmatar este ponto não tão bem limado. Uma outra lição que retiro deste estágio é a necessidade de se fomentar a proatividade, já que tendo iniciativa se pode ir avançando nas tarefas ainda que com algumas limitações. Inicialmente, uma das dificuldades sentidas foi duvidar da própria motivação quando as decisões que tomava não eram as adequadas, por exemplo, no que respeita à seleção de uma metodologia ou comando do *software*, por outro lado era nessas ocasiões que encontrava ainda mais alento por ver espaço para melhorar, crescer e aprender. Quanto a este projeto/ estudo em que colaborei é, pessoalmente, algo que considero interessante e que suscitou a minha curiosidade, abordando uma doença conhecida mundialmente e apresentando uma nova abordagem inovadora que visa melhorar a qualidade de vida das pessoas com esclerose múltipla, equacionando-se a sua aplicação com outros fins distintos.

## 4.2 Conclusão

As pontuações totais do Brain on Track (BoT) apresentaram uma trajetória quadrática, tendo apresentado a ferramenta capacidade de discriminar pacientes com esclerose múltipla e défice cognitivo de pacientes com a performance cognitiva preservada, mostrando igualmente capacidade para identificar pacientes com recaídas da performance cognitiva através de dados longitudinais. A ferramenta BoT apresenta-se então como verdadeiramente promissora uma vez que mantém a sua capacidade de discriminação e vai de encontro às recomendações de alguns profissionais de ser necessário efetuar um diagnóstico de forma longitudinal. Uma das maiores limitações deste estudo foi o tamanho amostral que inviabilizou algumas das comparações que se pretenderam fazer, pelo que, nos próximos passos relacionados com a ferramenta BoT será benéfico o recrutamento de uma amostra de maiores dimensões de modo a contornar o abandono do estudo e a possibilitar uma maior representatividade das categorias definidas nas comparações, permitindo aumentar o poder estatístico da análise.

O ISPUP é um instituto que comecei a seguir de perto já há algum tempo, pelo que ter a oportunidade de ter trabalho lá, ainda que enquanto estagiário, foi deveras gratificante, para além de que foi uma oportunidade que me permitiu conhecer e contactar em primeira mão com o que é feito na área de investigação em saúde pública, em Portugal. Estando ainda o país a atravessar um tempo de pandemia com todas as restrições que isso implica, não me foi possível contactar com todas as pessoas com quem seria esperado se o estágio tivesse decorrido numa modalidade presencial, tendo tido a oportunidade de conhecer e trabalhar com o Doutor Milton Severo e o Doutor Luís Ruano que me propiciaram um ambiente de trabalho inclusivo e algo relaxado, que me deu a sensação de fazer parte da equipa.

Este estágio foi uma experiência única e bastante enriquecedora do ponto de vista profissional, mas que me permitiu também crescer enquanto ser humano. Para além de ter tido a possibilidade de aplicar conhecimentos adquiridos no Mestrado a um cenário não ficcional (como os que são algumas vezes criados para aplicar os conhecimentos em contexto de aula) tive ainda a oportunidade de melhorar as minhas capacidades de gestão de tempo, comunicação e de trabalho autónomo, continuando ainda a haver espaço para continuar a melhorar.

Como já referido, é incontornável não identificar o Mestrado como um dos fatores que me permitiu levar o estágio avante. A diversidade de informação e temas que são abordados no plano curricular permitiram-me reconhecer e sentir-me preparado para aplicar a quase

totalidade das metodologias estatísticas utilizadas. Para além deste ponto mais teórico, a própria metodologia de ensino que é adotada pelos professores permitiu-me estar confiante dos meus conhecimentos e saber que ter dúvidas ou não saber algo não deve constituir um entrave, mas antes ser um estímulo ao nosso trabalho e pesquisa autónoma.

Futuramente, espero poder continuar a desenvolver os meus conhecimentos e competências, querendo continuar o meu percurso académico com realização de um Doutoramento, e pretendo também continuar a contribuir para a melhoria contínua e reconhecimento do Mestrado em Estatística Médica da Universidade de Aveiro da melhor forma que conseguir.

## Bibliografia

1. ISPUP - Instituto de Saúde Pública da Universidade do Porto. Apresentação - Missão. <https://ispup.up.pt/about/school/mission/>. Accessed June 16, 2021.
2. ISPUP - Instituto de Saúde Pública da Universidade do Porto. Estatutos Do Instituto de Saúde Pública Da Universidade Do Porto. <http://asset.youoncdn.com/ab296ab30c207ac641882479782c6c34/40d51d9f1d89f31d4d466ba37d0b4239.pdf>. Accessed June 16, 2021.
3. Fundação para a Ciência e a Tecnologia. FCT — Unidades de I&D — Resultados da Avaliação de Unidades 2017. <https://www.fct.pt/apoios/unidades/avaliacoes/2017/resultados.phtml.pt>. Accessed June 21, 2021.
4. Tyagi H. How to Choose the Best Programming Language for your Data Science Project. <https://www.freecodecamp.org/news/how-to-choose-the-best-programming-language-for-your-data-science-project/>. Published July 1, 2020. Accessed June 21, 2021.
5. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020. <https://www.R-project.org/>.
6. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, 2019. <http://www.rstudio.com/>.
7. Alenius M, Koskinen S, Hallikainen I, Ngandu T, Lipsanen J, Sainio P, Tuulio-Henriksson A, Hänninen T. Cognitive Performance among Cognitively Healthy Adults Aged 30–100 Years. *Dement Geriatr Cogn Disord Extra* 2019; 9:11-23. doi: 10.1159/000495657
8. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med*. 2013 Nov;29(4):737–52.
9. Harvey PD. Domains of cognition and their assessment. *Dialogues Clin Neurosci*. 2019;21(3):227-237. doi:10.31887/DCNS.2019.21.3/pharvey
10. Weber D, Skirbekk V, Freund I, Herlitz A. The changing face of cognitive gender differences in Europe. *Proc Natl Acad Sci U S A*. 2014 Aug 12;111(32):11673-8. doi: 10.1073/pnas.1319538111.
11. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol*. 2012 Nov;11(11):1006-12. doi: 10.1016/S1474-4422(12)70191-6.



12. Ruano, L., Sousa, A., Severo, M. et al. Development of a self-administered web-based test for longitudinal cognitive assessment. *Sci Rep.* 2016; 6, 19114. <https://doi.org/10.1038/srep19114>
13. Ruano L, Severo M, Sousa A, et al. Tracking Cognitive Performance in the General Population and in Patients with Mild Cognitive Impairment with a Self-Applied Computerized Test (Brain on Track). *J Alzheimers Dis.* 2019;71(2):541-548. doi: 10.3233/JAD-190631.
14. McDowell, I. In *Measuring health: A guide to rating scales and questionnaires*, 3rd Edition 394–469 (Oxford University Press, 2006)
15. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. *Alzheimers Dement.* 2008;4(6):428-437. doi:10.1016/j.jalz.2008.07.003
16. Schlegel, R. E. & Gilliland, K. Development and quality assurance of computer-based assessment batteries. *Archives of Clinical Neuropsychology.* 2007; 22 Suppl 1, S49–61, doi: 10.1016/j.acn.2006.10.005.
17. Milne A, Culverwell A, Guss R, Tuppen J, Whelton R. Screening for dementia in primary care: a review of the use, efficacy and quality of measures. *Int Psychogeriatr.* 2008 Oct;20(5):911-26. doi: 10.1017/S1041610208007394.
18. Twisk, Jos W. R. *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide*. 1st ed. Cambridge: Cambridge UP, 2013.
19. Pinheiro, José C, and Douglas M. Bates. *Mixed-effects Models in S and S-Plus*. New York: Springer, 2000.
20. Hosmer, D. W., & Lemeshow, S. *Applied Logistic Regression*. Wiley Interscience, 2000.



## **Anexos**

## Anexo A – Dicionário de variáveis da base de dados da bateria de exames neuropsicológicos inicial

ID Variável	Variável BD	Variável (nome)	Descrição adicional	Categorias
1	id	Indivíduo		
2	Idade	Idade	em anos	NA S/ Informação
3	Escolaridade	Escolaridade	em anos	NA S/ Informação
4	Data_de_avaliacao	Data de avaliação	dd/mm/aaaa	NA S/ Informação
5	MoCA	MoCA [screening]	Montreal Cognitive Assessment test (MoCA)	NA S/ Informação
6	GD_MoCA	Grau de defeito MoCA [screening]		1 Dentro dos parâmetros normativos 2 Abaixo dos parâmetros normativos NA S/ Informação
7	Trilhas_A	Trilhas A [Atenção/velocidade de processamento]	Trail Making Test (TMT), em segundos	NA S/ Informação
8	GD_Trilhas_A	Grau de defeito Trilhas A [Atenção/velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
9	Trilhas_B	Trilhas B [Atenção/velocidade de processamento]	Trail Making Test (TMT), em segundos	NA S/ Informação
10	GD_Trilhas_B	Grau de defeito Trilhas B [Atenção/velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
11	SLN	SLN [Memória]	Letter-Number Sequencing (LNS)	NA S/ Informação
12	GD_SLN	Grau de defeito SLN [Memória]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave

				NA S/ Informação
13	CVLT_lista_A_total_imediata	CVLT_lista A_total(imediata) [Memória]	Lista A imediata - California Verbal Learning Test (CVLT)	NA S/ Informação
14	CVLT__lista_B_imediata	CVLT_ lista B(imediata) [Memória]	Lista B imediata - California Verbal Learning Test (CVLT)	NA S/ Informação
15	CVLT_diferida	CVLT_diferida [Memória]	Diferida - California Verbal Learning Test (CVLT)	NA S/ Informação
16	CVLT_dif_categ_semantica	CVLT_dif_categ.semântica [Memória]	Categoria semântica - California Verbal Learning Test (CVLT)	NA S/ Informação
17	BVMT_R_imediata	BVMT-R imediata [Memória]	Imediata - Brief Visuospatial Memory Test (BVMT)	NA S/ Informação
18	BVMT_R_diferida	BVMT-R diferida [Memória]	Diferida - Brief Visuospatial Memory Test (BVMT)	NA S/ Informação
19	Matrizes	Matrizes [Funções executivas]	Matrix Reasoning Test (MRT)	NA S/ Informação
20	GD_Matrizes	Grau de defeito Matrizes [Funções executivas]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
21	Fluencia_Verbal_semantica	Fluência Verbal semântica [Funções executivas]		NA S/ Informação
22	GD_Fluencia_Verbal_semantica	Grau de defeito Fluência Verbal semântica [Funções executivas]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
23	Fluencia_verbal_fonologica	Fluência verbal_fonológica [Funções executivas]		NA S/ Informação
24	GD_Fluencia_verbal_fonologica	Grau de defeito Fluência verbal_fonológica [Funções executivas]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
25	Relogio	Relógio [Funções executivas]	Clock Drawing Test	NA S/ Informação
26	GD_Relogio	Grau de defeito Relógio [Funções executivas]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação

27	Stroop_nomeacao	Stroop_nomeação [Funções executivas]	Stroop test	NA S/ Informação
28	GD_Stroop_nomeacao	Grau de defeito Stroop_nomeação [Funções executivas]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
29	Stroop_leitura	Stroop_leitura [Velocidade de processamento]	Stroop test	NA S/ Informação
30	GD_Stroop_leitura	Grau de defeito Stroop_leitura [Velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
31	Pesquisa_de_simbolos	Pesquisa de símbolos [Velocidade de processamento]	Symbol Search Test (SST)	NA S/ Informação
32	GD_Pesquisa_de_simbolos	Grau de defeito Pesquisa de símbolos [Velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
33	SDMT	SDMT [Velocidade de processamento]	Symbol Digit Modalities Test (SDMT)	NA S/ Informação
34	BDI	BDI [Humor]	Beck Depression Inventory (BDI)	NA S/ Informação
35	MFIS_Total	MFIS_Total [Fadiga]	Modified Fatigue Impact Scale (MFIS)	NA S/ Informação
36	Questao_1	Questão 1 [Questionário_Tecnologias]	Telefone fixo	NA S/ Informação
37	Questao_2	Questão 2 [Questionário_Tecnologias]	Telemóvel / PDA	NA S/ Informação
38	Questao_2_1	Questão 2.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
39	Questao_3	Questão 3 [Questionário_Tecnologias]	Computador	NA S/ Informação
40	Questao_3_1	Questão 3.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
41	Questao_4	Questão 4 [Questionário_Tecnologias]	Acesso_internet	NA S/ Informação
42	Questao_4_1	Questão 4.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
43	Questao_5	Questão 5 [Questionário_Tecnologias]	Uilizador_internet_próx.	NA S/ Informação
44	Questao_6	Questão 6	Utilização difícil de	NA S/ Informação

		[Questionário_Tecnologias]	tecnologias	
45	Questao_7	Questão 7 [Questionário_Tecnologias]	Aprender tecnologias	NA S/ Informação
46	Questao_8	Questão 8 [Questionário_Tecnologias]	Grau de confiança com tecnologias	NA S/ Informação
47	Questao_9	Questão 9 [Questionário_Tecnologias]	Pagamento de tecnologias	NA S/ Informação
48	Tempo_de_evolucao	Tempo de evolução [Avaliação clinica inicial]		
49	EDSS_ini	EDSS [Avaliação clinica inicial]	Expanded Disability Status Scale (EDSS) à entrada do estudo	
			à entrada do estudo	1 Acetato de glatiramer 2 Fingolimod 3 Interferon beta 1a 4 Interferon beta 1b 5 Natalizumab 6 Teriflunamida 7 Dimetil-furamato 8 Interferon pegilado beta 1a
50	Medicacao_ini	Medicação [Avaliação clinica inicial]		
51	Surtos_1_ano_ini	Surtos 1 ano [Avaliação clinica inicial]	à entrada do estudo	
52	Surtos_5_anos_ini	Surtos 5 anos [Avaliação clinica inicial]	à entrada do estudo	
53	EDSS_fu	EDSS [Avaliação Clinica 12 meses]	Expanded Disability Status Scale (EDSS)	NA S/ Informação
			Após decorridos 12 meses do estudo	0 Sem medicação 1 Acetato de glatiramer 2 Fingolimod 3 Interferon beta 1a 4 Interferon beta 1b 5 Natalizumab 6 Teriflunamida 7 Dimetil-furamato 8 Interferon pegilado beta 1a 9 Rituximab 10 Ocrelizumab 11 Cladribina
54	Medicacao_fu	Medicação [Avaliação Clinica 12 meses]		
55	Surtos_1_ano_fu	Surtos 1 ano [Avaliação Clinica 12 meses]	Após decorridos 12 meses do estudo	NA S/ Informação

56	Score2	Nº de Graus de Defeito iguais ou superiores a 2
57	Score3	Nº de Graus de Defeito iguais ou superiores a 3
58	Score4	Nº de Graus de Defeito iguais ou superiores a 4



## Anexo B – Dicionário de variáveis da base de dados da bateria de exames neuropsicológicos final

ID Variável	Variável BD	Variável (nome)	Descrição adicional	Categorias
1	id	Indivíduo		
2	Idade	Idade	em anos	NA S/ Informação
3	Escolaridade	Escolaridade	em anos	NA S/ Informação
4	Data_de_avaliacao	Data de avaliação	dd/mm/aaaa	NA S/ Informação
5	MoCA	MoCA [screening]	Montreal Cognitive Assessment test (MoCA)	NA S/ Informação
6	GD_MoCA	Grau de defeito MoCA [screening]		1 Dentro dos parâmetros normativos 2 Abaixo dos parâmetros normativos NA S/ Informação
7	Trilhas_A	Trilhas A [Atenção/velocidade de processamento]	Trail Making Test (TMT), em segundos	NA S/ Informação
8	GD_Trilhas_A	Grau de defeito Trilhas A [Atenção/velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
9	Trilhas_B	Trilhas B [Atenção/velocidade de processamento]	Trail Making Test (TMT), em segundos	NA S/ Informação
10	GD_Trilhas_B	Grau de defeito Trilhas B [Atenção/velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito

				grave NA S/ Informação
11	SLN	SLN [Memória]	Letter-Number Sequencing (LNS)	NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
12	GD_SLN	Grau de defeito SLN [Memória]		NA S/ Informação
13	"CVLT_lista_A_total_imediata	CVLT_lista A_total(imediata) [Memória]	Lista A imediata - California Verbal Learning Test (CVLT)	NA S/ Informação
14	"CVLT__lista_B_imediata	CVLT_ lista B(imediata) [Memória]	Lista B imediata - California Verbal Learning Test (CVLT)	NA S/ Informação
15	CVLT_diferida	CVLT_diferida [Memória]	Diferida - California Verbal Learning Test (CVLT)	NA S/ Informação
16	CVLT_dif_categ_semantica	CVLT_dif_categ.semântica [Memória]	Categoria semântica - California Verbal Learning Test (CVLT)	NA S/ Informação
17	BVMT_R_imediata	BVMT-R imediata [Memória]	Imediata - Brief Visuospatial Memory Test (BVMT)	NA S/ Informação
18	BVMT_R_diferida	BVMT-R diferida [Memória]	Diferida - Brief Visuospatial Memory Test (BVMT)	NA S/ Informação
19	Matrizes	Matrizes [Funções executivas]	Matrix Reasoning Test (MRT)	NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
20	GD_Matrizes	Grau de defeito Matrizes [Funções executivas]		NA S/ Informação
21	Fluencia_Verbal_semantica	Fluência Verbal_semântica [Funções executivas]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado
22	GD_Fluencia_Verbal_semantica	Grau de defeito Fluência Verbal_semântica [Funções executivas]		

				4 Defeito grave NA S/ Informação
23	Fluencia_verbal_fonologica	Fluência verbal_fonológica [Funções executivas]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
24	GD_Fluencia_verbal_fonologica	Grau de defeito Fluência verbal_fonológica [Funções executivas]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
25	Relogio	Relógio [Funções executivas]	Clock Drawing Test	NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
26	GD_Relogio	Grau de defeito Relógio [Funções executivas]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
27	Stroop_nomeacao	Stroop_nomeação [Funções executivas]	Stroop test	NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
28	GD_Stroop_nomeacao	Grau de defeito Stroop_nomeação [Funções executivas]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
29	Stroop_leitura	Stroop_leitura [Velocidade de processamento]	Stroop test	NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
30	GD_Stroop_leitura	Grau de defeito Stroop_leitura [Velocidade de processamento]		NA S/ Informação
				1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
31	Pesquisa_de_simbolos	Pesquisa de símbolos	Symbol Search Test	NA S/ Informação

		[Velocidade de processamento]	(SST)	Informação
32	GD_Pesquisa_de_simbolos	Grau de defeito Pesquisa de símbolos [Velocidade de processamento]		1 Sem defeito 2 Defeito ligeiro 3 Defeito moderado 4 Defeito grave NA S/ Informação
33	SDMT	SDMT [Velocidade de processamento]	Symbol Digit Modalities Test (SDMT)	NA S/ Informação
34	BDI	BDI [Humor]	Beck Depression Inventory (BDI)	NA S/ Informação
35	MFIS_Total	MFIS_Total [Fadiga]	Modified Fatigue Impact Scale (MFIS)	NA S/ Informação
36	Questao_1	Questão 1 [Questionário_Tecnologias]	Telefone fixo	NA S/ Informação
37	Questao_2	Questão 2 [Questionário_Tecnologias]	Telemóvel / PDA	NA S/ Informação
38	Questao_2_1	Questão 2.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
39	Questao_3	Questão 3 [Questionário_Tecnologias]	Computador	NA S/ Informação
40	Questao_3_1	Questão 3.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
41	Questao_4	Questão 4 [Questionário_Tecnologias]	Acesso_internet	NA S/ Informação
42	Questao_4_1	Questão 4.1 [Questionário_Tecnologias]	Frequência de utilização	NA S/ Informação
43	Questao_5	Questão 5 [Questionário_Tecnologias]	Utilizador_internet_próx.	NA S/ Informação
44	Questao_6	Questão 6 [Questionário_Tecnologias]	Utilização difícil de tecnologias	NA S/ Informação
45	Questao_7	Questão 7 [Questionário_Tecnologias]	Aprender tecnologias	NA S/ Informação
46	Questao_8	Questão 8 [Questionário_Tecnologias]	Grau de confiança com tecnologias	NA S/ Informação
47	Questao_9	Questão 9 [Questionário_Tecnologias]	Pagamento de tecnologias	NA S/ Informação
48	Score2	Nº de Graus de Defeito iguais ou superiores a 2		
49	Score3	Nº de Graus de Defeito iguais ou superiores a 3		
50	Score4	Nº de Graus de Defeito iguais ou superiores a 4		

51	Score2_var	Varição do Nº de Graus de Defeito iguais ou superiores a 2	Nº de Grau de Defeitos iguais ou superiores a 2 da avaliação final menos os da avaliação inicial
52	Score3_var	Varição do Nº de Graus de Defeito iguais ou superiores a 3	Nº de Grau de Defeitos iguais ou superiores a 3 da avaliação final menos os da avaliação inicial
53	Score4_var	Varição do Nº de Graus de Defeito iguais ou superiores a 4	Nº de Grau de Defeitos iguais ou superiores a 4 da avaliação final menos os da avaliação inicial

## Anexo C – Dicionário de variáveis da base de dados dos resultados do Brain on Track (BoT)

Excerto do dicionário de variáveis, sendo que este se replica durante as 48 repetições do Teste 2 e as 66 repetições do Teste 3.

ID Variável	Variável BD	Variável (nome)	Categorias
1	id	Indivíduo	
2	Doente	Grupo de tratamento	0 Controlo 1 Doente da Feira 2 Doente de Braga
3	Utilizadores	Utilizadores	
4	Data_de_nascimento	Data de Nascimento (aaaa/mm/dd)	
5	Escolaridade	Escolaridade (anos)	
6	Genero	Género	1 Feminino 2 Masculino
7	Teste_1_0_Total_de_testes	Total de repetições do Teste 1	
8	Teste_2_0_Total_de_testes	Total de repetições do Teste 2	
9	Teste_3_0_Total_de_testes	Total de repetições do Teste 3	
10	Teste_1_0_Testes_completos	Repetições do Teste 1 completas	
11	Teste_2_0_Testes_completos	Repetições do Teste 2 completas	
12	Teste_3_0_Testes_completos	Repetições do Teste 3 completas	
13	Teste_1_0_Testes_incompletos	Repetições do Teste 1 incompletas	
14	Teste_2_0_Testes_incompletos	Repetições do Teste 2 incompletas	
15	Teste_3_0_Testes_incompletos	Repetições do Teste 3 incompletas	
16	Teste_1_0_Testes_nao_realizados	Repetições do Teste 1 não realizadas	
17	Teste_2_0_Testes_nao_realizados	Repetições do Teste 2 não realizadas	
18	Teste_3_0_Testes_nao_realizados	Repetições do Teste 3 não realizadas	
19	Teste_2_0_1_Data	Data da repetição 1 do Teste 2 (aaaa/mm/dd)	
20	Teste_2_0_1_Igual_ou_diferente_C	Exercícios corretos do Subteste Attention task III da repetição 1	

	ertos	do Teste 2
21	Teste_2_0_1_Igual_ou_diferente_Er rados	Exercícios errados do Subteste Attention task III da repetição 1 do Teste 2
22	Teste_2_0_1_Cubos_irrequietos_C ertos	Exercícios corretos do Subteste Visual memory task II da repetição 1 do Teste 2
23	Teste_2_0_1_Cubos_irrequietos_Er rados	Exercícios errados do Subteste Visual memory task II da repetição 1 do Teste 2
24	Teste_2_0_1_Longa_memoria_Cert os	Exercícios corretos do Subteste Delayed verbal memory task da repetição 1 do Teste 2
25	Teste_2_0_1_Longa_memoria_Erra dos	Exercícios errados do Subteste Delayed verbal memory task da repetição 1 do Teste 2
26	Teste_2_0_1_Mente_rapida_Certos	Exercícios corretos do Subteste Calculus task da repetição 1 do Teste 2
27	Teste_2_0_1_Mente_rapida_Errado s	Exercícios errados do Subteste Calculus task da repetição 1 do Teste 2
28	Teste_2_0_1_Combine_a_cor_Cert os	Exercícios corretos do Subteste Colour interference task da repetição 1 do Teste 2
29	Teste_2_0_1_Combine_a_cor_Erra dos	Exercícios errados do Subteste Colour interference task da repetição 1 do Teste 2
30	Teste_2_0_1_Rapida_memoria_Cer tos	Exercícios corretos do Subteste Verbal memory I da repetição 1 do Teste 2
31	Teste_2_0_1_Rapida_memoria_Err ados	Exercícios errados do Subteste Verbal memory II da repetição 1 do Teste 2
32	Teste_2_0_1_Contrario_Certos	Exercícios corretos do Subteste Opposite task da repetição 1 do Teste 2
33	Teste_2_0_1_Contrario_Errados	Exercícios errados do Subteste Opposite task da repetição 1 do Teste 2
34	Teste_2_0_1_Correspondencias_C ertos	Exercícios corretos do Subteste Attention task III da repetição 1 do Teste 2
35	Teste_2_0_1_Correspondencias_Er rados	Exercícios errados do Subteste Attention task III da repetição 1 do Teste 2
36	Teste_2_0_1_Arrume_as_palavras_ Certos	Exercícios corretos do Subteste Word categories da repetição 1 do Teste 2
37	Teste_2_0_1_Arrume_as_palavras_	Exercícios errados do Subteste

	Errados	Word categories da repetição 1 do Teste 2
38	Teste_2_0_1_Sequencias_Certos	Exercícios corretos do Subteste Sequences da repetição 1 do Teste 2
39	Teste_2_0_1_Sequencias_Errados	Exercícios errados do Subteste Sequences da repetição 1 do Teste 2
40	Teste_2_0_1_Puzzles_Certos	Exercícios corretos do Subteste Puzzles da repetição 1 do Teste 2
41	Teste_2_0_1_Puzzles_Errados	Exercícios errados do Subteste Puzzles da repetição 1 do Teste 2
42	Teste_2_0_2_Data	Data da repetição 2 do Teste 2 (aaaa/mm/dd)
...	...	...