Universidade de Évora - Escola de Ciências e Tecnologia

Mestrado em Engenharia Informática

Dissertação

Plataforma Web para Análise Normativa de Testes Neuropsicológicos utilizando um Modelo de Dados EAV

Diogo Filipe Cristo Mestre

Orientador(es) | Pedro Salgueiro

Teresa Gonçalves

Évora 2026

Universidade de Évora - Escola de Ciências e Tecnologia

Mestrado em Engenharia Informática

Dissertação

Plataforma Web para Análise Normativa de Testes Neuropsicológicos utilizando um Modelo de Dados EAV

Diogo Filipe Cristo Mestre

Orientador(es) | Pedro Salgueiro

Teresa Gonçalves

Évora 2026

A dissertação foi objeto de apreciação e discussão pública pelo seguinte júri nomeado pelo Diretor da Escola de Ciências e Tecnologia:

Presidente	|	Luís Rato (Universidade de Évora)
Vogais	|	José Saias (Universidade de Évora) (Arguente) Pedro Salgueiro (Universidade de Évora) (Orientador)

Évora 2026

Dedicado a...

Agradecimentos

Gostaria de agradecer a todos os professores que me acompanharam ao longo do mestrado, pelo conhecimento transmitido e pelo contributo fundamental para o meu desenvolvimento académico e pessoal. Em particular, agradeço aos meus orientadores pela orientação rigorosa e constante disponibilidade. O seu apoio, sugestões e motivação foram essenciais e cruciais para a conclusão desta dissertação. Gostaria igualmente de expressar o meu reconhecimento a todos os colegas e amigos que me acompanharam. A vossa ajuda e apoio foram cruciais para tornar este percurso mais fácil e enriquecedor. Finalmente, a minha gratidão estende-se à minha família. O seu apoio incondicional, a compreensão demonstrada e a confiança depositada foram a base essencial para a conclusão deste trabalho.

Declaração sobre o uso de IA na elaboração da dissertação

Declaro que recorri à utilização de ferramentas de Inteligência Artificial Generativa (ChatGPT (OpenAI), Gemini, Claude) no desenvolvimento desta dissertação, nomeadamente como suporte à redação, reformulação de texto, descrição da finalidade de forma sucinta e esclarecimento de conceitos técnicos. A utilização destas ferramentas foi sempre de caráter auxiliar, não substituindo o processo de análise, desenvolvimento e tomada de decisão científica.

Sempre que aplicável, a sua utilização encontra-se descrita na metodologia adotada e, sempre que necessário, explicitada no próprio texto. Todo o conteúdo produzido ou assistido por IA foi revisto, validado e editado por mim, pelo que assumo plena responsabilidade pelo rigor, originalidade, transparência e integridade científica do trabalho que aqui se materializa. Considero-me, neste sentido, pleno Autor do texto submetido.

Conteúdo

Lista de Figuras

Lista de Tabelas

Lista de Acrónimos

ADCC Alzheimers Disease Coordinating Center

ADEPT ANAM Data Extraction and Presentation Tool

ADRCs Alzheimers Disease Research Centers

ANAM Automated Neuropsychological Assessment Metrics

ANDI Advanced Neuropsychological Diagnostics Infrastructure

ANOVA Analysis of Variance

APR ANAM Performance Report

BFC Baringhaus and Franz's Cramér statistic

BNT Boston Naming Test

CNNS Calibrated Neuropsychological Normative System

COWA Controlled Oral Word Association Test

CSV Comma-Separated Values

EAV Entity-Attribute-Value

FWER Family-Wise Error Rate

GAMLSS Generalized Additive Models for Location, Scale, and Shape

I-ANDI Indonesian Advanced Neuropsychological Diagnostics Infrastructure

I-BNT Indonesian Boston Naming Test

JSON JavaScript Object Notation

NoSQL Not Only SQL

ORM Object-Relational Mapper

ROCF Rey-Osterrieth Complex Figure Test

SCWT Stroop Color-Word Interference Test

TCE Traumatismos Cranioencefálicos

TMT Trail Making Test

UDS Uniform Data Set

Sumário

A análise e interpretação precisa dos resultados dos testes neuropsicológicos é fundamental para a prática clínica e investigação. No entanto, a complexidade e variabilidade destes testes muitas vezes dificultam a recolha de dados consistentes e a realização de comparações normativas significativas, especialmente para populações culturalmente diversas. Esta dissertação apresenta uma base de dados especialmente desenvolvida para lidar com a escalabilidade e variabilidade de diferentes testes neuropsicológicos e uma aplicação web para recolher, armazenar e analisar resultados de testes neuropsicológicos de diversos estudos. A aplicação permite ao utilizador fazer o *upload* de dados de um estudo, realizar pesquisas e fazer comparações normativas tendo em consideração todas as amostras armazenadas. Para comparações normativas, a aplicação classifica os resultados dos testes de um paciente tendo como norma os valores existentes na base de dados, permitindo aos utilizadores avaliar os dados do sujeito em relação aos valores normativos e determinar se os primeiros se encontram dentro dos intervalos típicos. Um aspeto central do desenvolvimento é o sistema da base de dados, que requer uma estrutura dinâmica e escalável para acomodar uma ampla gama de testes e parâmetros. Esta dissertação também descreve as principais considerações no design da base de dados e aborda os desafios da estruturação de dados neuropsicológicos para a flexibilidade e precisão analítica. A aplicação web visa simplificar a gestão dos dados, oferecendo ferramentas robustas para a comparação normativa e capacidades preditivas.

Palavras chave: Testes neuropsicológicos, Comparação Normativa de dados, Aplicação Web, Base de Dados, Modelo EAV

Abstract

Web Platform for Normative Analysis of Neuropsychological Tests using an EAV Data Model

Accurate analysis and interpretation of neuropsychological test results are fundamental for clinical practice and research. However, the complexity and variability of these tests often make it difficult to collect consistent data and make meaningful normative comparisons, especially for culturally diverse populations. This dissertation presents a specially developed database to handle the scalability and variability among different neuropsychological tests and a web application to collect, store, and analyze neuropsychological test results from various studies. The application allows users to upload data from a study, perform searches, and make normative comparisons taking into consideration all stored samples. For normative comparisons, the application classifies a patient's test results using the values in the database as the norm, allowing users to evaluate the subject's data against normative values and determine whether the former fall within the typical ranges. A central aspect of the development is the database system, which requires a dynamic and scalable structure to accommodate a wide range of tests and parameters. This dissertation also describes key considerations in database design and addresses the challenges of structuring neuropsychological data for flexibility and analytical accuracy. The web application aims to simplify data management by offering robust tools for normative comparison and predictive capabilities.

Keywords: Neuropsychological Tests, Normative data comparison, Web Application, Database, Model EAV

Capítulo 1

Introdução

Os testes neuropsicológicos referem-se a uma série de testes que os profissionais de saúde utilizam para obter informações sobre o funcionamento do cérebro humano e são uma ferramenta fundamental para ajudar a avaliar as funções cognitivas e compreender o impacto das condições neurológicas, médicas, psicológicas e sociais no cérebro.

Existem vários tipos de testes neuropsicológicos, dirigidos, por exemplo, à memória, cognição, comunicação verbal ou habilidades motoras, que normalmente envolvem escrever, desenhar, resolver quebra-cabeças ou responder a perguntas ou questões sobre imagens apresentadas num computador [Cleveland Clinic, 2023; Thayer and Robokos, 2025].

Ao realizar vários testes, os profissionais de saúde podem obter um perfil detalhado das capacidades cognitivas de um indivíduo. Estes testes podem ajudar os médicos a determinar um diagnóstico (uma vez que os seus resultados ajudam a compreender a causa de alguns sintomas), a identificar pontos fortes e fracos nos processos cognitivos, a fornecer planos para tratamento futuro (esquema de reabilitação, por exemplo) e compreender o risco de alterações após um acidente/cirurgia cerebral [Cleveland Clinic, 2023; Thayer and Robokos, 2025].

Portanto, a análise e interpretação precisas dos resultados dos testes neuropsicológicos desempenham um papel crucial tanto na avaliação clínica quanto na investigação. No entanto, a natureza complexa dos testes neuropsicológicos apresenta desafios significativos para a recolha e análise consistente de dados. Por exemplo, o *Trail Making Test* (TMT), teste que avalia a atenção e a função executiva, inclui duas condições: *Trail A* e *Trail B*. *Trail A* avalia a atenção e a função executiva; os seus parâmetros incluem o tempo de conclusão, erros de sequenciamento e erros de perda de conjunto. *Trail B* avalia a atenção, a flexibilidade cognitiva e a função executiva, também medidas pelo tempo de conclusão, erros de sequenciamento e erros de perda de conjunto [Ashendorf et al., 2008; Tsiakiri et al., 2024; Lähde et al., 2024].

A comparação normativa é utilizada para determinar o desempenho de um sujeito num determinado teste, com base no desempenho de um grupo de referência. Estas comparações permitem aos médicos verificar se os resultados do indivíduo desviam-se significativamente do esperado. No entanto, enfrentam alguns desafios devido a fatores como diferenças no estatuto socioeconómico, género, idade, nível de educação e contexto cultural, que podem afetar significativamente o desempenho, dificultando o estabelecimento de normas culturalmente específicas [Tripathi et al., 2014]. Os dados normativos frequentemente agrupam os sujeitos em amplas faixas etárias, o que pode levar à representação incorreta das capacidades cognitivas, uma vez que indivíduos da mesma faixa etária podem ter desempenhos muito diferentes [Arampatzi et al., 2024]. Além disso, podem não existir dados normativos para algumas populações específicas, o que torna a interpretação dos resultados dos testes num contexto significativo mais complicada [Lavrencic et al., 2019].

Para alguns testes, especialmente aqueles que verificam a memória, não podemos criar versões diferentes que sejam igualmente desafiadoras. O que significa que não podemos testar indivíduos com frequência sem que os mesmos saibam as respostas. Esta limitação restringe a frequência com que estes testes podem ser administrados, o que contribui para a escassez de dados disponíveis para análises. Outras causas que podem contribuir para a falta de dados são a longa duração dos testes e o grande número de testes que falham por vários motivos durante a sua administração, o que compromete a clareza da interpretação e restringe a validade ou a confiabilidade dos resultados, e, por sua vez, dificulta significativamente a comparação normativa. Os diferentes idiomas e culturas também contribuem para a falta de dados, uma vez que alguns testes precisam ser traduzidos e adaptados antes de serem aplicados [Howieson, 2019].

Para resolver estas limitações, incluindo o tamanho e a disponibilidade limitada dos conjuntos de dados normativos existentes, esta dissertação apresenta o desenho e o desenvolvimento de uma aplicação web para a recolha, armazenamento e análise dos resultados de testes neuropsicológicos em vários estudos. Esta aplicação permite que os investigadores importem os seus dados de estudos neuropsicológicos, realizem pesquisas direcionadas e comparações normativas utilizando todos os dados guardados no sistema. Esta ferramenta foi projetada para ser capaz de suportar um grupo crescente e diversificado de testes, indivíduos e parâmetros de teste.

A arquitetura da base de dados é um componente crítico desta solução. Foi cuidadosamente projetada para armazenar dados provenientes de diversos testes, com diferentes condições, parâmetros e intervalos de resultados, garantindo a integridade dos dados e facilitando o processo analítico.

1.1 Motivação

A análise normativa de testes neuropsicológicos continua a enfrentar múltiplos obstáculos, tanto do ponto de vista clínico como científico. Estes desafios incluem a

heterogeneidade dos indivíduos, ou seja, a manifestação dos sintomas cognitivos em indivíduos com doenças raras ou condições específicas varia bastante, o que torna a aplicação de testes padronizados mais complicada. Outro desafio é a escassez de normas e ferramentas. A maioria dos dados e testes disponíveis foram desenvolvidos para doenças comuns, o que torna a avaliação de indivíduos com condições raras muito mais complexa. Como já foi mencionado, os conjuntos de dados normativos disponíveis atualmente possuem limitações, em termos de dimensão, diversidade e disponibilidade, o que dificulta a obtenção de normas robustas e adaptadas a diferentes populações [NeuronUP, 2025]. Estes fatores dificultam a comparação normativa precisa dos resultados de um indivíduo com a sua população de referência.

De um ponto de vista técnico, a gestão dos dados de testes neuropsicológicos apresenta desafios adicionais. Devido às diferentes condições, parâmetros e níveis de desempenho de cada teste, é fundamental desenhar uma estrutura flexível capaz de armazenar informações diversas, sem comprometer a sua consistência ou integridade. Modelos tradicionais de bases de dados são, por norma, baseados em esquemas fixos, o que os torna menos ideais para lidar com dados muito dinâmicos e variáveis. Esta característica dificulta a adição de novos testes ou a alteração de variantes sem uma reestruturação completa do sistema.

Já os clínicos e os investigadores necessitam de uma interface simples e intuitiva, que lhes permita efetuar comparações e análises complexas, de forma eficiente, precisa e sem a necessidade de conhecimento técnico prévio sobre a aplicação. A constante inserção de novos dados e de novos testes requer uma atualização constante do modelo normativo, e para tal, é necessária uma arquitetura escalável, estável e robusta, capaz de suportar o crescimento contínuo da base de dados.

Assim, este trabalho foi motivado pela necessidade de desenvolver uma solução tecnológica capaz de gerir dados neuropsicológicos de forma sistemática e padronizada e oferecer ferramentas que facilitem a comparação normativa, minimizando as limitações das soluções atuais.

1.2 Oportunidades

O desenvolvimento de uma plataforma web para a recolha, armazenamento e análise de testes neuropsicológicos oferece um conjunto significativo de oportunidades.

No domínio científico, esta solução permite a centralização de dados provenientes de múltiplos estudos, permitindo comparações normativas mais robustas e representativas. Esta aplicação pode, ainda, impulsionar a reprodutibilidade de uma investigação e, ao mesmo tempo, facilitar a colaboração entre instituições e melhorar o processo de comparação.

No domínio clínico, esta plataforma irá ajudar na tomada de decisão, permitindo oferecer uma comparação normativa sólida, onde a análise é feita com subconjuntos da população mais rigorosos, que têm em conta o contexto cultural, socioeconómico

e educacional dos sujeitos. Com o aumento da quantidade e variedade dos dados e com comparações mais dinâmicas, os profissionais de saúde podem interpretar os resultados de forma mais exata e fundamentada.

Do ponto de vista tecnológico, a solução adotada permite alcançar a flexibilidade necessária para lidar com a diversidade de testes e parâmetros, assegurando a escalabilidade e a adaptabilidade. Além disso, a plataforma estabelece uma base sólida para futuras integrações, permitindo análises robustas e personalizadas.

Por fim, a proposta abre perspetivas para o desenvolvimento de análises normativas dinâmicas, que evoluem à medida que novos dados são introduzidos, ampliando a utilidade da plataforma para diferentes populações e contextos de aplicação.

1.3 Objetivos e Abordagem proposta

O presente trabalho foi desenvolvido no âmbito do projeto Apples to Apples, que tem como objetivo a criação de uma plataforma para a recolha, gestão e análise de dados neuropsicológicos, de forma a apoiar os investigadores e clínicos na organização e interpretação dos resultados. Para o atingir foram definidos os seguintes objetivos específicos:

- Desenvolver uma base de dados capaz de integrar a grande quantidade e variabilidade de dados e parâmetros dos testes neuropsicológicos.
- Implementar um modelo Entidade-Atributo-Valor para a modelação da base de dados, garantindo a flexibilidade necessária para acomodar a variabilidade dos testes;
- Desenvolver uma aplicação web para a inserção, gestão e consulta dos dados recolhidos;
- Implementar mecanismos de análise normativa para ajudar na interpretação dos resultados;
- Avaliar a aplicação final com dados reais, a fim de a melhorar e analisar o seu desempenho e viabilidade face à diversidade dos testes.

O modelo Entidade-Atributo-Valor (EAV) é utilizado para descrever entidades que têm uma variedade potencialmente vasta de atributos, mas que utilizam apenas uma pequena parte deles. Como este modelo é otimizado para lidar com a evolução frequente do esquema e a esparsidade dos dados, é ideal para situações em que os atributos das entidades são altamente dinâmicos. [Batra et al., 2018; Dominik Liebler, 2020].

Para lidar com os desafios associados à recolha e análise de dados de testes neuropsicológicos, o presente trabalho propõe uma abordagem baseada no desenvolvimento

de uma aplicação web integrada com uma base de dados relacional. A estrutura da base de dados foi concebida para suportar os diferentes tipos de testes e parâmetros, recorrendo ao modelo de dados já referido.

1.4 Estrutura da Dissertação

Esta dissertação encontra-se organizada em vários capítulos. O primeiro capítulo apresenta a introdução ao tema, incluindo o contexto do problema/limitações a resolver, os objetivos do trabalho e a abordagem. Neste é ainda efetuada uma descrição do que são os testes neuropsicológicos, do que é a comparação normativa, da motivação do trabalho e das oportunidades do mesmo nos domínios científico, clínico e tecnológico. O segundo capítulo apresenta o estado da arte, onde são discutidos os processos atuais na avaliação neuropsicológica, juntamente com as limitações associadas à recolha e interpretação dos resultados. Serão ainda mencionados alguns dos testes mais utilizados, bem como sistemas e plataformas web destinados à gestão e análise dos respetivos resultados. Este capítulo foca-se ainda no desenho e implementação da base de dados numa perspetiva mais técnica e prática e nos métodos e técnicas atualmente aplicados na comparação normativa.

O terceiro capítulo descreve o trabalho realizado, onde é feita uma análise dos dados que serão utilizados ao longo deste trabalho, juntamente com os métodos/técnicas utilizados na sua limpeza, a fim de melhor entender os dados e explorar técnicas para os normalizar durante a execução da aplicação. Neste capítulo, são ainda apresentadas e justificadas as tecnologias selecionadas para a implementação da solução, nomeadamente a base de dados, a linguagem de programação e a *framework web* adotadas. É ainda descrito o desenho da base de dados, a arquitetura da aplicação e as suas principais funcionalidades, com particular destaque para o *upload* dos dados e a comparação normativa, onde é explicada a metodologia da implementação. Por fim, o quarto capítulo apresenta as conclusões do trabalho, onde são discutidas as principais contribuições da dissertação e possíveis direções para o trabalho futuro.

Capítulo 2

Fundamentação Teórica e Estado da Arte

O presente capítulo apresenta a fundamentação teórica e o estado da arte relacionados com o desenvolvimento desta dissertação. Na Secção 2.1 será apresentada a fundamentação teórica necessária para compreender os conceitos subjacentes à avaliação neuropsicológica. Serão descritos na Secção 2.1.1 alguns dos testes neuropsicológicos mais utilizados, bem como o processo de avaliação neuropsicológica na Secção 2.1.2. De seguida, na Secção 2.1.3 será abordada a comparação normativa, distinguindo-se as abordagens univariada e multivariada, e serão apresentados na Secção 2.1.4 os principais métodos estatísticos utilizados para a interpretação dos resultados. De seguida, a Secção 2.1.5 descreve alguns dos modelos e infraestruturas de armazenamento de dados mais utilizados para lidar com informação heterogénea e dinâmica.

Na Secção 2.2, será apresentado o estado da arte onde serão abordadas as soluções existentes nesta área. Serão analisadas nas Secções 2.2.1 e 2.2.2 bases de dados de testes neuropsicológicos e sistemas ou plataformas web destinados à gestão, armazenamento e análise de dados neuropsicológicos. Para cada solução serão identificadas as suas principais características, vantagens e limitações, permitindo compreender os desafios atuais e as oportunidades de melhoria que motivam o desenvolvimento do presente trabalho.

2.1 Fundamentação Teórica

A avaliação neuropsicológica envolve o uso de testes baseados no desempenho, cujos resultados são interpretados com base em normas estabelecidas, permitindo avaliar várias capacidades cognitivas [Harvey, 2012]. Esta avaliação refere-se a um processo de teste de hipóteses multi-método, multi-informativo e multi-conceitual, com o objetivo de quantificar o desempenho cognitivo de um indivíduo. Embora possam ser utilizados testes de rastreio cognitivo, uma bateria mais extensa de testes

neuropsicológicos fornece uma avaliação mais fiável e válida do estado cognitivo de um indivíduo [Kessels and Hendriks, 2023]. Atualmente, os resultados dos testes neuropsicológicos são interpretados através da comparação estatística da pontuação padrão ajustada à idade, educação e/ou género de um indivíduo com uma distribuição normal, utilizando dados normativos, para determinar a probabilidade de um défice cognitivo [Kessels and Hendriks, 2023].

Em alguns casos, as avaliações neuropsicológicas podem até utilizar instrumentos padronizados para avaliar as funções cognitivas, o comportamento e o funcionamento socioemocional [Schaefer et al., 2023].

2.1.1 Testes Neuropsicológicos

Os testes neuropsicológicos englobam uma vasta variedade de instrumentos que permitem avaliar inúmeros domínios cognitivos como: a atenção, a memória, a concentração, a linguagem, o raciocínio lógico, as funções executivas e a perceção visual e espacial, entre outros [Jaqueline Terres Borges, 2025].

Nesta secção são descritos quatro exemplos de testes neuropsicológicos mais utilizados: Trail Making Test, Stroop Color-Word Interference Test, Controlled Oral Word Association Test e Rey-Osterrieth Complex Figure Test, onde cada um foca-se na avaliação de domínios cognitivos específicos.

Trail Making Test

Entre os testes mais utilizados encontra-se o Trail Making Test (TMT), utilizado para avaliar capacidades cognitivas como atenção, flexibilidade cognitiva, velocidade de processamento, reconhecimento de letras e números, leitura visual e função executiva. Está dividido em duas partes, Parte A e Parte B. A Parte A foca-se na atenção visual e na velocidade de processamento, onde o indivíduo efetua a ligação sequencial de números; já a Parte B concentra-se na alternância mental, uma vez que o indivíduo necessita de conectar números e letras em ordem crescente e alternada. O método destaca-se pela sua simplicidade, sensibilidade e aplicabilidade em diversas áreas clínicas e de investigação. É ainda aplicável a várias populações, possui um custo acessível e é rápido e fácil de aplicar [Matheus Santos, 2024; Katie Marvin, 2012].

Stroop Color-Word Interference Test

Outro teste amplamente conhecido é o Stroop Color-Word Interference Test (SCWT). O teste avalia a atenção seletiva (capacidade de ignorar distrações), a inibição (capacidade de controlar impulsos), a flexibilidade cognitiva, a velocidade de processamento, a inteligência fluida e o sistema semântico. É composto por três partes: na primeira, o indivíduo tem de ler o nome das cores escritas a preto; na segunda, tem

de nomear as cores de diversas manchas; e por fim, na terceira, é-lhe solicitado que nomeie a cor da tinta em que as palavras estão escritas, quando as próprias palavras indicam uma cor diferente. As suas medições consistem no número de respostas corretas em quarenta e cinco segundos e o número total de erros para cada parte do teste [Espírito Santo et al., 2015; Scarpina and Tagini, 2017].

Controlled Oral Word Association Test

Para medir a fluência verbal, são normalmente aplicados testes como o Controlled Oral Word Association Test (COWA), que proporciona uma avaliação rápida e eficiente, exigindo que os indivíduos produzam o maior número possível de palavras que comecem com uma letra específica do alfabeto ou que pertençam a uma categoria definida, como animais ou alimentos, num certo período de tempo [Spreen and Risser, 1998]. Está assim dividido em duas partes, letras e categoria, onde em cada uma são medidos três parâmetros: número de respostas corretas, o número de valores repetidos e o número de erros de perda de conjunto.

Rey-Osterrieth Complex Figure Test

As capacidades visuoconstrutiva e memória visual podem ser examinadas através de testes como o Rey-Osterrieth Complex Figure Test (ROCF). Este teste avalia o planeamento, a organização e a atenção através da cópia e reprodução de uma figura complexa, tanto de imediato como após um intervalo de tempo [Zhang et al., 2021]. Está dividido em quatro condições, como: *direct copy*, *immediate recall*, *delayed recall* e *recognition score*. Para as três primeiras são medidos o tempo de conclusão e o resultado. No *direct copy* e no *immediate recall* é ainda medido o intervalo de tempo entre a execução da condição atual e a próxima. Por fim, o *recognition score* mede o número de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos.

Cada um destes testes inclui medidas específicas como o tempo de execução e o número de erros, entre outras. Estas medidas, em conjunto, permitem criar um perfil detalhado do funcionamento cognitivo de um indivíduo.

2.1.2 Processo de Avaliação Neuropsicológica

O método para avaliar um indivíduo começa com uma aferição abrangente, que geralmente é iniciada por uma revisão médica completa, cobrindo o histórico médico e psiquiátrico, medicamentos, resultados laboratoriais e relatórios de neuro-imagem, juntamente com entrevistas clínicas aprofundadas. Em seguida, é administrada uma variedade de testes neuropsicológicos de acordo com os respetivos manuais. Por fim, dependendo dos testes específicos, os resultados são interpretados comparando-os com dados normativos selecionados, tendo em conta o género, a idade, a escolaridade e a etnia do indivíduo [Schaefer et al., 2023].

Atualmente, a recolha dos resultados dos testes neuropsicológicos é frequentemente realizada de forma manual, através de ferramentas tradicionais em papel ou de *softwares* específicos de cada teste [Gallagher et al., 2023]. A pandemia do Covid--19 acelerou a adoção de ferramentas digitais na área da saúde, estendendo-se à neuropsicologia clínica. Esta mudança levou à exploração de novas soluções digitais na avaliação neuropsicológica. Apesar da utilização destes *softwares* ser cada vez mais comum, as ferramentas tradicionais em papel e lápis continuam a dominar. No entanto, as ferramentas tradicionais possuem limitações, tais como [Maggio et al., 2025; Kessels, 2019]:

- Coletam dados de forma infrequente, pois apenas identificam o estado do indivíduo no momento, sem considerar fatores em tempo real como a fadiga e a hora do dia. Desta forma, não são capturadas as variações que ocorrem ao longo do tempo, que podem influenciar significativamente os resultados;
- As suas avaliações normalmente são mais demoradas, reduzindo o número de indivíduos avaliados ao longo do tempo e aumentando as listas de espera;
- São mais propensas a erros humanos, tanto na sua utilização, como na transcrição dos resultados, afetando a confiabilidade e a precisão do diagnóstico;
- Não oferecem funcionalidades de teste adaptativo, mantendo a dificuldade fixa, independentemente do desempenho do indivíduo;
- Alguns testes em papel estão desatualizados e podem resultar em diagnósticos imprecisos.

As ferramentas digitais para testes neuropsicológicos também têm desvantagens. Por exemplo, alguns aspetos do desempenho podem não ser registados, como a recordação livre de material memorizado [Howieson, 2019].

Quando vários destes testes são administrados em conjunto, como parte de uma avaliação, é possível efetuar afirmações sobre o perfil cognitivo de um indivíduo. Estes testes em papel já existem há décadas e são amplamente utilizados e disponíveis publicamente. A necessidade de armazenamento físico torna os registos em papel vulneráveis a perdas e danos. Esta limitação compromete o rastreio e a recolha de dados, restringindo assim a realização de comparações normativas. A adoção de soluções digitais pode resolver várias destas limitações [Kessels, 2019].

Normalmente, em muitos contextos clínicos, após a recolha manual destes dados em formato de papel, os resultados são transcritos para folhas de cálculo, como o Microsoft Excel, ou armazenados em bases de dados. O Microsoft Excel pode ajudar a organizar, analisar e apresentar dados de testes neuropsicológicos: reduz o risco de erros humanos, pode ser utilizado para armazenar dados normativos e possui fórmulas que podem ser utilizadas para comparar o desempenho dos indivíduos [Bryant, A. and Freilich, B. and Papova, A., 2025].

Em termos de análise, a avaliação é por norma efetuada com base na administração de múltiplos testes. No entanto, o aumento no número de testes realizados aumenta

a probabilidade de pelo menos um deles apresentar um resultado falso positivo, ou seja, indicar incorretamente uma anomalia. Isto é conhecido como aumento da taxa de erro familiar (FWER - *Family-Wise Error Rate*), termo que se refere à probabilidade de obter pelo menos um valor falso positivo durante a realização de múltiplos testes. Para diminuir a FWER, é necessário aplicar algumas correções antes da comparação normativa. No entanto, estas correções podem diminuir a capacidade de detetar desvios verdadeiros [Zadelaar et al., 2017].

Existem vários métodos de correção, como a Correção de *Bonferroni*, que reduz o aumento do FWER através da divisão da probabilidade de obter pelo menos um valor falso positivo pelo número de testes realizados. Este método é amplamente utilizado devido à sua simplicidade, apesar de ser excessivamente conservador, pode aumentar o número de falsos negativos e diminuir o poder estatístico [Zadelaar et al., 2017; Physiotutors, 2025].

2.1.3 Comparação Normativa

A comparação normativa constitui uma etapa fundamental na interpretação dos resultados dos testes neuropsicológicos. Tem como objetivo verificar se os resultados de um indivíduo se desviam significativamente do esperado, com base num grupo com características semelhantes. Para realizar esta comparação, existem abordagens diferentes, que variam na forma como os resultados dos testes são considerados. É importante distinguir entre abordagens de análise univariada e multivariada. Nas secções seguintes, será apresentada uma descrição de cada uma delas, destacando as suas vantagens e limitações.

Comparação Univariada

Na análise univariada, a comparação é feita individualmente para cada variável ou parâmetro do teste, ou seja, o resultado do indivíduo é comparado com as pontuações do grupo normativo para a variável ou parâmetro do teste específico. Por comparar cada resultado separadamente, esta abordagem recomenda a aplicação de métodos de correção [van Rentergem et al., 2017].

Com os dados normativos selecionados, no caso da comparação univariada, um resultado é considerado anormal quando a pontuação do indivíduo fica abaixo de um certo percentil [Huizenga et al., 2007].

Existem vários métodos para determinar o resultado da comparação normativa univariada, que serão explicados mais à frente neste capítulo na secção 2.1.4. Alguns métodos tradicionais consistem na divisão da amostra normativa em subgrupos com base em fatores demográficos, como idade, género e escolaridade e, de seguida, é calculada a média e o desvio padrão de cada subgrupo. Para o subgrupo correspondente ao do indivíduo são calculadas métricas (diretamente relacionado com os percentis), que permitem entender o desempenho do indivíduo. Outro método é o

baseado na regressão, que utiliza a análise de regressão para modelar a relação entre o resultado do teste e as variáveis demográficas [delCacho Tena et al., 2023].

Para além disso, o modelo de regressão prevê o resultado normativo, com base nas suas características demográficas. Este valor é obtido através da distribuição cumulativa de resíduos padronizados. Uma das vantagens deste método é a sua capacidade de controlar parâmetros importantes, como a normalidade e a multicolinearidade, e de identificar possíveis casos influentes nos dados. Permite ainda a utilização da amostra na sua totalidade para calcular os dados normativos, garantindo assim uma maior precisão. Outro benefício é que o método possibilita visualizar a significância e o valor das variáveis sociodemográficas incluídas no modelo para a amostra em questão [delCacho Tena et al., 2023].

Comparação Multivariada

Na análise multivariada, o perfil cognitivo completo (o conjunto de resultados em vários testes) é comparado com a distribuição multivariada desses testes no grupo normativo, considerando a relação entre eles. A comparação multivariada é mais eficaz na identificação de desvios e dispensa a utilização de métodos de correção, uma vez que se baseia numa única comparação [van Rentergem et al., 2017].

O perfil de resultados dos testes do indivíduo é comparado com a distribuição dos resultados na amostra normativa, através da estimativa da covariância entre os diferentes resultados do teste [van Rentergem et al., 2017]. Normalmente, os dados normativos são recolhidos individualmente para cada teste, resultando em poucos dados normativos multivariados [de Vent et al., 2016]. Dadas estas restrições, vários *datasets* de indivíduos saudáveis foram e continuam a ser desenvolvidos, frequentemente agregados a partir de vários estudos de investigação [van Rentergem et al., 2017].

2.1.4 Métodos Estatísticos para Comparação Normativa

Tal como mencionado na secção 2.1.3, as comparações normativas classificam-se como univariadas ou multivariadas. Na abordagem univariada, cada variável é analisada individualmente, enquanto que as multivariadas recorrem ao perfil completo do indivíduo para efetuar a comparação com os respetivos grupos normativos. Após a recolha dos dados do indivíduo e a seleção do grupo de dados normativos, são aplicados métodos para realizar as comparações e avaliar o indivíduo.

Comparação Normativa Univariada

Para efetuar comparações normativas univariadas existem vários métodos/medidas estatísticas como: *T-test*, *Mann-Whitney U Test*, *Analysis of Variance* (ANOVA), métodos baseados na regressão, *Z-scores*, *T-scores*.

T-test é um teste estatístico usado para determinar se a diferença entre as médias de dois grupos é estatisticamente significativa. Este teste calcula o `t-value` com base nas médias, desvios padrão e tamanhos das amostras dos dois grupos. O valor calculado é utilizado para definir o `p-value`, através da sua comparação com um valor crítico da distribuição `t`. O `p-value` representa a probabilidade de obter um resultado tão extremo assumindo que não existe diferença entre as médias; assim, um valor baixo normalmente indica que existem diferenças significativas [Pandey, 2015; Wadhwa and Marappa-Ganeshan, 2023; Emerson, 2023; Kim, 2019; Analytical Methods Committee AMCTB No. 93, 2020]. Existem vários tipos de *T-test*, como [Pandey, 2015; Wadhwa and Marappa-Ganeshan, 2023; Kim, 2019; Mishra et al., 2019]:

- *One-sample t-test* - Compara a média de uma única amostra com a média de uma população, a fim de verificar se a amostra poderia pertencer a essa população;
- *Two-sample t-test* (independent samples) - Compara as médias de dois grupos distintos e separados;
- *Paired/dependent t-test* - Compara as médias de duas medições relacionadas, por exemplo, calcula a diferença entre duas observações de um indivíduo, antes e depois do tratamento.

Este teste possui algumas limitações: é inapropriado para comparações que envolvem mais do que um grupo; é sensível aos valores atípicos (observações que se afastam significativamente da maioria dos dados de um conjunto); se os dados não possuírem uma distribuição normal, ou se as variações dos dois grupos não forem iguais, ou se os dois grupos não forem independentes, então os resultados não são confiáveis [Martin, 2025].

Mann-Whitney U Test é um teste estatístico não paramétrico utilizado para comparar duas amostras ou grupos, avaliando se ambos provêm da mesma população. Este teste calcula o `U-value` para cada grupo após atribuir classificações ordinais aos valores da amostra combinada (ambos os grupos) por ordem crescente. Posteriormente, utiliza o tamanho das amostras e o nível de significância para determinar o valor crítico através da respetiva tabela de referência. Se o `U-value` calculado for inferior ou igual ao valor crítico, conclui-se que existem diferenças significativas entre os grupos [Elliot McClenaghan, 2024].

Estes testes possui algumas limitações: não é adequado para dados emparelhados; apenas permite comparar dois grupos de cada vez; apenas indica a existência de diferenças entre os grupos, mas não especifica qual deles possui os valores mais baixos ou mais altos; requer amostras com maior dimensão para garantir resultados mais robustos, quando comparado com outros métodos estatísticos [Pannmie, 2023].

Analysis of Variance (ANOVA) é um teste estatístico utilizado para avaliar a diferença entre as médias de dois ou mais grupos. Tem como objetivo determinar se as variações observadas são aleatórias ou se representam diferenças genuínas e estatisticamente significativas. Existem dois tipos de ANOVA: *one-way* e *two-way*. *One-way* compara as médias de três ou mais grupos independentes, visando determinar diferenças significativas com base numa única variável independente. *Two-way* analisa as médias considerando duas variáveis independentes em simultâneo, o que possibilita determinar os efeitos individuais de cada variável e a existência de interações entre elas [Kenton, 2025].

Para avaliar se as médias são diferentes, este teste compara a variância explicada (causada pelos campos de entrada) com a variância inexplicada (causada pela fonte de erro). Uma divergência elevada entre elas indica diferenças estatisticamente significativas entre as médias [IBM, 2025a]. Este método também apresenta algumas limitações: exige que os dados em cada grupo sigam uma distribuição normal; assume que a variância dos dados é homogénea em todos os grupos; a sua fiabilidade é reduzida quando a dimensão da amostra é insuficiente ou os dados não são recolhidos aleatoriamente; requer que as variáveis independentes não afetem outros pontos de dados [testbook, 2025].

Z-scores é uma medida estatística que descreve a posição relativa de um valor face à média de um grupo de valores. A sua computação envolve a diferença entre o valor individual e a média, dividida pelo desvio padrão, sendo que um `Z-score` mais próximo de zero indica proximidade à média [Andrade, 2021]. Embora seja uma ferramenta simples e poderosa, a fiabilidade do `Z-score` está condicionada à qualidade e à assunção da normalidade da distribuição dos dados. Pode, ainda, ser comprometido por assimetrias ou pela presença de valores atípicos nos dados [fawwazmts, 2024].

T-score é uma pontuação padronizada, utilizada na comparação normativa, com uma média de 50 e um desvio padrão de 10. Permite comparar o resultado de um indivíduo com o da sua população de referência, através do cálculo inicial do `Z-score`, seguido da determinação do `T-score` correspondente. O `T-score` é calculado pela fórmula: $t_{score} = 10 * z_{score} + 50$. Consequentemente, quanto mais próximo o `T-score` for de 50, menores serão as diferenças significativas em relação à média normativa [Iverson, 2011]. Esta medida apresenta as seguintes limitações: a fiabilidade depende da qualidade do grupo normativo; o `T-score` pode ser complexo de interpretar, particularmente para utilizadores menos experientes; perde relevância entre populações, uma vez que o grupo de referência pode ser diferente da população real a nível cultural ou demográfico, comprometendo a validade da comparação [McKee, 2025].

Métodos baseados na regressão são um tipo de normalização contínua. Para além deste tipo de normalização, ainda existem outros, como: normalização semi-

paramétrica [Timmerman et al., 2019; Urban et al., 2025] e normalização inferencial [Timmerman et al., 2019; Urban et al., 2025]. A normalização contínua é um método que utiliza a amostra total de um estudo para calcular normas através de uma função contínua, em vez de dividir os dados em grupos isolados. Utiliza explicitamente a natureza contínua ou categórica ordenada das variáveis independentes (como a idade ou o nível de escolaridade) para o cálculo das normas, podendo também ser aplicado às variáveis independentes discretas (como o sexo) [Timmerman et al., 2019; Urban et al., 2025]. A normalização tradicional difere bastante da contínua, uma vez que consiste em dividir variáveis contínuas, como a idade, em intervalos fixos (categorias) e calcular normas específicas para cada um desses grupos isoladamente [Timmerman et al., 2019].

Os métodos baseados em regressão permitem comparar uma pontuação com um grupo padrão utilizando um modelo de regressão para prever a distribuição das pontuações com base em variáveis independentes/características, como a idade. Um dos métodos baseados na regressão é o *Generalized Additive Models for Location, Scale, and Shape* (GAMLSS), que possibilita modelar a distribuição dos resultados brutos. Após o ajuste do modelo, a função de distribuição cumulativa (CDF) permite transformar os dados brutos em dados normalizados. Deste modo, é possível determinar se o resultado de um indivíduo se enquadra num determinado intervalo percentual, considerando as suas características demográficas [Blagus et al., 2025; Urban et al., 2025].

As vantagens destes métodos baseados na regressão sobre a normalização tradicional são: maior flexibilidade ao gerar normas mais realistas e maior eficiência ao permitir obter a mesma precisão com menos dados [Timmerman et al., 2019].

Comparação Normativa Multivariada

Para efetuar comparações normativas multivariadas existem vários métodos/medidas estatísticas como: *Mahalanobis Distance*, *Hotelling's* T^2 *statistic*, abordagens baseadas na regressão multilinear e modelos de *deep learning*, *Bayesian Frameworks*, *Baringhaus and Franz's Cramér statistic* (BFC).

Mahalanobis Distance é uma medida estatística que indica a posição de um perfil de dados num modelo de calibração multivariado, ou seja, é uma medida que calcula a distância entre um ponto e uma distribuição num espaço multivariado [Mark and Workman, 2018]. Na prática, esta métrica quantifica a distância de um ponto ao centroide do conjunto de dados multivariados. O centroide é definido como o ponto de interseção das médias de todas as variáveis independentes. Assim, uma distância maior indica um resultado menos favorável face ao padrão médio do grupo normativo [Stephanie Glen, 2025b].

A fórmula para calcular a distância é: $d = \sqrt{(x-\mu)^T C^{-1}(x-\mu)}$, onde x corresponde ao vetor que representa os dados do indivíduo, μ corresponde ao vetor

da média da distribuição dos dados e C corresponde à matriz da covariância da distribuição. Um dos seus maiores problemas consiste na impossibilidade de calcular o inverso da matriz da covariância para distribuições altamente correlacionadas [Stephanie Glen, 2025b].

Hotelling's T^2 statistic é um teste multivariado que generaliza o `t-test`, já referido, para duas ou mais variáveis dependentes. O resultado estatístico é obtido a partir da distância de *Mahalanobis*, considerando a covariância entre as variáveis e a `distribuição F`. Existem duas versões diferentes para este teste: uma verifica se um vetor da média da amostra é igual a um vetor da média hipotético e específico; a outra determina se as médias de dois grupos são significativamente diferentes, considerando várias variáveis simultaneamente [Stephanie Glen, 2025a].

O funcionamento deste teste é análogo ao do `t-test`. Após o cálculo das médias, determina-se a estatística T^2 de *Hotelling* (equivalente ao `t-value` multivariado), que é de seguida comparada com um valor crítico da `distribuição F`. Se o valor calculado for superior ao da tabela `F` para as condições da população, conclui-se que as duas amostras provêm de populações com médias multivariadas diferentes [Stephanie Glen, 2025a]. Este teste é bastante sensível aos valores atípicos, o que reduz a confiabilidade dos resultados e pode levar a classificações incorretas [Mokhtar et al., 2021].

Abordagens baseadas na regressão multilinear são também aplicadas na comparação normativa multivariada. Utilizam modelos de regressão para comparar o perfil completo de resultados de um indivíduo com uma amostra normativa, considerando as relações entre as variáveis. Assim, estas abordagens normalizam vários testes num só modelo, assumindo a normalidade multivariada e a igualdade das variâncias dos resíduos [Innocenti et al., 2024]. Estas abordagens podem ser combinadas com estatísticas normativas multivariadas, como *Hotelling's* e *Mahalanobis Distance*, para classificar o desempenho dos indivíduos [Innocenti et al., 2024].

Este método permite uma melhor avaliação dos indivíduos, dado que considera a inter-relação entre os resultados, conduzindo a diagnósticos mais precisos e à minimização de erros. Reduz também o erro Tipo I, evitando o aumento de falsos positivos através da diminuição do número de testes desnecessários [Innocenti et al., 2024].

Baringhaus and Franz's Cramér statistic (BFC) é um teste baseado em permutações para o problema multivariado com duas amostras, que é livre de pressupostos distributivos, ou seja, permite comparar um indivíduo com um grupo normativo, sem assumir que os dados seguem uma distribuição específica. Mede a distância euclidiana dos resultados de um indivíduo com cada sujeito do grupo normativo. A distância é calculada através da diferença das pontuações em vários testes. Posteriormente, o teste compara a distância média do indivíduo em relação aos normativos

com a distância média entre os próprios normativos. Um indivíduo é considerado normal se a sua distância média aos normativos é semelhante à distância média entre quaisquer dois normativos [Grasman et al., 2010; Baringhaus and Franz, 2004].

O BFC pode ser utilizado como uma alternativa viável a testes como o *T-squared* de *Hotelling*. Permite controlar a taxa de erro tipo I (a taxa de rejeição de uma hipótese nula incorretamente) e é sensível o suficiente para detetar padrões anormais nos dados [Grasman et al., 2010].

Modelos de Aprendizagem Profunda é outra abordagem usada recentemente para realizar comparações normativas multivariadas. Utiliza redes neuronais artificiais (ANNs), com arquiteturas recorrentes e/ou convolucionais, que são mais complexas, flexíveis e poderosas estatisticamente. A aprendizagem profunda suporta análises multivariadas complexas que podem ser aplicadas a comparações normativas, através da criação de modelos mais complexos e sofisticados, sendo capaz de identificar e distinguir fontes de ruído e sinais de interesse complexos [Kuntzelman et al., 2021].

Esta abordagem possui um bom desempenho e é adequado para conjuntos de dados complexos. Apresenta uma alta flexibilidade, que permite a sua utilização em diversas tarefas, para além da classificação multivariada. No entanto, apresenta limitações como: mais propenso ao *overfitting*, fenómeno que ocorre quando um modelo se ajusta excessivamente aos dados de treino, memorizando-os e captando não apenas os padrões relevantes, mas também o ruído (informação irrelevante, erros ou valores atípicos), o que compromete a sua capacidade de generalização; desempenho inferior em conjuntos de dados mais pequenos; exigência de maior poder computacional, para processar um maior volume de dados; necessidade de arquiteturas mais complexas, que podem dificultar a sua compreensão e a interpretação dos resultados obtidos [Kuntzelman et al., 2021].

Bayesian Frameworks são abordagens estatísticas que aplicam o *Teorema de Bayes* para atualizar a probabilidade de uma hipótese ou parâmetro, combinando conhecimentos prévios com novas evidências. A regra de *Bayes* estabelece que a probabilidade de uma proposição ser verdadeira é refinada à medida que novos dados são obtidos, integrando a probabilidade prévia com a das novas evidências. O resultado é uma probabilidade mais informada, que une o conhecimento prévio com as novas observações [Mirzaei and Adeli, 2022].

Esta abordagem pode ser aplicada na comparação normativa multivariada para calcular as probabilidades de diagnóstico, com base no perfil de resultados de um indivíduo e utilizando estatísticas descritivas de amostras de referência. É uma ferramenta prática e flexível para apoiar a tomada de decisões clínicas e de investigação em neuropsicologia, embora a sua principal limitação seja a necessidade de um maior volume de dados representativos para a comparação, que nem sempre estão disponíveis para certas populações. [Goette et al., 2023].

Para além das abordagens e modelos já mencionados, existem muitas outras metodologias que permitem realizar comparações normativas multivariadas. Esta área de investigação está em constante evolução, dada a sua importância e a necessidade de superar limitações como o acesso a certos grupos normativos. Consequentemente, novos modelos são continuamente desenvolvidos, seja através da melhoria de estatísticas/abordagens existentes com novas estratégias, seja pela introdução de novos modelos analíticos.

2.1.5 Modelos e infraestruturas de armazenamento de dados

Para armazenar os dados, podem ser utilizadas várias abordagens, como bases de dados relacionais ou NoSQL [Wahyuningrum et al., 2019]. Uma base de dados relacional é um tipo de base de dados que organiza os dados em tabelas, cada uma com linhas (representando diferentes observações) e colunas (representando os atributos de cada observação), estabelecendo relações entre elas. As tabelas podem ser unidas através de chaves primárias ou chave estrangeiras. Estas chaves estabelecem diferentes relações entre as tabelas [IBM, 2025c]. A estrutura deste tipo de base de dados é representada através de um esquema, que define como os dados estão organizados, que ilustra a arquitetura da mesma, ou seja a representação visual das tabelas e das suas relações e que inclui restrições lógicas, como o nome das tabelas e dos atributos [IBM, 2026].

As bases de dados NoSQL são sistemas não relacionais criados para grandes volumes de dados, diversificados e não estruturados. Armazenam os dados numa única estrutura sem esquema, oferecendo uma rápida escalabilidade e alta disponibilidade através de armazenamento distribuído e replicado em vários sistemas. Melhoram a disponibilidade e a tolerância a falhas, uma vez que a base de dados permanece operacional mesmo que algumas fontes de dados estejam *offline*, embora possa introduzir inconsistências, uma vez que os dados não estão normalizados [IBM, 2025b].

No entanto, para cenários em que a variabilidade dos atributos é elevada, existem abordagens ao nível do modelo de dados. O modelo Entidade-Atributo-Valor EAV é um padrão de *design* de bases de dados flexível, concebido para o armazenamento de dados [Brandt et al., 2002; Batra et al., 2018]. Este modelo será explicado posteriormente, na secção 3.4.

2.2 Estado da Arte

Após a apresentação dos conceitos fundamentais relacionados com a avaliação neuropsicológica e a comparação normativa, esta secção irá analisar as soluções atualmente disponíveis nesta área. Serão identificadas algumas das principais bases de dados, plataformas e sistemas desenvolvidos para o armazenamento, gestão e análise de dados neuropsicológicos. A análise destas soluções ajuda a perceber como funcionam, as suas vantagens e limites, permitindo identificar desafios e oportunidades

de melhoria que justificam este trabalho.

2.2.1 Bases de Dados de Testes Neuropsicológicos

Embora existam *softwares* para interpretar dados de testes neuropsicológicos, existem também diferentes soluções para a sua gestão e armazenamento. Estas incluem bases de dados centralizadas, bases de dados integradas nas plataformas e bases de dados locais geridas pelos clínicos [Gordon et al., 2005]. Estas bases de dados consistem em vastos repositórios de dados estruturados de indivíduos saudáveis, organizados por variáveis demográficas como idade, género e escolaridade. Permitem que os clínicos comparem os resultados individuais dos indivíduos com a população representativa, facilitando o diagnóstico e a monitorização do progresso.

As bases de dados na área da neuropsicologia são essenciais para a comparação normativa, uma vez que, sem elas, seria quase impossível efetuar comparações normativas para certas populações. Fornecem a base para interpretar o desempenho de um indivíduo e proporcionam uma melhor compreensão das populações, ao analisar tendências e resultados em diferentes condições e grupos demográficos [Robinson and Finley, 2025]. Possibilitam também a criação de normas locais para certas variáveis demográficas, facilitando a deteção de doenças mais cedo e aumentando a precisão das comparações. Ao oferecerem acesso a dados abrangentes e anonimizados, impulsionam a investigação, permitindo que os investigadores desenvolvam estudos originais e ganhem experiência em análise de dados. Também podem facilitar colaborações entre diferentes centros e proporcionar oportunidades de investigação, permitindo a partilha de dados e estudos, especialmente em contextos onde estes são limitados [Robinson and Finley, 2025].

Exemplos de plataformas e repositórios de dados neuropsicológicos

Existe uma vasta diversidade de bases de dados normativas: algumas são dedicadas a doenças e populações específicas, outras a testes específicos, enquanto outras são mais genéricas, abrangendo múltiplos testes e populações. Adicionalmente, distinguem-se aquelas para testes realizados remotamente, as que compilam dados de avaliações presenciais, as que integram ambos os formatos, e ainda aquelas que possibilitam a agregação de dados provenientes de diversas outras bases. Existem ainda bases de dados para comparações univariadas e multivariadas.

Alguns exemplos de bases de dados normativas para testes neuropsicológicos são: Advanced Neuropsychological Diagnostics Infrastructure (ANDI) [de Vent et al., 2016], Distributed Archives for Neurophysiology Data Integration (DANDI), Indonesian Boston Naming Test Database (I-BNT) [Sulastri et al., 2018], Brain Resource International Database (BRID) [Gordon et al., 2005; Poznanski, 2006; Neuroimaging Research for Veterans Center, 2025] e Uniform Data Set (UDS) [National Alzheimer's Coordinating Center (NACC), 2025; Health Center Program, 2016].

A Advanced Neuropsychological Diagnostics Infrastructure (ANDI) é um sistema de base de dados dinâmico, construído a partir da combinação de múltiplos *datasets* de diversos grupos de investigação. Cada um contém resultados de testes neuropsicológicos de indivíduos saudáveis, variando em termos de variáveis demográficas, testes aplicados e objetivos dos estudos. A sua estrutura e a vasta variedade de dados normativos possibilitam comparações mais precisas, superando as bases de dados tradicionais, nomeadamente para os testes mais populares [de Vent et al., 2016].

Para integrar várias bases de dados numa só e efetuar comparações normativas, válidas e precisas sobre todos os dados, é necessário efetuar alguns ajustes aos dados, que passam por várias etapas, como: remoção de valores impossíveis, integração e codificação demográfica, determinação do grau de importância para cada teste, remoção de valores atípicos corrigidos demograficamente e normalização das distribuições [de Vent et al., 2016].

A Brain Resource International Database (BRID) é uma base de dados padronizada e extensa, com mais de 4000 participantes, incluindo indivíduos normativos e aqueles com doenças neurológicas e psiquiátricas. Foi a primeira a integrar, de forma padronizada, uma vasta gama de dados: de neuro-imagens (Electroencephalogram (EEG), ressonâncias magnéticas e Functional Magnetic Resonance Imaging (fMRI)), de genética (como, por exemplo, dados de polimorfismos genéticos), dados fisiológicos, avaliações neuropsicológicas e cognitivas, resultados de testes neuropsicológicos e de personalidade, informações demográficas e clínicas, históricos médicos, medições elétricas das funções cerebrais e corporais, e outros fatores psicológicos [Gordon et al., 2005; Poznanski, 2006; Neuroimaging Research for Veterans Center, 2025].

Os dados são adquiridos através de equipamentos e procedimentos idênticos em mais de 70 laboratórios internacionais, permitindo o agrupamento robusto de grandes conjuntos de dados para investigação. Os principais objetivos deste repositório são: identificar e quantificar as diferenças individuais na função cerebral normativa, comparar o desempenho de um indivíduo com a sua população de referência e oferecer uma estrutura normativa robusta para a avaliação clínica, diagnóstico e prognóstico de tratamento [Gordon et al., 2005; Poznanski, 2006]. Grande parte dos dados processados são disponibilizados gratuitamente a cientistas para fins de investigação e publicação científica, aumentando os seus benefícios [Neuroimaging Research for Veterans Center, 2025; Poznanski, 2006].

A determinação dos protocolos de padronização foi um dos maiores desafios no desenvolvimento desta base de dados, uma vez que foi necessário lidar com vários obstáculos como: dificuldade em obter um consenso global entre cientistas sobre as informações necessárias a recolher. A seleção dos testes neuropsicológicos e dos modelos de investigação a utilizar, bem como de outras informações relevantes, mostrou ser um processo complexo. [Gordon et al., 2005].

Indonesian Boston Naming Test Database (I-BNT) é uma base de dados para armazenar dados normativos recolhidos para o Indonesian Boston Naming Test (I-BNT). I-BNT é uma versão culturalmente adaptada do Boston Naming Test (BNT) para

a população da Indonésia [Sulastri et al., 2018]. BNT é um teste para a avaliação neuropsicológica, que consiste em nomear várias figuras baseadas em linhas, apresentadas por ordem crescente de dificuldade [Han et al., 2011]. As adaptações culturais realizadas envolveram a alteração de algumas imagens e palavras-alvo. Estas modificações foram baseadas no consenso de linguistas, neuropsicólogos e investigadores indonésios, que selecionaram elementos pertencentes ao conhecimento local e consideraram o nível de dificuldade dos mesmos [Sulastri et al., 2018].

Esta base de dados disponibiliza resultados normativos ajustados por diversas variáveis demográficas, como idade e nível de educação, permitindo aos neuropsicólogos avaliar as capacidades linguísticas na Indonésia com maior precisão. Muitos dos seus dados, ou até mesmo o repositório completo, são incorporados na infraestrutura I-ANDI, uma versão da ANDI adaptada para a Indonésia [Wahyuningrum et al., 2023].

Uniform Data Set (UDS) é um sistema que funciona como uma base de dados para a recolha de dados padronizados, destinados a relatórios e a investigações específicas (por exemplo, no estudo do envelhecimento e da demência). Inclui dados longitudinais, recolhidos desde 2005 através de avaliações anuais padronizadas aos participantes do UDS nos Alzheimers Disease Research Centers (ADRCs). Este sistema fornece informações padronizadas sobre o desempenho dos indivíduos ao longo dos anos e sobre o funcionamento dos centros de saúde dedicados ao estudo e à prestação de serviços a populações afetadas pela doença do Alzheimer. Os participantes do estudo podem apresentar diferentes estados cognitivos, desde demência e comprometimento cognitivo leve até cognição normal [National Alzheimer's Coordinating Center (NACC), 2025; Health Center Program, 2016].

As avaliações anuais são realizadas nos ADRCs através de um conjunto de testes neuropsicológicos que medem domínios cognitivos como funcionamento global, executivo, aprendizagem, memória verbal e auditiva, atenção, linguagem e velocidade de processamento [Rockholt et al., 2025; Weintraub et al., 2009]. A base de dados inclui dados demográficos, características do início e evolução dos sintomas, histórico médico pessoal, medicações, histórico familiar de demência e resultados de exames neurológicos e neuropsicológicos [Weintraub et al., 2009]. Ao garantir a recolha consistente de dados cognitivos por todos os ADRCs e anualmente aos mesmos participantes, o UDS facilita a investigação em grande escala e fornece informações valiosas sobre as alterações cognitivas ao longo do tempo.

Foi ainda desenvolvido um *web site calculator*, que estima uma gama de `z-scores` para o desempenho individual nos testes neuropsicológicos da UDS. É uma calculadora web interativa, baseada na regressão para pontuações normativas, que serve como um recurso adicional aos investigadores clínicos do UDS. Esta ferramenta facilita a interpretação do desempenho individual, especialmente em situações onde a distinção entre dificuldades leves e problemas mais sérios é ambígua. A representatividade da população geral é um desafio nesta aplicação, uma vez que os participantes estudados são maioritariamente indivíduos com elevado nível educacional e da mesma etnia. Adicionalmente, o número insuficiente de resultados para

subpopulações minoritárias limita a sua capacidade de estabelecer normas e, consequentemente, limita a sua utilização [Shirk et al., 2011].

Para além destes repositórios, existem ainda outras bases de dados com dados normativos de testes neuropsicológicos, incluindo aquelas pertencentes a plataformas online, onde as avaliações são realizadas de forma remota.

Para os repositórios privados de cada clínico/investigador, é possível utilizar certas ferramentas estatísticas para a comparação normativa, como, por exemplo, o *software R*. Este *software* destina-se à computação estatística e incorpora diversos pacotes, como o *cNorm* e o *NormData*, que permitem determinar resultados normalizados e efetuar comparações através da derivação de dados normativos baseados na regressão. O *NormData* inclui funções para ajustar modelos de regressão à média e à variância, testar premissas, gerar tabelas normativas ou folhas de resultados automáticas e estimar intervalos de confiança [CRAN, 2025]. Já o *cNORM* foca-se no cálculo de pontuações normativas a partir de variáveis como idade ou escolaridade, recorrendo à regressão polinomial para modelar a relação entre pontuações brutas, normativas e explicativas de forma contínua [Gary et al., 2021].

Repositórios de literatura científica

Para além dos sistemas e bases de dados focados na avaliação neuropsicológica, existem também repositórios de literatura científica que desempenham um papel fundamental na divulgação de informação nesta área.

PubMed, Pub Psych e Web of Science são repositórios e sistemas de recolha de informação destinados a clínicos, investigadores e estudantes para fins de investigação académica, embora cada um possua âmbitos e finalidades distintas. PubMed consiste num recurso gratuito, que serve de suporte à investigação e à recolha de literatura, visando melhorar a saúde a nível global e individual. Possui uma vasta coleção de citações e resumos de artigos, com acesso aos textos completos, cobrindo as áreas da biomedicina, saúde e disciplinas relacionadas, nomeadamente ciências da vida, ciências comportamentais, ciências químicas e bioengenharia [PubMed, 2025]. Pub Psych é um sistema gratuito para a recolha de informações sobre recursos psicológicos. Disponibiliza um vasto conjunto de recursos provenientes de uma variedade de bases de dados internacionais, maioritariamente europeias, atendendo às necessidades de psicólogos académicos e profissionais [Scolary, 2019]. Web of Science é uma base de dados bibliográfica que agrega artigos académicos a nível mundial. Oferece ferramentas para pesquisa avançada, comparação/avaliação de investigações científicas, análise de citações e bibliometria. A plataforma indexa artigos de uma ampla variedade de áreas, incluindo publicações do Google Scholar [EUI(European University Institute), 2025].

2.2.2 Sistemas e Plataformas Web

Já existem vários sistemas para fazer a gestão de dados de testes neuropsicológicos. São também utilizados *benchmarks* para determinar a pontuação normativa de uma pessoa, derivada da média e do desvio padrão de um grupo populacional saudável. Nas secções seguintes são apresentados alguns dos trabalhos mais relevantes na área.

PsiNorm

O PsiNorm é um *software* gratuito e de código aberto disponível para investigadores e clínicos que realizam testes neuropsicológicos. Foi concebido principalmente para distúrbios neurocognitivos em adultos na Turquia e para simplificar as suas avaliações cognitivas. Incorpora testes padronizados frequentemente utilizados, é compatível com a maioria dos sistemas operativos (Microsoft Windows, Linux e MacOs), utiliza o Python 3.6 e o sistema de dados é baseado em JSON, permitindo que todos os dados, testes e até mesmo menus sejam modificados. É licenciado através da GNU General Public Licence (GNU GPL v3) [Ayhan et al., 2022].

Esta ferramenta reduz significativamente o tempo necessário para calcular percentis e normas, além de ajudar a gerar relatórios preliminares rapidamente. Permite ainda armazenar de forma adequada os resultados dos indivíduos. Os testes incluídos foram escolhidos entre aqueles que são aplicados para a avaliação neuropsicológica. O *software* calcula os `z-scores` e apresenta os percentis correspondentes (quando as normas estão disponíveis) ou determina a posição relativa de um resultado com base num valor limite [Ayhan et al., 2022].

A aplicação possui uma interface maioritariamente gráfica. Primeiramente, é inserida a informação do indivíduo. De seguida, selecionam-se os testes a serem aplicados e registam-se os seus resultados. Uma vez concluído, o sistema gera automaticamente relatórios preliminares personalizados para cada teste. Estes relatórios abrangem dados demográficos, datas dos testes, os resultados obtidos e informações detalhadas sobre as normas e referências utilizadas. Para maior flexibilidade, os dados da base de dados com os resultados para cada indivíduo podem ser exportados para Microsoft Excel ou CSV, facilitando a análise e a partilha [Ayhan et al., 2022].

Para determinar as pontuações, o *software* integra um calculador de normas de Shirk [Shirk et al., 2011], que opera via Microsoft Excel e utiliza a base de dados normativa do Alzeimer's Disease Coordinating Center (ADCC) [Ayhan et al., 2022; Shirk et al., 2011] e do Uniform Data Set (UDS) [Ayhan et al., 2022; Shirk et al., 2011]. Caso as normas para o grupo demográfico do indivíduo não existam, o PsiNorm recorrerá ao grupo demográfico mais próximo para o cálculo dos valores. Contudo, o sistema não valida nem indica se os dados do indivíduo se enquadram nos critérios demográficos das normas utilizadas [Ayhan et al., 2022].

Algumas das suas vantagens são [Ayhan et al., 2022]:

- Cria automaticamente uma base de dados organizada;
- É uma ferramenta simples de utilizar, e ao mesmo tempo, simplifica o processo da avaliação cognitiva;
- É uma ferramenta adaptável, que reduz o trabalho, os erros e a perda de dados no campo da avaliação neurocognitiva.

No entanto, possui algumas desvantagens/limitações [Ayhan et al., 2022]:

- Não inclui todos os testes neurocognitivos utilizados na Turquia, nem os testes que não foram estudados. Apesar de permitir adicionar novos testes, não existem dados normativos suficientes sobre os mesmos para efetuar as comparações;
- A interface de navegação e as informações operacionais da aplicação/testes e explicações encontram-se em turco. Assim, a sua utilização internacional é limitada;
- Herda as limitações dos testes incluídos;
- A precisão dos cálculos está limitada aos dados normativos da população;
- O sistema não valida, nem indica se os dados do indivíduo se enquadram nos critérios demográficos das normas utilizadas. Assim, acarreta um risco significativo de diagnósticos incorretos, resultados inválidos e, consequentemente, uma perda da confiança e credibilidade na aplicação.

Automated Neuropsychological Assessment Metrics (ANAM)

O Automated Neuropsychological Assessment Metrics (ANAM) é um sistema de *software* comercial concebido para avaliações neurocognitivas e gestão de traumatismos cranioencefálicos (TCE) ao longo do tempo. É uma ferramenta para a avaliação cognitiva, concebida para detetar a velocidade e a precisão da atenção, da memória e da capacidade de raciocínio. É composta por uma biblioteca de testes computadorizados de domínios cognitivos, incluindo atenção, concentração, tempo de reação, memória, velocidade de processamento e tomada de decisões, utilizada para avaliações cognitivas e comportamentais. Fornece aos clínicos e aos investigadores dados para avaliar a evolução do estado cognitivo de um indivíduo ao longo do tempo [Defense Health Agency, 2024; Defense and Veterans Brain Injury Center, 2009].

Inclui recursos para testes, geração de relatórios e gestão de dados. Atualmente, possui 22 testes que podem ser agrupados em baterias flexíveis ou padronizadas, altamente sensíveis a alterações cognitivas [VistaLifeSciences, 2025].

A principal característica do ANAM é a capacidade de definir um conjunto de resultados padrão para um teste ou avaliação, que servirão como base de comparação para interpretar o desempenho individual. Desta forma, permite que os profissionais

de saúde comparem os resultados subsequentes dos testes ANAM para indivíduos com suspeita ou confirmação de TCE ou de outras doenças, ajudando a acompanhar a recuperação ou a identificar deficiências contínuas. Estas pontuações também são utilizadas para identificar funcionários cujo desempenho pode estar comprometido, através da comparação com valores normativos ou avaliações previamente realizadas [Defense Health Agency, 2024].

Este *software* integra duas ferramentas essenciais: o ANAM Performance Report (APR) e o ANAM Data Extraction and Presentation Tool (ADEPT). O APR foca-se na apresentação dos resultados dos testes, permitindo a comparação com desempenhos anteriores do indivíduo ou com diversas normas/grupos de referência. Por outro lado, o ADEPT permite aos clínicos e investigadores selecionar, organizar e visualizar dados ANAM de forma eficiente em folhas de cálculo, ou exportá-los para *softwares* estatísticos e gráficos. Este *software* possui ainda uma base de dados relacional baseada em SQL Server e um sistema de relatórios projetado para armazenar, indexar e possibilitar vários requisitos de relatórios [Defense Health Agency, 2024; VistaLifeSciences, 2025].

Algumas das suas vantagens são [Defense Health Agency, 2024; VistaLifeSciences, 2025]:

- Permite a realização de vários testes de forma computorizada, tornando o processo de administração mais eficiente e rápido;
- É sensível a mudanças cognitivas subtis;
- Permite armazenar dados de testes realizados e compará-los com padrões estabelecidos, visando diagnosticar possíveis lesões ou monitorizar a evolução;
- É capaz de efetuar comparações em várias áreas cognitivas.

No entanto, possui algumas desvantagens/limitações [VistaLifeSciences, 2025; Defense and Veterans Brain Injury Center, 2009]:

- Possui um número limitado de testes e não permite a inserção de novos;
- Foi concebida especialmente para lidar com traumatismos cranioencefálicos e para avaliar o desempenho de membros do serviço militar;
- O resultado não é um diagnóstico. Fornece informações ao clínico sobre o funcionamento neurocognitivo que devem ser consideradas, juntamente com outras informações clinicamente relevantes, para auxiliar na tomada de decisões;
- Está limitada aos dados normativos da população, que neste caso, são maioritariamente dados de membros do serviço militar, antes e depois de lesões.

NeurOn

É uma plataforma online para testes cognitivos, onde os clínicos e os investigadores são donos dos seus próprios testes, facilitando a sua gestão. A plataforma disponibiliza os testes mais reconhecidos e utilizados globalmente. Graças à sua execução extremamente eficiente e à avaliação de resultados em apenas segundos, a rapidez é uma característica central da aplicação [NeurOn Neuropsychology Online, 2025].

Uma das suas principais características é a flexibilidade, uma vez que permite integrar e aplicar tanto testes neuropsicológicos padrão já estabelecidos como outros desenvolvidos recentemente de forma remota. Possui ainda a capacidade de desenvolver conjuntos de testes personalizados, adaptados às necessidades específicas do clínico ou investigador. Suporta eficazmente a avaliação transversal e o acompanhamento longitudinal dos indivíduos avaliados, garantindo que nenhuma informação identificável dos mesmos seja recolhida. O sistema faculta o acesso e a exportação de dados brutos, incluindo os dados de cada ensaio para cada participante [NeurOn Neuropsychology Online, 2025].

Algumas das suas vantagens são [NeurOn Neuropsychology Online, 2025]:

- Permite a gestão de milhares de participantes, tornando-a escalável;
- Oferece uma maior flexibilidade;
- Permite a criação de perfis cognitivos de indivíduos e a sua exportação para futuros estudos;
- Garante a segurança dos dados, pois não armazena nenhuma informação pessoalmente identificável dos participantes ou dos clínicos;
- Apresenta um custo reduzido, com o objetivo de promover a adoção generalizada de testes online, aumentando a sua acessibilidade;
- Fornece resultados instantâneos e acesso a dados brutos.

No entanto, possui algumas desvantagens/limitações [Peterson et al., 2025; Morrissey et al., 2024]:

- Falta de validação/ e dados normativos, nomeadamente devido à realização dos testes de forma remota;
- A realização dos testes em ambiente não supervisionado, ou seja, no domicílio do indivíduo e fora de um contexto clínico, aumenta o risco dos resultados serem afetados por fatores diversos, como distrações, comprometendo a sua confiabilidade;
- Os resultados não podem ser considerados equivalentes aos de outras plataformas onde os testes são realizados presencialmente;

- Os estudos nesta plataforma têm sido limitados a indivíduos saudáveis, pelo que poderão ter dificuldades em analisar resultados para certos indivíduos com doenças mais raras;
- Para certas populações, a exigência de realizar testes de forma remota e eletrónica pode constituir um desafio.

Advanced Neuropsychological Diagnostics Infrastructure (ANDI)

É uma plataforma online criada para ajudar os clínicos e os investigadores a avaliar os seus pacientes. A avaliação consiste em identificar se os resultados obtidos são normais ou anormais. O seu componente principal consiste numa base de dados diversificada, composta por dados de múltiplas equipas de investigação, doados por investigadores neuropsicológicos, principalmente dos Países Baixos e da Bélgica flamenga. O conjunto de dados é formado pela agregação de resultados de diversas investigações em indivíduos saudáveis, tornando-os mais representativos. Para os testes neuropsicológicos mais populares, esta combinação oferece uma quantidade e variedade de dados superiores aos normativos convencionais, permitindo avaliações neuropsicológicas mais precisas. Abrange ainda informações de vários testes neuropsicológicos [de Vent, 2021; de Vent et al., 2016].

Graças à sua estrutura, esta base de dados simplifica a comparação normativa, em particular para a avaliação multivariada que envolve o perfil cognitivo completo do indivíduo. É abrangente e representativa, contendo dados normativos de homens e mulheres de todas as idades e níveis de escolaridade [de Vent, 2021; de Vent et al., 2016].

Esta plataforma permite que clínicos e investigadores utilizem os dados normativos que foram especificamente calibrados e recolhidos na sua própria língua e contexto cultural/geográfico, tornando a comparação mais precisa, válida e relevante. O ANDI utiliza normas baseadas na regressão, o que significa que são contínuas e não estratificadas pela idade, género ou nível de escolaridade [de Vent, 2021].

A plataforma web permite aos clínicos e investigadores o *upload* dos dados dos indivíduos que desejam analisar. Após o *upload*, é gerado um relatório detalhado para cada indivíduo, que inclui comparações normativas univariadas e multivariadas [de Vent, 2021].

Algumas das suas vantagens são [de Vent, 2021; de Vent et al., 2016]:

- Devido à vasta variedade de dados normativos de indivíduos saudáveis na base de dados, os clínicos conseguem comparar os resultados de um indivíduo com valores esperados para a população de referência, o que permite uma deteção mais fiável de quaisquer anomalias;
- Lida com as variações na interpretação dos resultados, nomeadamente entre diferentes grupos etários;

- Fornece uma comparação mais representativa para cada indivíduo, uma vez que utiliza um maior volume de dados normativos e permite ajustar a comparação com base nas características demográficas;
- Efetua tanto comparações normativas univariadas, como multivariadas, tornando-a mais eficaz na deteção de anomalias;
- Possui um maior controlo das variáveis demográficas, uma vez que utiliza dados normativos estruturados com base nas características como a idade e o nível de educação;
- Trata-se de uma infraestrutura de código aberto, o que permite a criação de plataformas semelhantes a partir da sua base;
- Promove a partilha de dados, minimizando a necessidade de pesquisa individual por conjuntos de dados normativos para estudos específicos.

No entanto, possui algumas desvantagens/limitações [de Vent et al., 2016]:

- A combinação de vários conjuntos num repositório central de dados normativos, requer um esforço acrescido para assegurar a consistência e o formato dos dados, essenciais para análises fiáveis;
- A qualidade do conjunto de dados depende da qualidade dos dados inseridos. Os vários conjuntos doados podem conter erros, que afetam a confiabilidade;
- Os dados normativos baseiam-se em testes neuropsicológicos tradicionais, que apresentam certas limitações que novas abordagens, como testes que avaliam o comportamento do indivíduo no dia a dia, conseguem superar.

Calibrated Neuropsychological Normative System (CNNS)

Este sistema foi desenvolvido com o propósito de auxiliar clínicos e investigadores na interpretação dos testes nele integrados, utilizando as normas disponibilizadas. Consiste em 24 testes amplamente utilizados e a interpretação dos resultados é efetuada com base em calibrações demográficas e classificações qualitativas, fornecidas pelo utilizador. Os resultados brutos dos indivíduos são processados para gerar valores padronizados, como resultados escalados, `T-scores` (pontuações padronizadas, utilizadas na comparação normativa, com uma média de 50 e um desvio padrão de 10), percentis para 73 resultados de medidas e pontuações fatoriais [Schretlen et al., 2025]. As pontuações fatoriais representam dimensões cognitivas latentes (não observáveis diretamente), isto é, capacidades não observáveis diretamente, cujos valores são estimados através da combinação dos resultados em múltiplos testes [Salkind, 2007]. Esta transformação permite que os resultados sejam facilmente interpretados e comparados com a população normativa de referência.

Após a geração dos resultados de medida, o *software* organiza-os em seis domínios cognitivos. Adicionalmente, o sistema calcula 22 pontuações de discrepância, juntamente com testes de significância e análises de frequência, que permitem avaliar os pontos fortes e fracos do perfil cognitivo do indivíduo e fornecem uma compreensão aprofundada das suas capacidades [Schretlen et al., 2025].

O CNNS utiliza normas avançadas, baseadas na regressão, que permitem aos profissionais ajustar os resultados dos testes de um indivíduo. O ajuste é feito tendo em conta as características do indivíduo, como idade, género, educação, descendência e uma estimativa do Quociente de Inteligência (QI) inicial. A personalização da comparação para cada teste aumenta a precisão do diagnóstico. Como as normas para todos os testes foram recolhidas da mesma forma, garante-se que, ao comparar diferentes capacidades cognitivas, as discrepâncias encontradas representam os verdadeiros pontos fortes e fracos do indivíduo, e não meras variações na medição [Schretlen et al., 2025].

Algumas das suas vantagens são [Schretlen et al., 2025]:

- Permite efetuar comparações normativas de forma mais precisa, através da filtragem da população de referência com base nas características do indivíduo;
- Inclui um conjunto alargado de testes neuropsicológicos, como o Trail Making Test, o Brief Test of Attention e o Modified Wisconsin Card Sorting Test, amplamente utilizados, permitindo uma avaliação cognitiva mais abrangente;
- Automatiza o processo do cálculo dos `T-scores` e das pontuações fatoriais;
- Os resultados das classificações não se limitam a valores numéricos, são acompanhadas por uma descrição qualitativa, tornando-os mais fáceis de interpretar.

No entanto, possui algumas desvantagens/limitações [DiPasquale, 2024]:

- Apesar dos ajustes demográficos, a amostra original utilizada para criar as normas pode não ser verdadeiramente representativa de todas as populações às quais os testes serão aplicados;
- Inclusão da etnia como fator de ajuste, uma vez que a etnia é uma construção social. O seu uso pode perpetuar enviesamentos e não reflete a diversidade interna das etnias;
- Não permite a inserção de novos testes;
- Falta de dados normativos padronizados para diversas populações internacionais.

Neuropsychological assessment platform (NPAP)

É um sistema computorizado desenvolvido como uma ferramenta profissional para a administração de testes psicológicos e neuropsicológicos, bem como para a recolha, armazenamento e disponibilização de dados de diversas populações/estudos a clínicos e investigadores [Chervinsky, 2007].

Recorrendo a *hardware* e *software* de multimédia, a plataforma administra testes padronizados da função cerebral, recolhe dados (audiovisuais e de resposta grafo motora) e armazena dados clínicos dos indivíduos avaliados, bem como dados normativos utilizados para a comparação. Esta informação é armazenada numa base de dados centralizada, acessível remotamente através da Internet para análise e comparação de resultados neurológicos [Chervinsky, 2007].

O sistema integra módulos de teste locais que administram avaliações aos sujeitos com mínima intervenção do examinador. Estes módulos apresentam testes de forma audiovisual e recolhem respostas audiovisuais, gráficas e táteis. Os dados dos indivíduos são depois enviados para uma base de dados centralizada, possibilitando o acesso remoto para efetuar comparações clínicas individuais (diagnósticos) ou para análises de grupo destinadas à investigação [Chervinsky, 2007].

Esta plataforma foi ainda desenhada para permitir a adição de novos módulos/aplicações destinadas à administração de testes neuropsicológicos e a atualização dos atuais [Chervinsky, 2007].

Algumas das suas vantagens são [Wärn et al., 2025]:

- Aumenta a acessibilidade, pois facilita a realização de testes neuropsicológicos, uma vez que são efetuados online, sem a necessidade de deslocação a um centro médico;
- Permite a realização de vários testes de forma computorizada, tornando o processo de administração mais eficiente e rápido;
- Permite armazenar dados de testes realizados e compara-los com padrões estabelecidos;
- Aumenta a precisão ao reduzir os erros comuns dos testes em papel.

No entanto, possui algumas desvantagens/limitações [Sahoo and Grover, 2022]:

- Dados recolhidos através de dispositivos ligados à Internet, especialmente os pessoais, suscita questões significativas de privacidade;
- Não possui dados normativos suficientes para certas populações;
- A administração remota ou diferenças no *hardware* dos computadores pode levar a desvios das condições de teste padronizadas originais, e desta forma reduzir a precisão dos diagnósticos;

- Alguns testes podem ser culturalmente específicos e inadequados para indivíduos de diferentes origens culturais ou grupos linguísticos.

2.2.3 Limitações

Tradicionalmente, os resultados são comparados com os dados normativos publicados nos manuais dos testes neuropsicológicos ou em estudos. No entanto, existem várias limitações com esta abordagem. Os dados podem estar desatualizados, levando a uma representação incorreta dos indivíduos atuais. Muitos dos testes publicados não possuem normas para certos grupos, tendo de recorrer a normas de outras populações ou a normas que os próprios clínicos recolheram. Os resultados normativos dos manuais são frequentemente corrigidos apenas para a idade, mas não para outras variáveis demográficas, como nível de educação ou género [de Vent et al., 2016].

Embora as abordagens mencionadas permitam análises concretas, o processo de comparação normativo possui várias limitações, como já foi referido. Algumas destas limitações advêm da forma como os dados são recolhidos, das abordagens/métodos selecionados, dos testes executados, do conjunto normativo ou das ferramentas selecionadas. As limitações da comparação normativa são variadas e correlacionadas. Do ponto de vista dos dados, destacam-se a reduzida dimensão dos conjuntos normativos e a ausência de normas para certas populações. Relativamente aos testes, muitos não possuem versões alternativas com dificuldade equivalente à original para repetir a avaliação e nem sempre refletem o real funcionamento do indivíduo no quotidiano, sendo ainda insuficientes para avaliar condições mais específicas, necessitando de um maior número de testes. Muitos deles enfrentam várias dificuldades na adaptação linguística e cultural e problemas de confiabilidade entre avaliadores. Alguns não representam adequadamente as exigências cognitivas atuais ou populações diversas e o desempenho nos testes pode ainda ser condicionado por fatores individuais e contextuais, nomeadamente o estado psicológico e a motivação do indivíduo durante a avaliação, bem como por características sociodemográficas como a idade, o nível de educação e o contexto cultural [Huizenga et al., 2007; Howieson, 2019; delCacho Tena et al., 2023; Sbordone, 2001; Shadowfourati, 2023].

Capítulo 3

Trabalho desenvolvido

O sistema apresentado neste trabalho é uma aplicação Web concebida para facilitar a recolha, armazenamento e análise de dados de testes neuropsicológicos de diferentes estudos. O público-alvo da aplicação são os investigadores e clínicos que necessitam de uma solução escalável para gerir dados de resultados de testes e realizar comparações normativas entre várias populações. A aplicação desenvolvida procura responder aos requisitos definidos pelo cliente para o armazenamento, gestão e análise de dados neuropsicológicos. As funcionalidades implementadas e as decisões tomadas durante o desenvolvimento serão descritas nas secções seguintes.

A arquitetura da base de dados é um componente crítico desta solução proposta. Foi desenvolvida com o objetivo de lidar com a alta escalabilidade dos dados (testes e resultados) e a sua variabilidade. Para tal, optou-se pela utilização de um modelo EAV. Esta plataforma tem como propósito abordar as limitações da comparação normativa de testes neuropsicológicos, identificadas na Secção 2.2.3, com o objetivo de as resolver.

3.1 Seleção das Tecnologias

Nesta secção é apresentada a seleção das tecnologias utilizadas no desenvolvimento da aplicação. Será abordada a escolha do sistema de base de dados, da linguagem de programação e da *framework* Web. São discutidas as principais opções consideradas e analisadas para cada componente, tendo em consideração o desempenho, flexibilidade e adequação aos requisitos do projeto. Será justificada cada escolha com base na compatibilidade com os objetivos/requisitos da aplicação e na capacidade de suportar a escalabilidade, variabilidade e complexidade dos dados neuropsicológicos.

3.1.1 Base de dados

Como referido na Secção 2.1, a avaliação neuropsicológica de um indivíduo envolve uma ampla variedade de testes, cada um concebido com objetivos, estruturas e métodos de execução distintos. Estes testes diferem significativamente não apenas no número de parâmetros que utilizam, mas também no tipo de dados que medem, cobrindo domínios como memória, atenção, funções executivas e processamento emocional. Alguns testes podem basear-se em apenas alguns indicadores-chave, enquanto outros geram grandes volumes de dados em várias variáveis. Além disso, os estudos frequentemente combinam vários testes para capturar um perfil abrangente da condição cognitiva de um indivíduo. Devido a esta alta variabilidade, rápido crescimento, natureza dinâmica e à necessidade de permitir a adição de novos testes no futuro, uma estrutura estática da base de dados é insuficiente. Em vez disso, um *design* da base de dados flexível e escalável é essencial para acomodar os diversos conjuntos de parâmetros e requisitos em evolução dos dados neuropsicológicos [Wahyuningrum et al., 2019]. Existem muitas abordagens para lidar com estes requisitos. Assim, foi decidido analisar duas soluções de bases de dados que foram consideradas mais adequadas, uma relacional e outra NoSQL (não relacional), com o objetivo de selecionar a que melhor se adequasse às necessidades do projeto.

Base de dados relacional (SQL)

Uma base de dados relacional é um tipo de base de dados que organiza os dados em tabelas, cada uma com linhas e colunas, estabelecendo relações entre as tabelas através de chaves. Cada linha representa um registo único e cada coluna um atributo [OVHcloud, 2025]. As tabelas podem ser unidas através de uma chave primária e de uma chave estrangeira. Estas chaves primárias estabelecem diversas relações entre as tabelas, facilitando a ligação e a organização estruturada dos dados [IBM, 2025c; OVHcloud, 2025]. Para manter a precisão e a acessibilidade, as bases de dados aplicam regras como a prevenção de linhas duplicadas [Oracle, 2021a].

SQL é a linguagem utilizada para interagir com estas bases de dados, possibilitando aos utilizadores criar, visualizar, atualizar e eliminar dados. Adicionalmente, esta linguagem permite a criação de consultas para extrair informações, manipular tabelas, efetuar cálculos e implementar regras de integridade e segurança [OVHcloud, 2025].

Este tipo de base de dados tem várias vantagens, incluindo [IBM, 2025c; Google-Cloud, 2025b; Hassan, 2021; CodeNet, 2025]:

- Possibilidade de combinar dados de múltiplas tabelas, através de operações de junção (`joins`) numa única consulta;
- Mais fácil de utilizar quando comparado com outros tipos de bases de dados (por exemplo, base de dados não relacionais), devido a uma vasta comunidade;

- Reduz a redundância por meio da normalização e de procedimentos armazenados;
- Cópia de segurança fácil em caso de desastre;
- Mais fácil de adicionar, alterar, atualizar ou eliminar tabelas, sem ser necessário alterar a estrutura;
- Maior escalabilidade e flexibilidade, permitindo efetuar consultas complexas;
- São ideais para aplicações que exigem precisão;
- Garantem a integridade e segurança dos dados;
- A segurança baseada em perfis garante que o acesso aos dados seja limitado a utilizadores específicos.

Apesar das vantagens das bases de dados relacionais, também apresentam algumas desvantagens, como [Milvus, 2025; Hassan, 2021; Abdellaoui et al., 2025; CodeNet, 2025]:

- São projetadas para a escalabilidade vertical (servidores mais potentes) em vez da escalabilidade horizontal (mais servidores), o que limita a sua eficiência em grandes conjuntos de dados;
- Inflexibilidade devido ao seu esquema rígido;
- A escalabilidade das bases de dados relacionais pode ser cara, principalmente devido à necessidade de *hardware* robusto especializado;
- A sua eficiência é reduzida para dados não estruturados ou altamente distribuídos.

Base de dados NoSQL

As bases de dados NoSQL são bases de dados não relacionais criadas para grandes volumes de dados, diversificados e não estruturados. Os dados são armazenados numa única estrutura não tabular sem um esquema [IBM, 2026] (planta da base de dados, que descreve como os dados se podem relacionar) fixo, oferecendo escalabilidade rápida, alto desempenho e alta disponibilidade através do armazenamento distribuído e replicado em vários servidores. Melhora a disponibilidade e a fiabilidade dos dados, uma vez que a base de dados permanece operacional mesmo em casos em que algumas fontes de dados estejam offline [IBM, 2025b]. As bases de dados NoSQL são utilizadas para aplicações Web em tempo real e de *big data* devido às suas principais vantagens: alta escalabilidade e alta disponibilidade [Oracle, 2021b].

Suportam diversos modelos de dados, como documentos, chave-valor e modelos orientados por colunas, cada um com características únicas. Bases de dados de documentos são utilizadas para armazenar e consultar dados semi-estruturados. Os dados são normalmente representados em documentos JSON, facilitando a criação e atualização de aplicações, sem a necessidade de recorrer a um esquema fixo. Bases de dados de chave-valor são o tipo mais simples de base de dados NoSQL, onde os dados são armazenados numa estrutura em que a chave única é ligada a um valor (string, número, booleano ou um objeto). A chave é utilizada para armazenar ou consultar o valor associado. Bases de dados orientadas por colunas armazenam os dados em linhas organizadas como um conjunto de colunas. Elas são ideais para consultas e agregações de dados rápidas. Bases de dados de gráficos organizam os dados como nós num gráfico, focando-se nas suas relações. Elas representam melhor as relações entre os dados e permitem um armazenamento e mecanismos de pesquisa mais simplificados. As bases de dados em memória caracterizam-se por armazenar os dados diretamente na memória principal, o que possibilita latências ultrabaixas em tempo real [GoogleCloud, 2025a].

Este tipo de base de dados tem várias vantagens, tais como [IBM, 2025b; Oracle, 2021b]:

- Mais flexível, uma vez que os dados podem ser armazenados de forma mais livre, sem esquemas rígidos, permitindo lidar facilmente com qualquer formato de dados num único armazenamento de dados;
- Maior escalabilidade, permitindo acomodar o aumento do tráfego através de arquiteturas distribuídas e da escalabilidade horizontal, sem tempo de inatividade;
- Alto desempenho, garantindo tempos de resposta rápidos quando o tráfego e os volumes de dados aumentam;
- Melhor disponibilidade, uma vez que replicam automaticamente os dados em vários servidores, centros de dados ou recursos na nuvem. Esta funcionalidade também reduz a carga da gestão da base de dados e minimiza a latência para os utilizadores;
- São projetadas para armazenar dados extremamente grandes com custos mínimos.

As desvantagens das bases de dados NoSQL são as seguintes [Abdellaoui et al., 2025; Hassan, 2021]:

- Não possui uma linguagem de consulta universal como o SQL, o que significa que cada um tem a sua própria, o que complica o desenvolvimento, a manutenção e a integração, e exige que os programadores aprendam novas ferramentas;

- Frequentemente, prioriza a disponibilidade em detrimento da consistência imediata dos dados, ou seja, os dados podem não ser instantaneamente consistentes em todas as partes da base de dados, o que pode ser uma desvantagem para aplicações que exigem precisão rigorosa dos dados em tempo real;
- A segurança nestas bases de dados é inferior à das bases de dados relacionais. A ausência de encriptação nos ficheiros de dados e a fragilidade do sistema de autenticação são as principais preocupações;
- Lidar com transações complexas pode ser mais desafiante.

Decisão

A diferença fundamental entre bases de dados relacionais e não relacionais (NoSQL) reside nos métodos de armazenamento e organização de dados. Como mencionado anteriormente, as bases de dados relacionais utilizam uma abordagem estruturada e baseada em esquemas. Em contrapartida, as bases de dados NoSQL oferecem flexibilidade ao armazenar dados como unidades individuais sem um esquema predefinido. Isto permite-lhes lidar com dados que mudam frequentemente e processar diversos tipos de dados, definindo o modelo de dados conforme necessário [GoogleCloud, 2025b; Oracle, 2021b]. Em última análise, as bases de dados não relacionais visam resolver os desafios de flexibilidade e escalabilidade que surgem quando os modelos relacionais se deparam com formatos de dados não estruturados [IBM, 2025c].

A escolha entre bases de dados relacionais e NoSQL depende da natureza e da escala dos dados. As bases de dados relacionais são adequadas para aplicações que geram quantidades fixas e gerenciáveis de dados. No entanto, as bases de dados NoSQL são preferíveis para aplicações que lidam com dados muito grandes e complexos que exigem alterações dinâmicas no esquema [Wahyuningrum et al., 2019].

Após analisar ambas as abordagens, tendo em conta os objetivos da aplicação e a complexidade, variabilidade e estrutura hierárquica dos dados dos testes neuropsicológicos, decidi utilizar uma base de dados relacional. As razões que me levaram a esta decisão foram:

- Os dados neuropsicológicos envolvem várias entidades relacionadas (utente, investigador, estudo, sessão). Desta forma, uma base de dados relacional assegura a coerência e integridade dos dados, através de chaves primárias e estrangeiras, evitando inconsistências.
- Possui uma estrutura lógica organizada e de fácil leitura e compreensão;
- Utiliza SQL, permitindo a criação de consultas complexas para filtrar, agregar e combinar dados de múltiplas tabelas de forma eficiente. Esta capacidade é fundamental para a pesquisa de informações na aplicação e para a análise normativa, onde é necessário selecionar grupos normativos para avaliar o in-

divíduo alvo, exigindo um processamento eficiente em grandes conjuntos de dados de consultas complexas;

- Possui regras de integridade de dados, que garantem a unicidade das chaves primárias, a validação de chaves estrangeiras que referenciam dados existentes, a restrição de valores e tipos de dados por coluna e a prevenção de linhas duplicadas. Este controlo é fundamental para a manutenção e para assegurar a qualidade dos dados e a fiabilidade num sistema com atualizações frequentes.
- Possui uma maior segurança e controlo de acessos;
- Compatibilidade com os requisitos do cliente.

No entanto, existem vários sistemas de gestão de bases de dados relacionais, sendo que decidi optar pelo PostgreSQL [The PostgreSQL Global Development Group, 2026]. Esta escolha deve-se a um conjunto de fatores. Em primeiro lugar, suporta uma grande variedade de tipos de dados, possibilitando uma estrutura mais flexível e adequada à natureza heterogénea dos dados neuropsicológicos. Apresenta também alto desempenho através de indexação avançada, especialmente com grandes volumes de dados e consultas complexas, comuns na análise normativa. Além disso, garante a integridade dos dados e dispõe de recursos de segurança avançada, incluindo controlo de acesso baseado em papéis/cargos e encriptação. Oferece ainda alta robustez e escalabilidade através da replicabilidade, lidando melhor com consultas complexas. Por fim, é *open source* e pode ser facilmente ampliada com funções personalizadas, operadores e agregados. [Quest, 2025].

Relativamente à modelação da base de dados, optou-se por usar um modelo EAV, já mencionado, implementado sobre PostgreSQL. Este modelo foi utilizado em tabelas onde a estrutura dos dados é heterogénea e pode sofrer alterações ao longo do tempo. Desta forma, permite uma maior flexibilidade na representação dos dados, centralizando-os numa estrutura mais genérica, garantindo a integridade e a eficiência do armazenamento.

3.1.2 Linguagem e Web Framework

Relativamente à linguagem de programação, optou-se por utilizar o Python, uma vez que: suporta um vasto número de bibliotecas e módulos de terceiros, incluindo ferramentas cruciais para a análise comparativa; é bastante utilizada no desenvolvimento Web, devido à sua eficiência e adaptabilidade; é conhecida pela sua simplicidade e legibilidade, pois oferece flexibilidade na codificação, uma vez que suporta a programação orientada a objetos e procedimental e não exige a declaração explícita dos tipos de dados, mantendo-se confiável [Krishna and Pritibala, 2025].

Depois de definir a linguagem de programação, o passo seguinte consistiu em escolher uma Web *framework* que se integrasse e permitisse cumprir os requisitos da aplicação. Para isso, foram consideradas soluções como o Flask [Pallets Projects, 2026] e o

Django [Django Software Foundation, 2026]. O Flask é um *micro-framework* para o desenvolvimento Web, escrito em Python. *Micro-Frameworks* são *frameworks* mais simples e modulares, com uma estrutura inicial muito menos complexa do que as *frameworks* convencionais. Apresenta um núcleo simples e expansível, possuindo apenas os recursos necessários para a execução da aplicação. No entanto, permite adicionar novos pacotes para adicionar funcionalidades, conforme necessário [Ghimire, 2020; Aggarwal, 2014; Singh et al., 2019; Patryk, 2024; FUIOR, 2021].

O Flask destina-se principalmente a pequenas aplicações, com requisitos mais simples, uma vez que possui uma arquitetura simples. Distingue-se pela sua simplicidade e rapidez no desenvolvimento de um projeto, pois foca-se apenas em desenvolver as funcionalidades principais, sem se preocupar em realizar configurações desnecessárias [Ghimire, 2020; Aggarwal, 2014; Singh et al., 2019; Patryk, 2024; FUIOR, 2021]. No entanto, possui algumas desvantagens [FUIOR, 2021; Patryk, 2024]:

- Exige bibliotecas externas, uma vez que carece de funcionalidades integradas;
- Menos adequado para aplicações de grande escala;
- Custos de manutenção elevados para sistemas mais complexos;
- Manutenção complexa para aplicações de maior dimensão;
- Ausência de um *Object-Relational Mapper* (ORM).

O Django é uma Python Web *framework* de alto nível, projetado para o rápido desenvolvimento de sites seguros e de fácil manutenção. Uma vez que é gratuito e de código aberto (código-fonte disponibilizado publicamente), conta com uma comunidade ativa e excelente documentação, simplificando o processo de desenvolvimento. Inclui dezenas de ferramentas extras para lidar com tarefas comuns (autenticação dos utilizadores e administração dos conteúdos), o que aumenta a versatilidade, permitindo o desenvolvimento de quase todo o tipo de sites. Além disso, oferece alta segurança, evitando erros comuns de programação, e é extremamente escalável, adaptando-se com rapidez e flexibilidade às necessidades da aplicação, permitindo a adição de novas funcionalidades conforme necessário [MDN, 2025; Django, 2025]. Destaca-se assim pela sua simplicidade no desenvolvimento de soluções escaláveis de forma rápida, proporcionando ferramentas robustas e eficientes.

Após comparar estas duas *frameworks*, foi escolhido o Django para o desenvolvimento da aplicação. Esta decisão foi baseada nas seguintes características: inclui um *Object-Relational Mapper* (ORM) que permite interagir com bases de dados relacionais, como o PostgreSQL, de forma simples; rapidez no desenvolvimento; alta segurança através de vários mecanismos; permite uma fácil administração/gestão dos dados e do controlo de acesso dos utilizadores; compatível com inúmeras bibliotecas; extremamente escalável e flexível. Atende assim plenamente às necessidades do projeto e permite o desenvolvimento de uma aplicação Web robusta, escalável e cientificamente orientada [Katie McLaughlin, 2017].

3.2 Análise dos dados

O conjunto de dados reais utilizado para testar as funcionalidades da aplicação foi fornecido no âmbito do projeto Apples2Apples mencionado anteriormente. Estes dados foram recolhidos num estudo neuropsicológico na Roménia, onde foram utilizados seis testes neuropsicológicos diferentes (`Rey-Osterrieth Complex Figure Test`, `Trail Making Test`, `Controlled Word Association Test`, `Stroop Color--Word Interference Test`, `Brief Test of Attention` e `Mind in the Eyes Test`). Cada registo representa a aplicação dos seis testes a um determinado utente, incluindo dados demográficos (data de nascimento, data da realização do teste, género, anos de escolaridade, município, localização, grupo étnico, e mão dominante), informações contextuais (nome do investigador) e resultados quantitativos associados aos parâmetros de cada condição dos testes.

O conjunto de dados é composto por 420 utentes e os seis testes neuropsicológicos englobam múltiplos domínios cognitivos, como a memória (visual e de trabalho), a atenção, as funções executivas, a linguagem e a cognição social. Esta diversidade do conjunto foi determinante para testar a flexibilidade e robustez da aplicação.

3.2.1 Indivíduos

Como já referi anteriormente, o conjunto possui 420 instâncias, o que equivale a 420 indivíduos, cada um com a sua informação demográfica. O identificador do indivíduo, como o nome indica, corresponde às suas iniciais, que o irão identificar. A Tabela 3.1, apresenta os diferentes valores em cada variável demográfica.

Variável	**Representações dos valores**
Sex	Masculin, Feminin, masculin, F, M
County	Ialomia, Bucureti, Prahova, Ilfov, Constana, Cluj, Piteti, Bucuresti, Botoani, Constanta, Iasi, Bacau, Vaslui, Maramures, Botosani, missing, Suceava, braov, Iai, bucureti, Valcea, bucuresti, Rep. Moldova, Brila, Neam, Arge, Brasov, BUCURESTI, TULCEA, Buzau, Teleorman, Dambovia, bacau, Galati, Mures, Harghita
Urban/Rural	Urban, Rural, urban, rural, missing
DOB (Date of Birth)	Transformado em idade: 17–87 anos
DOE (Date of Examination)	NULL
Ethnic Group	Român, Roman, Rrom, Moldovean, roman, turc, rrom, missing
Dominant Hand	Dreapta, Ambidextru, missing, dreapta, dreptaci, stângaci, Stanga, Stangaci
Years of Education	4–20 anos

Tabela 3.1: Representações dos valores das variáveis demográficas.

Com base nela, é possível verificar que existem alguns problemas com a qualidade dos dados. Existem várias inconsistências na codificação dos resultados, ou seja, existe

uma alta variabilidade textual. Por exemplo, a variável `Sex` apresenta diferentes nomes/nomenclaturas para o mesmo valor, ou seja, a mesma categoria é apresentada de diferentes formas. Existem problemas de capitalização (*Masculin* e *masculin*) e abreviações misturadas com a forma extensa, podendo gerar categorias a mais e prejudicar a análise estatística.

A variável `County` (município) apresenta problemas de inconsistência geográfica, uma vez que possui variações ortográficas, valores capitalizados, diacríticos e valores em falta. Estas inconsistências indicam que os resultados foram inseridos manualmente sem um controlo fechado. A variável `Urban/Rural` também contém problemas de capitalização. Já a variável `Ethnic Group` para além da capitalização e da questão dos diacríticos, também dispõe de uma confusão entre identidade nacional ("Roman") e etnia ("Rrom"), e de uma falta de padronização semântica. Por fim, a variável `Dominant Hand` também detém problemas de capitalização, de diacríticos e de mistura entre categoria ("Dreapta", que significa direita) e descrição do sujeito ("dreptaci", que significa destro).

A Tabela 3.1 permite me concluir que existe uma elevada variabilidade, o que é comum neste tipo de dados. No entanto, esta variabilidade pode gerar ruído na análise, o que reduz a qualidade e confiabilidade dos resultados, caso o pré-processamento não seja o adequado. Este conjunto de dados apresenta uma variabilidade categórica substancial devido à inconsistência na utilização de maiúsculas, variação diacrítica, utilização de sinónimos e valores como *missing* para representar dados em falta. Assim é necessário realizar uma limpeza/pré-processamento nos dados após/durante o *upload* e antes de efetuar as comparações normativas, de forma a obter os melhores resultados com o menor ruído possível.

Variável Demográfica	**Valores Ausentes**
Sex	0
Age	0
Dominant Hand	4
Urban/Rural	13
County	12
Years of Education	0
Ethnic Group	2

Tabela 3.2: Número total de valores ausentes por variável demográfica.

A Tabela 3.2 permite concluir que o conjunto de dados possui um baixo número de valores ausentes nas variáveis demográficas, no entanto, é necessário tratar aqueles que existem. As variáveis *Sex*, *Age* e *Years of Education* não possuem nenhum valor ausente, o que indica que não haverá perda de amostras ao usá-las. As variáveis *Dominant Hand* e *Ethnic Group* apresentam um baixo número de valores ausentes, o que permite constatar que terão um impacto mínimo. Por fim, as variáveis *County* e *Urban/Rural* são as que apresentam um maior problema, uma vez que considerando o tamanho da amostra, estes valores podem diminuir o desempenho da análise. De forma geral, o conjunto de dados apresenta poucos valores ausentes nas variáveis

demográficas, no entanto, a sua existência sugere possíveis limitações nas análises de representatividade regional.

Uma forma de mitigar este problema seria remover as observações com valores em falta, embora isso reduza o tamanho da amostra e possa introduzir algum enviesamento. Outra alternativa, seria atribuir aos valores ausentes o valor da média ou da moda. Outra opção seria criar uma categoria Unknown, especialmente útil no caso de variáveis categóricas, permitindo manter a amostra na integra sem perder informação potencialmente relevante.

A Tabela 3.3 apresenta o número de indivíduos que realizou cada teste neuropsicológico e o número de indivíduos que não apresentam valores em falta em todos os parâmetros essenciais para a análise. Sabendo que o tamanho da amostra é 420 instâncias, esta tabela permite concluir o número de indivíduos que não realizaram cada teste.

Teste	**Nº (casos completos)**	**Nº total de indivíduos que o realizaram**
Rey-Osterreith Complex Figure Test (ROCF)	139	350
Trail Making Test (TMT)	389	412
Controlled Word Association Test (COWA)	420	420
Stroop Color-Word Interference Test (SCWT)	324	391
Brief Test of Attention (BTA)	408	408
Mind in the Eyes Test (RMET)	416	416

Tabela 3.3: Número de indivíduos com dados completos por teste neuropsicológico.

A análise exploratória e estrutural do conjunto de dados revelou uma elevada heterogeneidade entre testes, bem como a presença de numerosos valores em falta, tanto nas variáveis demográficas como nos resultados dos testes. Verificou-se também a necessidade de normalizar os resultados, de mapear as variáveis categóricas e de adotar uma estrutura flexível para o armazenamento. Desta forma, estes fatores influenciam a arquitetura da base de dados, as validações automáticas e a preparação dos dados antes da comparação.

3.2.2 Testes Neuropsicológicos

Cada teste permite avaliar domínios cognitivos específicos definidos por um objetivo prévio; por isso, cada um utiliza instrumentos próprios e mede parâmetros distintos. Estes parâmetros correspondem às variáveis medidas em cada tarefa/condição, onde cada um avalia um conjunto específico de capacidades cognitivas e possui interva-

los de referência próprios considerados normais. Esta secção permite realizar uma análise dos testes, dos seus intervalos e do número de valores ausentes[1] no conjunto de dados fornecido, com o objetivo de adquirir uma melhor compreensão sobre os mesmos. Para cada um dos testes, foi realizada uma análise exploratória dos resultados, com o objetivo de identificar valores em falta e padrões relevantes. Esta análise permitiu definir estratégias de validação durante o processo de *upload*, bem como identificar as transformações necessárias para a comparação normativa.

Rey-Osterreith Complex Figure Test (ROCF)

Este teste foca-se na precisão do desenho e no tempo de execução em diferentes tarefas, sendo elas *Direct Copy*, *Immediate Recall*, *Delayed Recall* e *Recognition*. Cada parâmetro de desempenho tem um intervalo fechado que varia entre 0 e 36 pontos, enquanto os temporais variam entre 0 e 300 segundos. Neste teste é realizada ainda uma análise detalhada dos erros, identificando o número de falsos negativos, falsos positivos, verdadeiros negativos e verdadeiros positivos, em que cada um possui um intervalo entre 0 e 12 pontos. Para além destas variáveis, ainda existem outras duas que medem o intervalo de tempo em minutos entre duas tarefas, isto é, entre *Direct Copy* e Immediate Recall, e entre *Immediate Recall* e *Delayed Recall*, com intervalos de 2-3 e 20-30, respetivamente.

Após uma análise detalhada, verificou-se que alguns parâmetros continham valores inválidos, incluindo mensagens incorretas ou resultados fora dos intervalos normativos conhecidos e alguns sem qualquer registo. A ausência total de valores para determinados indivíduos indica que estes não realizaram o teste ou que não obtiveram qualquer resultado para esse parâmetro, justificando a ausência de qualquer valor. Para este teste foi possível verificar que 70 indivíduos não o realizaram, uma vez que o conjunto de dados apresenta valores em falta em todos os parâmetros para estes indivíduos.

Esta observação é consistente com a natureza do ROCF, que é um teste complexo que envolve capacidades visuoconstrutivas e de memória, tornando-o mais exigente do ponto de vista cognitivo e motor. Assim, este número de indivíduos com valores em falta pode ser explicado pela dificuldade na realização, pela incapacidade dos indivíduos ou por limitações na recolha dos dados. A existência de valores inválidos sugere problemas na execução ou no registo dos resultados. Dadas estas características e a análise realizada, será necessário um tratamento mais rigoroso durante a limpeza dos dados e preparação para a análise normativa.

Trail Making Test (TMT)

Este teste foca-se na velocidade de processamento, no nível de atenção, na flexibilidade cognitiva, no reconhecimento de letras e números, na leitura visual, na função

[1] Do inglês, *missing values*.

executiva e na deteção de erros sequenciais e de perda de controlo. Possui apenas duas condições, *Trail Making A* e *Trail Making B*, que variam na sequência a desenvolver. Na tarefa A, o indivíduo conecta números de forma ascendente, na tarefa B é necessário conectar tanto números como letras de forma alternada num padrão ascendente. Cada condição possui estes três parâmetros: `completionTime`, `sequencingErrors` e `setLossErrors`. A variável temporal possui um intervalo considerado normal que varia entre 0-180 segundos para a condição *Trail Making A* e 0-300 segundos para a condição *Trail Making B*. Para as restantes variáveis, o intervalo normal varia entre 0 e 3 erros, excepto para o parâmetro *sequencing errors* da condição *Trail Making A*, que varia entre 0 e 2 erros.

Neste teste, foi verificado também a existência de mensagens inválidas e valores em falta, o que sugere que houve indivíduos que não o realizaram ou que por alguma razão o resultado em certos parâmetros foi inválido ou impossível de medir. A Tabela 3.4 permite ter uma ideia do número de valores ausentes presentes no conjunto de dados.

Parâmetro	**Valores Ausentes**
Trail Making A: Completion Time (seconds)	8
Trail Making A: Sequencing Errors	8
Trail Making A: Near-miss Errors	8
Trail Making B: Execution Time (seconds)	30
Trail Making B: Sequencing Errors	30
Trail Making B: Near-miss Errors	30

Tabela 3.4: Valores ausentes do teste TMT

Com base na Tabela 3.4, foi possível verificar que 8 indivíduos não realizaram ou que não completaram corretamente a parte A do teste e 30 a parte B. Esta análise é coerente com a estrutura do teste, uma vez que a parte B é substancialmente mais complexa do que a parte A, exigindo maior flexibilidade cognitiva e capacidade de alternar entre sequências. O maior número de resultados ausentes ou inválidos na parte B, sugere uma maior dificuldade de realização por parte dos indivíduos. É fundamental analisar as duas partes do teste separadamente e excluir valores inválidos para garantir a fiabilidade dos resultados.

Controlled Word Association Test (COWA)

Controlled Word Association Test (COWA) foca-se no funcionamento cognitivo e avalia essencialmente a fluência verbal fonética (letras F-A-S) e semântica (animais). É dividido em duas condições específicas, FAS e *Animal Test*. Este teste consiste em produzir o maior número de palavras diferentes iniciadas com letras específicas (F-A-S) e por verbalizar o maior número de nomes de animais em um minuto. Cada condição possui três parâmetros: *Correct responses* (número de palavras que seguem as regras impostas), *repetitions* (número de palavras repetidas) e *lost set errors* (número de palavras que não seguem as regras, por exemplo, palavras que

não começam pelas letras pré-estabelecidas). Todos os seus parâmetros possuem o mesmo intervalo de valores normais, 0-100 palavras.

Após analisar este teste, verificou-se que o conjunto de dados apresentava valores dentro do normal e sem quaisquer erros ou mensagens inválidas para todos os participantes, o que me permitiu avançar com segurança para as etapas seguintes de processamento e comparação normativa. Este comportamento pode ser explicado pela natureza do teste, que se baseia na produção verbal controlada, sendo mais simples de administrar e executar, tornando-o menos propenso a erros.

Stroop Color-Word Interference Test (SCWT)

Stroop Color-Word Interference Test (SCWT) permite medir a inibição de respostas rápidas. Mede o efeito Stroop que consiste em nomear a cor de uma palavra colorida (por exemplo, nomear a cor da palavra azul, redigida com tinta vermelha). Este processo é mais lento e está mais propenso a erros. Avalia a atenção, o funcionamento cognitivo e de execução e a memória de trabalho. SCWT está dividido em três condições: *Word*, *Color* e *Word-Color*. A tarefa *Word* consiste em ler o nome das cores redigidas a preto, a tarefa *Color* requer a identificação da cor de manchas e a tarefa *Word-Color* serve para medir o efeito Stroop já mencionado. Para cada uma das condições existem 2 parâmetros, respostas corretas e erros, sendo que apenas os erros possuem um intervalo de valores normativos definido, que varia entre 0 e 112. As respostas corretas não têm um limite definido, uma vez que não existe um limite fixo. Desta forma, é possível aferir que o mesmo depende do número de itens apresentados e que o desempenho corresponde à taxa de acertos relativa à velocidade ou pelo total de itens processados num tempo fixo.

Durante a análise, foi também verificada a existência de vários valores ausentes e de valores irregulares (mensagens inválidas), o que sugere que nem todos os indivíduos o realizaram e que houve alguma dificuldade na sua execução. A Tabela 3.5 apresenta o número de valores ausentes presentes no conjunto de dados.

Parâmetro	**Valores Ausentes**
Stroop Word: Correct Responses	29
Stroop Word: Errors	96
Stroop Color: Correct Responses	29
Stroop Color: Errors	96
Stroop Color-Word: Correct Responses	29
Stroop Color-Word: Errors	30

Tabela 3.5: Valores ausentes do teste SCWT

A Tabela 3.5 permite verificar que 29 indivíduos obtiveram valores irregulares na medição do número de respostas corretas, 96 no número de erros para as condições individuais e 30 no número de erros para a condição composta do teste. Este comportamento pode ser explicado pela natureza do teste, uma vez que avalia a atenção

seletiva e a capacidade de inibição cognitiva. A ocorrência de erros neste teste é esperada, especialmente em condições de maior interferência, justificando esta elevada variabilidade e quantidade. No entanto, os valores irregulares e inválidos, sugerem problemas de registo ou execução.

Brief Test of Attention (BTA)

Brief Test of Attention (BTA) avalia a atenção, a memória de trabalho e a função cognitiva. Tem como principal objetivo a avaliação da atenção de indivíduos com limitações motoras, isto é, com dificuldade em realizar os testes que requerem rastreamento visual ou destreza manual. É constituído por duas condições: números e letras. Na condição dos números, o participante ouve um conjunto de 10 listas reproduzidas, formadas por letras e números, e deve contar apenas os números apresentados. Já na condição das letras, ocorre o inverso, o participante ouve igualmente 10 listas, que aumentam progressivamente em comprimento, e deve contar apenas as letras em cada lista, ignorando os números. Cada condição possui apenas um parâmetro (respostas corretas), que varia entre 0 e 20.

Este teste também possui várias instâncias com valores em falta, indicando a sua não realização por parte de vários indivíduos. A Tabela 3.6 apresenta o número de valores ausentes presentes no conjunto de dados.

Parâmetro	**Valores Ausentes**
BTA Numbers: Correct Responses	12
BTA Letters: Correct Responses	12

Tabela 3.6: Valores ausentes do teste BTA

Com base na Tabela 3.6 é possível averiguar que existem 12 indivíduos que não realizaram este teste ou que obtiveram valores irregulares. Uma vez que a realização deste teste envolve a capacidade auditiva e o processamento de informação verbal, a ausência de resultados pode estar associada a dificuldades específicas neste domínio ou a limitações da realização do teste. A ausência de valores implica, portanto, a exclusão destes valores ou o seu tratamento. No entanto, o número reduzido de indivíduos com valores ausentes, sugere um baixo impacto na análise.

Mind in the eyes Test (RMET)

Mind in the eyes Test (RMET) é um teste que permite avaliar a teoria da mente e a cognição social, ou seja, permite reconhecer e compreender o estado mental/inteligência social através do olhar. Consiste em visualizar 36 imagens e, para cada, escolher a opção de entre quatro que melhor descreve o que a pessoa na imagem está a sentir ou pensar. Assim, apresenta apenas uma condição e um único parâmetro associado, número de respostas corretas, cujo valor pode variar entre 0 e 36.

Durante a análise, verifiquei que este teste apenas possuía quatro indivíduos que não obtiveram nenhum valor, ou seja, que não o realizaram. Este reduzido número pode ser explicado pela natureza simplista e rápida de realizar do teste. Esta menor complexidade contribui para um maior número de indivíduos que realizaram o teste corretamente. Este número reduzido de indivíduos com valores ausentes, permite a sua utilização direta em análises normativas, sem a necessidade de procedimentos complexos e extensos de limpeza.

Com estas análises foi possível entender melhor os dados com que estava a trabalhar e tomar as decisões mais informadas, tanto durante o *upload* como na comparação normativa.

3.3 Limpeza dos dados

Como foi mencionado na análise, os dados requerem alguma normalização antes de serem utilizados na comparação normativa. Desta forma, esta secção irá abordar algumas transformações necessárias para o pré-processamento do conjunto de dados, a serem aplicadas durante o processo de *upload* e comparação. Esta limpeza foi realizada antes do desenvolvimento da aplicação, com o objetivo de melhor entender os dados e explorar técnicas de normalização a aplicar automaticamente durante o processamento dos dados.

Em primeiro lugar, para lidar com as instâncias que continham valores ausentes nos parâmetros dos testes, para assegurar a consistência e o rigor estatístico e evitar enviesamentos complexos na comparação normativa, foram eliminadas estas instâncias na sua totalidade (incluindo todos os seus resultados e informações). No entanto, os valores ausentes serão permitidos nas variáveis demográficas.

Relativamente à existência de vários erros ortográficos nas variáveis demográficas (por exemplo, grupo étnico e município(*County*)), foi necessário corrigi-los. Para a correção dos municípios, foi utilizada a biblioteca *pycountry*[2] do Python. Esta biblioteca permite ter acesso a uma lista de cidades de cada país através da função *subdivisions*. Em conjunto com a biblioteca *rapidfuzz*, foi possível comparar e normalizar de forma eficiente todos os erros ortográficos nesta variável. Para as outras variáveis demográficas categóricas foi realizada uma normalização baseada na similaridade textual. Esta normalização consiste em percorrer todos os valores da variável e compara-los dois a dois utilizando a função `fuzz.ratio` da biblioteca *RapidFuzz*, para calcular a similaridade entre duas *strings*, retornando um valor entre 0 e 100. Antes desta normalização, os valores são convertidos para minúsculas através da função `lower()` e para o tipo *string* (`str()`), de forma a garantir a consistência. Caso os valores sejam semelhantes (similaridade calculada superior ou igual a 80), então são agrupados no mesmo *cluster*. Para cada grupo identificado, seleciona-se como forma padrão a variante mais frequente no conjunto de dados e cria-se um dicionário de mapeamento que liga todas os valores semelhantes à forma dominante.

[2] medium.com pycountry

Por fim, todos os elementos da variável são mapeados conforme o dicionário criado, substituindo automaticamente as variações pela versão padronizada.

Para além da correção dos erros ortográficos/diacríticos, também foram normalizados outros dados, como o mapeamento dos valores da variável género em apenas duas categorias: F e M. Foi também necessário converter a data de nascimento em idade, calculando a diferença em relação à data de execução do teste, de modo a obter a idade exata do indivíduo no momento da sua realização.

Após a limpeza, normalização, padronização, correção e formatação dos dados, o *dataset* utilizado para as comparações normativas ficou a possuir 269 instâncias completas. Esta abordagem foi realizada com base nos requisitos definidos pelo cliente, permitindo garantir a consistência nas comparações, embora reduza o número de observações disponíveis. No entanto, como os testes são independentes, análises futuras poderão utilizar amostras independentes para cada teste, aproveitando o maior número de observações. Por motivos visuais, optou-se por agrupar as variáveis numéricas, a fim de melhor compreender os dados. Os gráficos seguintes, Figura 3.1, Figura 3.2, Figura 3.3, Figura 3.5 e Figura 3.4, apresentam os resultados finais das variáveis demográficas após a limpeza. Nesta fase, não existem valores em falta, uma vez que foram tratados durante o pré-processamento. No entanto, os registos que contêm a designação 'missing' como valor próprio foram preservados nos resultados.

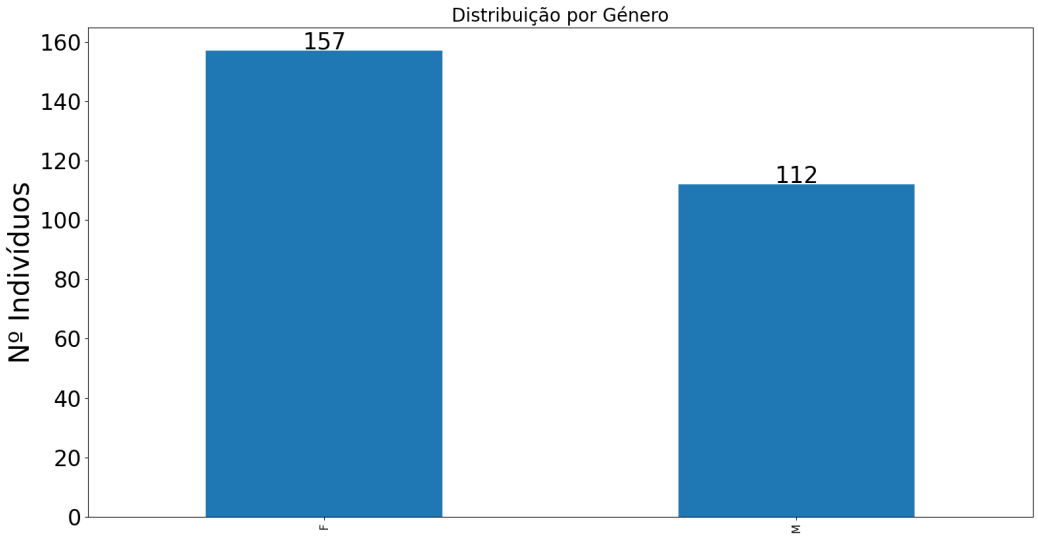

Figura 3.1: Distribuição da variável género após a limpeza e a normalização dos dados

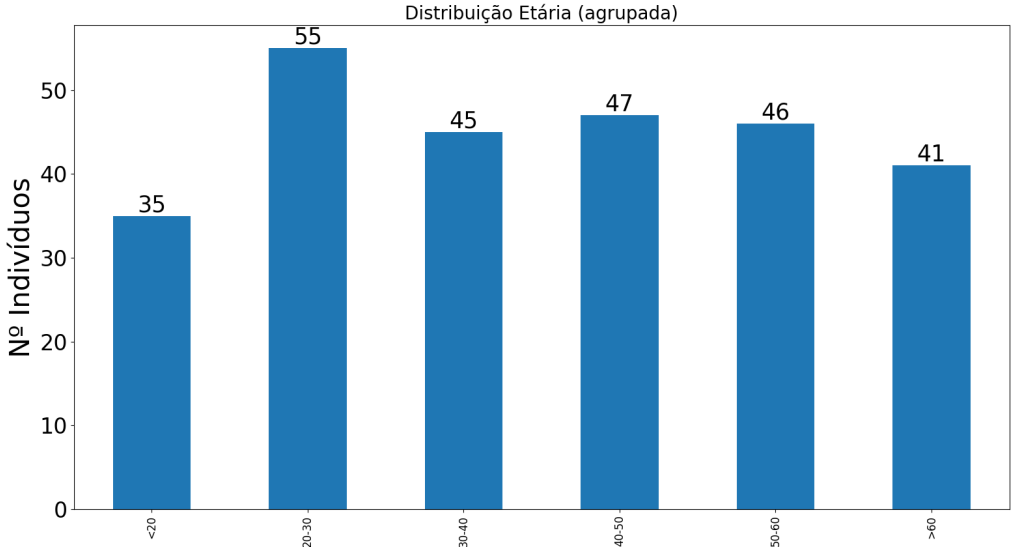

Figura 3.2: Distribuição da variável idade após a limpeza e a normalização dos dados

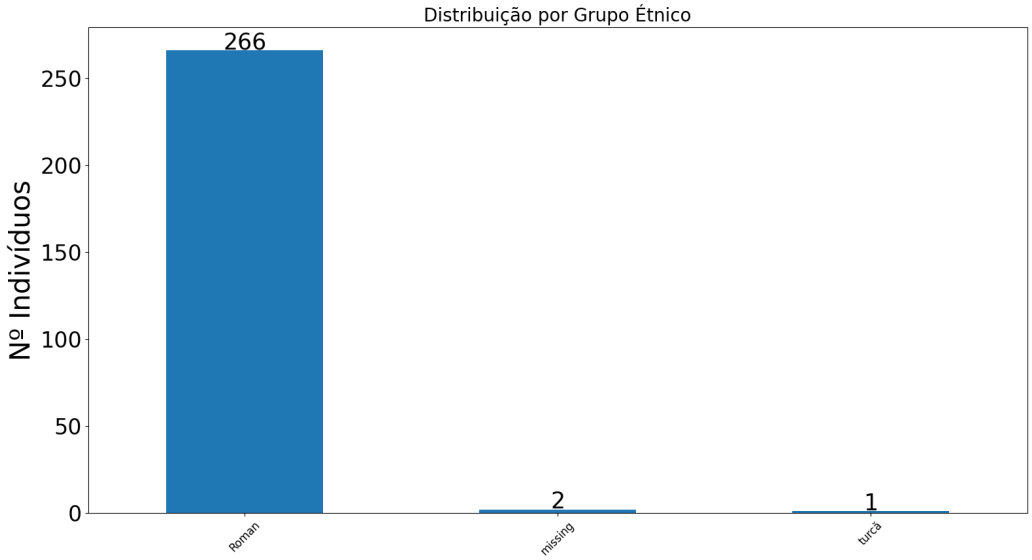

Figura 3.3: Distribuição da variável grupo étnico após a limpeza e a normalização dos dados

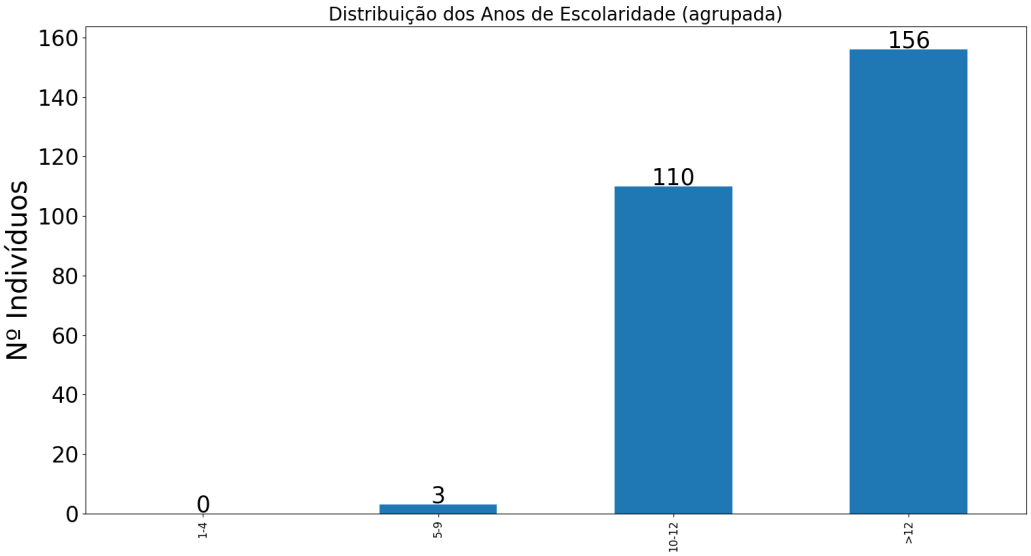

Figura 3.4: Distribuição da variável anos de educação após a limpeza e a normalização dos dados

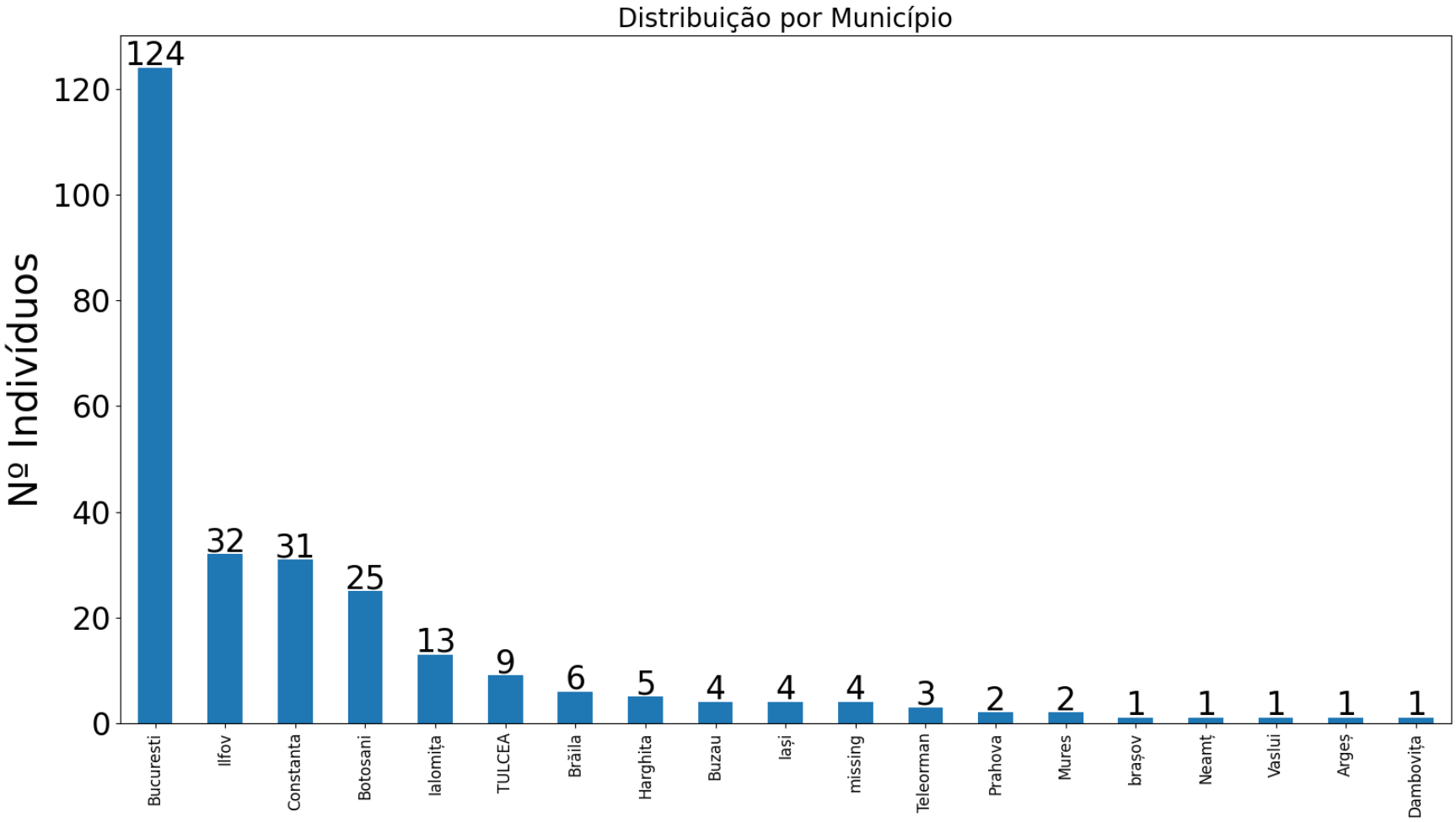

Figura 3.5: Distribuição da variável município após a limpeza e a normalização dos dados

3.4 Desenho da base de dados

A base de dados foi concebida para centralizar num único ambiente digital informação proveniente de diferentes estudos, testes e participantes, permitindo o armazenamento dos dados de forma estruturada. Dado o elevado número de testes neuropsicológicos existentes, a diversidade dos seus tipos, a variabilidade dos seus parâmetros e a sua complexidade inerente, foi necessário um *design* que pudesse acomodar estruturalmente a inserção contínua de diferentes testes. Os testes são estruturados com uma ou mais condições, que por sua vez contêm um ou mais parâmetros. Tomemos como exemplo o Rey-Osterrieth Complex Figure Test (ROCF), uma avaliação da memória visual e da função executiva: ele tem quatro condições (*direct copy*, *immediate recall*, *delayed recall*, e *recognition*). Dentro de cada condição, podem existir vários parâmetros, como o tempo para conclusão e a pontuação.

No âmbito da modelação de dados, selecionou-se o modelo Entidade-Atributo-Valor (EAV). Este modelo é particularmente útil em cenários onde os atributos e propriedades das entidades são frequentemente variáveis e escaláveis. É ideal para modelar situações com dados esparsos e personalizáveis, quando o número de atributos possíveis é elevado, mas apenas alguns se aplicam a cada entidade e quando é necessário adicionar consecutivamente novos atributos, uma vez que permite a sua inserção sem necessitar de alterar o esquema da base de dados [Brandt et al., 2002; Dinu and Nadkarni, 2007; Batra et al., 2018].

No contexto da modelação de dados, entidade é um objeto ou conceito real, identificável e em processo de descrição. Atributo é uma característica ou propriedade inerente à entidade, enquanto valor constitui um dado específico que um atributo assume para uma determinada entidade. O modelo armazena estes três componentes de forma correlacionada e interligada, permitindo representar os dados de forma versátil e ajustável [Brandt et al., 2002; Dinu and Nadkarni, 2007; Batra et al., 2018].

Como já foi mencionado anteriormente, este modelo permite armazenar dados não estruturados de forma eficiente, pois armazena os dados através de pares atributo-valor não vazios. Esta capacidade distingue-o dos modelos tradicionais de esquemas fixos, onde cada atributo possível ocupa um espaço de armazenamento, mesmo que vazio [Brandt et al., 2002; Dinu and Nadkarni, 2007; Batra et al., 2018].

O modelo EAV possui muitas vantagens, relativamente ao modelo relacional tradicional, algumas já referidas, como [Brandt et al., 2002; Dinu and Nadkarni, 2007; Batra et al., 2018]:

- Maior flexibilidade, uma vez que permite adicionar novos atributos sem a necessidade de alterar o esquema da base de dados;
- Mais eficiente, pois só armazena valores existentes, evitando espaço desperdiçado em colunas nulas;
- É ideal para situações em que o número de atributos é vasto e frequentemente

desconhecido à partida.

No entanto, também possui desvantagens, como [Brandt et al., 2002; Dinu and Nadkarni, 2007; Batra et al., 2018]:

- Pode levar a consultas mais complexas e a um desempenho mais lento, pois exige junções entre várias tabelas;
- Pode dificultar a imposição de restrições de integridade dos dados, visto que os atributos de uma entidade estão frequentemente dispersos por várias tabelas;
- A gestão pode tornar-se complexa e difícil com a adição de novos atributos.

Em termos de segurança, este modelo é tão robusto quanto o sistema de base de dados em que está implementado. Contudo, devido à sua inerente complexidade, pode requerer camadas adicionais de controlo de acesso, a fim de garantir a discrição no manuseamento dos dados [Batra et al., 2018].

Este modelo adapta-se perfeitamente a requisitos em constante evolução e a mudanças de esquema, tornando-o ideal para a área da neuropsicologia. Neste domínio, o elevado e crescente número de testes, a diversidade de atributos (alguns partilhados) e a recolha diária de um vasto volume de resultados para cada atributo, provenientes de múltiplos indivíduos, definem um ambiente onde a flexibilidade e a escalabilidade do modelo são essenciais.

Apesar de ter optado pelo modelo EAV num ambiente relacional, a utilização de uma base de dados NoSQL foi igualmente ponderada, como mencionado na secção 3.1.1. Estas abordagens apresentam várias diferenças, como: o modelo EAV armazena os componentes entidade, atributo e valor de forma correlacionada e interligada, já o NoSQL armazena os dados de forma não relacional (não estruturada). Quanto à flexibilidade, ambos os modelos permitem atributos dinâmicos, no entanto trata-se de uma característica nativa às bases de dados NoSQL. Quanto à leitura dos dados, ambos os modelos apresentam complexidade nas consultas, enquanto o EAV exige a execução de múltiplos *joins* para reconstruir a informação, o NoSQL requer o processamento de dados sem um esquema fixo, o que pode dificultar a estruturação dos resultados. Ambos são altamente escaláveis, sendo o NoSQL superior devido à sua escalabilidade horizontal. Por fim, o modelo EAV é mais confiável, uma vez que beneficia dos mecanismos de integridade e segurança inerentes ao modelo relacional.

Para acomodar estruturalmente a inserção contínua de diferentes testes, foi criado um modelo (representado na Fig. 3.6) em que cada teste 'Test' está associado a várias categorias 'TestCategory' e pode ter várias condições 'Condition', com várias categorias 'ConditionCategory', que estão ligadas a parâmetros mensuráveis 'Parameter'. As sessões 'Session' representam as avaliações dos participantes, e cada sessão pode envolver várias administrações de testes 'TestAdministration', cada uma ligada a testes específicos. Os resultados 'Result' armazenam os resultados dessas

administrações de testes, onde é referenciado tanto o teste administrado como o parâmetro medido, juntamente com valores como pontuações, indicadores de validade e mensagens descritivas.

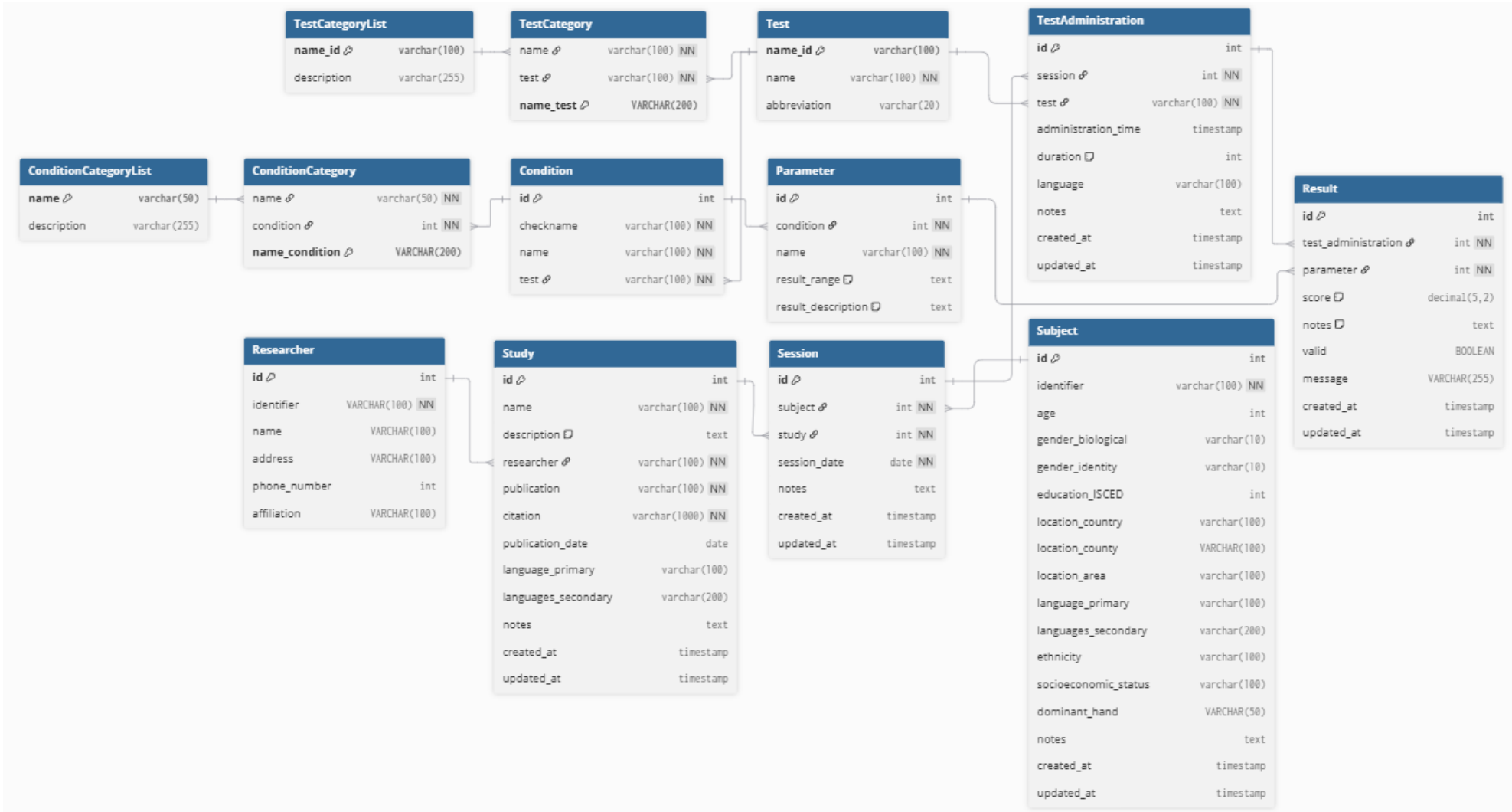

Figura 3.6: Desenho do esquema da base de dados

A base de dados também inclui outras tabelas para metadados, tais como 'Researcher', 'Subject' e 'Study'. Este modelo torna a base de dados mais flexível (lida com diferentes testes e parâmetros), escalável, consistente, fácil de compreender e implementar, e inclui mais validações automatizadas (graças aos diferentes tipos de relações e regras).

Na modelação da base de dados, o modelo EAV foi utilizado nas tabelas associadas à representação dos resultados dos testes neuropsicológicos ('Condition', 'Parameter' e 'Result').

Neste contexto, 'Test' (por exemplo, Trail Making Test) representa o instrumento de avaliação, ou seja, é uma entidade fixa que fornece o enquadramento conceptual da entidade avaliada. Cada 'Condition' representa a entidade avaliada (por exemplo, *Trail A* ou *Trail B*), cada parâmetro define o atributo (por exemplo, tempo de resposta, número de erros ou pontuação) e cada 'Result' armazena o valor para cada utente numa sessão onde o 'Test' foi administrado.

Além disso, esta combinação entre sistema relacional PostgreSQL e modelo EAV garante a integridade e consistência dos dados, mantendo as vantagens do sistema como a capacidade de efetuar consultas complexas. Esta abordagem híbrida é bastante adequada para armazenar dados neuropsicológicos, uma vez que equilibra a escalabilidade, flexibilidade, variabilidade e organização.

A gestão de utilizadores foi implementada utilizando o modelo de autenticação nativo do Django, que foi estendido com a tabela `UserProfile`. Enquanto o modelo `User` do Django trata das informações essenciais da autenticação (como *username/email* e palavra-passe encriptada), o modelo `UserProfile` estabelece uma relação um-para-um com o `User` para guardar dados complementares, como o nome completo e o *role* do utilizador (investigador ou clínico). Esta separação clara garante integridade da autenticação, segurança, flexibilidade e uma gestão mais eficiente dos diferentes tipos de utilizadores.

3.5 Aplicação Web

A aplicação Web desenvolvida tem como objetivo facilitar a recolha, gestão e análise de dados neuropsicológicos, proporcionando uma interface intuitiva e segura para investigadores e clínicos. Permite ainda a realização de comparações normativas.

Como já foi referido anteriormente, o desenvolvimento foi realizado em Django, um *framework* Python robusto, que assegura a segurança, modularidade e rapidez de desenvolvimento do sistema. Uma aplicação Web facilita a acessibilidade em diversas plataformas, a integração com bases de dados relacionais e a manutenção, garantindo soluções escaláveis e robustas.

Esta secção irá ainda apresentar algumas imagens das principais páginas da aplicação, com o objetivo de apresentar a interface, organização, funcionalidades e o fluxo

de utilização. Estas imagens permitem obter uma melhor compreensão da interação entre o utilizador e a aplicação. A interface gráfica foi concebida por uma equipa externa, tendo o desenvolvimento da aplicação seguido as suas diretrizes de *design*.

3.5.1 Arquitetura e Integrações

A arquitetura da aplicação separa o *frontend*, o *backend* e a base de dados, seguindo um *design* modular que suporta a escalabilidade. A arquitetura é modular, uma vez que segue o padrão Model-View-Template (MVT), do próprio Django. O modelo (Model) é responsável pela camada de dados, ou seja, na estrutura da base de dados (tabelas). A lógica/controlador (View) processa os pedidos HTTP, interage com a base de dados (Model) e envia informação ao utilizador. Os *templates* são responsáveis pela interface de utilizador (UI), ou seja, o aspeto da página. Esta separação torna a aplicação mais fácil de manter, escalar e reutilizar. Suporta um volume crescente de dados provenientes de vários estudos ao longo do tempo. A arquitetura é composta pelos seguintes módulos/componentes:

- Base de dados implementada em PostgreSQL, com um modelo EAV. Inclui *triggers* e *constraints* para assegurar integridade referencial e atualização automática de *timestamps*;
- *Backend*, desenvolvido em Django, que assegura a gestão de autenticação, permissões e perfis dos utilizador. Possui ainda funcionalidades como: importar dados; gerir utentes, investigadores e estudos; pesquisa avançada com múltiplos filtros e paginação de resultados; e um sistema de comparação normativa;
- *Frontend*, desenvolvido com *templates* Django, HTML, CSS e JavaScript, oferece uma interface simples e intuitiva. Esta inclui formulários, páginas de pesquisa e gráficos detalhados, com o objetivo de facilitar a compreensão dos resultados;
- Compatível e integrável com várias ferramentas e bibliotecas.

A integração entre os diferentes módulos/componentes da aplicação foi projetada para ser transparente ao utilizador e segura:

- **Integração da base de dados com a aplicação Web:** O ORM do Django garante que os dados provenientes dos formulários e *CSVs* sejam corretamente armazenados na base de dados;
- **Integração da aplicação com os utilizadores:** O sistema de autenticação garante que apenas utilizadores autorizados acedem às funcionalidades que os seus perfis permitem. Por exemplo, o *upload* de dados é restrito aos investigadores e administradores, as pesquisas filtradas podem ser efetuadas por qualquer utilizador e apenas os clínicos conseguem realizar comparações normativas;

- **Integração com modelos analíticos:** A infraestrutura permitiu a integração de métodos e modelos analíticos e encontra-se preparada para suportar a integração de novas abordagens, visando apoiar a comparação normativa;

- **Integração com ferramentas de visualização:** Permite a integração com bibliotecas como *Plotly* ou *Matplotlib* para visualização avançada de análises normativas.

3.5.2 Funcionalidades pedidas pelo cliente

A aplicação Web visa melhorar a acessibilidade, organização e interpretabilidade dos dados dos testes neuropsicológicos, ao mesmo tempo que apoia uma análise normativa flexível e culturalmente sensível. A fim de cumprir com os objetivos/funcionalidades solicitados pelo cliente para a aplicação, como permitir o *upload*, armazenamento, análise e pesquisa de dados, foi necessário desenvolver certas páginas Web.

Upload de dados

A página de *upload* dos dados permite aos utilizadores autenticados realizar o carregamento de dados de testes neuropsicológicos em formato CSV. O processo de carregamento inicia-se com o preenchimento de informações essenciais relativas ao estudo e ao investigador, incluindo o nome do estudo, citação, idioma, data de publicação, nome do investigador, identificador e afiliação.

Após esta etapa, os utilizadores selecionam os testes que serão carregados e o ficheiro CSV correspondente. Posteriormente ao *upload* do ficheiro, o sistema requer a normalização de determinados valores (como certas traduções) cruciais para a comparação normativa, garantindo a consistência, clareza, padronização e replicabilidade dos dados. Assim, o próximo passo consiste no mapeamento por parte do utilizador dos campos da base de dados com as colunas CSV, garantindo que cada coluna é corretamente associada ao respetivo campo na tabela. Este processo assegura a integridade, consistência, importação correta, compatibilidade dos dados, prevenção de potenciais erros e preserva o significado original dos dados.

Para facilitar esta tarefa, a aplicação realiza ainda, sempre que possível, um pré-mapeamento automático das colunas do ficheiro CSV para as variáveis existentes na base de dados, com base na semelhança entre os nomes. É também efetuado um pré-mapeamento das traduções, onde estas são associadas automaticamente aos termos em diferentes idiomas, com o propósito de reduzir a intervenção manual necessária por parte do utilizador. Este mapeamento permite automatizar o processo de importação e simplifica as validações.

Após o mapeamento, o sistema executa um conjunto de validações automáticas sobre os resultados dos testes neuropsicológicos para garantir a sua qualidade, consistência

e o alinhamento adequado com o modelo de dados do sistema. É verificada a integridade dos valores, ou seja, é verificado se os resultados de cada parâmetro de cada teste estão dentro dos intervalos de valores normais para esses parâmetros. Qualquer valor fora desse intervalo não é armazenado. É ainda efetuada uma validação do tipo de dados, ou seja, é validado se o resultado de um parâmetro é numérico. Se não for, o valor é armazenado como `NULL`, marcado como inválido e a mensagem de erro fornecida (pelo investigador) é registada.

Ainda durante o *upload* foi necessário padronizar o formato das datas, uma vez que o *dataset* continha múltiplas representações. Esta padronização garantiu um único formato, tanto na data de nascimento, como na data da examinação. As datas inválidas com combinações de formatos foram corrigidas manualmente, permitindo o cálculo preciso da idade e de outros atributos, caso seja necessário. O nível de escolaridade foi ainda calculado a partir dos anos de escolaridade. Foi ainda efetuada a normalização das variáveis demográficas, através da conversão dos valores para minúsculas, assegurando a consistência dos dados e a unificação das palavras que têm o mesmo significado (eliminação da sensibilidade ao maiúsculo/minúsculo).

Por fim, o sistema apresenta aos utilizadores as mensagens resultantes do processo de validação (erros/avisos) e a opção de prosseguir com o *upload*. Com base nestas mensagens, o utilizador pode optar por prosseguir com o *upload* ou cancela-lo.

Interface

Esta secção apresenta as interfaces das páginas que dizem respeito ao *upload* dos dados. As Figuras 3.7 e 3.8 ilustram a primeira página, onde são inseridos os dados relativos ao estudo, ao investigador e à seleção de testes que compõem a investigação. A Figura 3.9 ilustra a página onde o CSV é escolhido e carregado, e por fim, a Figura 3.10 ilustra a página do mapeamento dos campos da base de dados com as colunas do CSV.

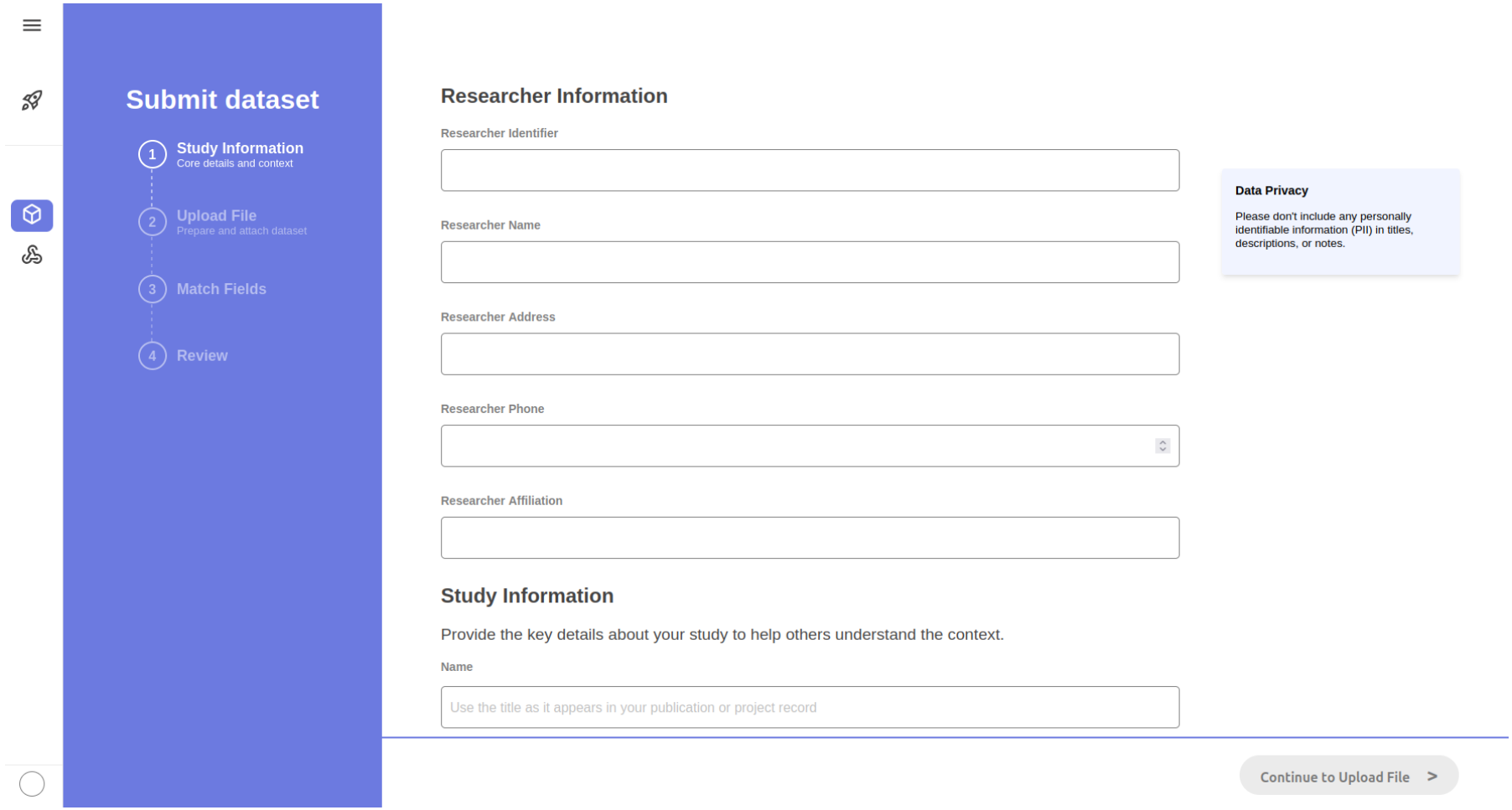

Figura 3.7: Interface da inserção dos dados do estudo e do investigador

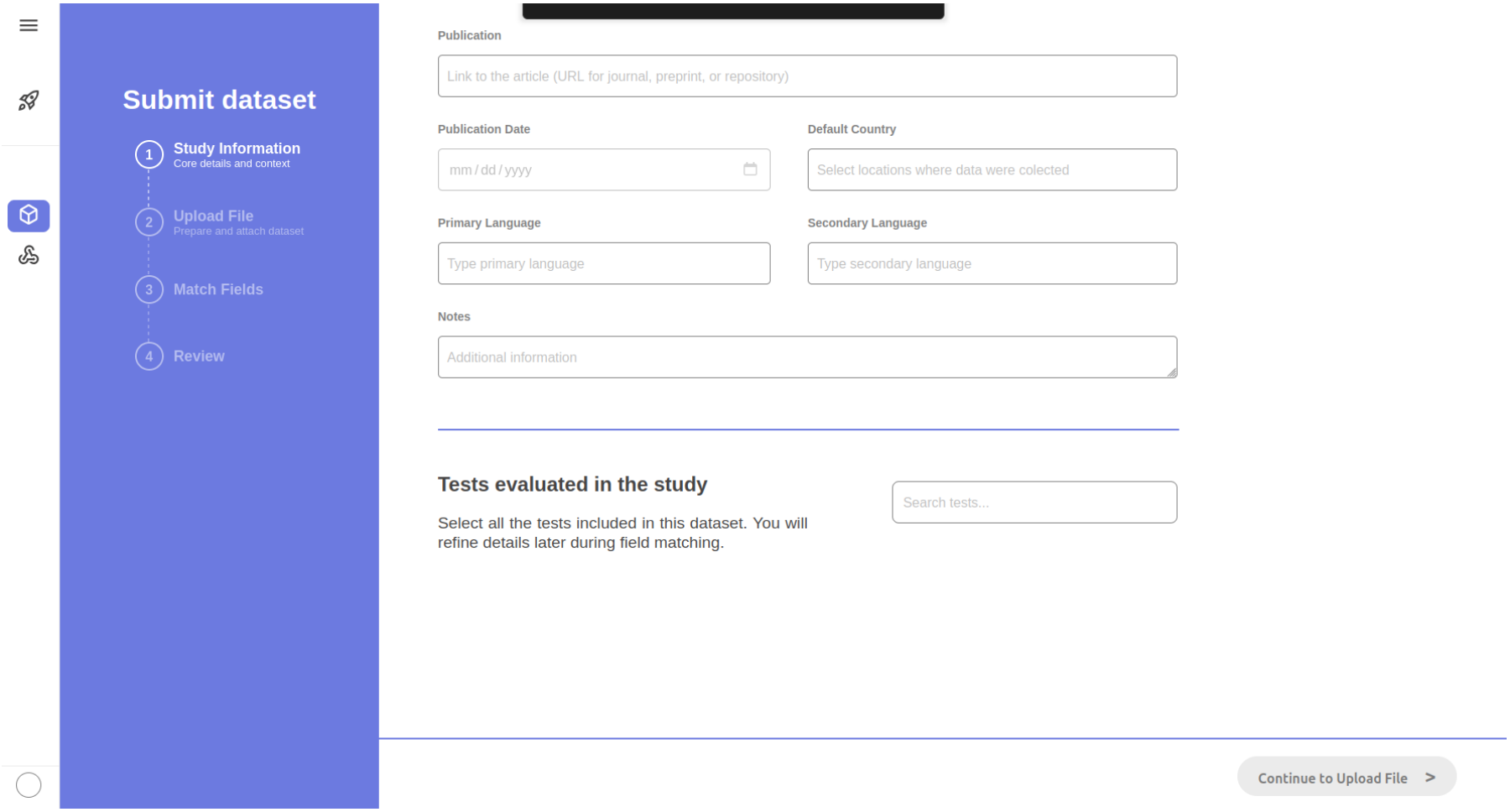

Figura 3.8: Interface da seleção dos testes utilizados no estudo

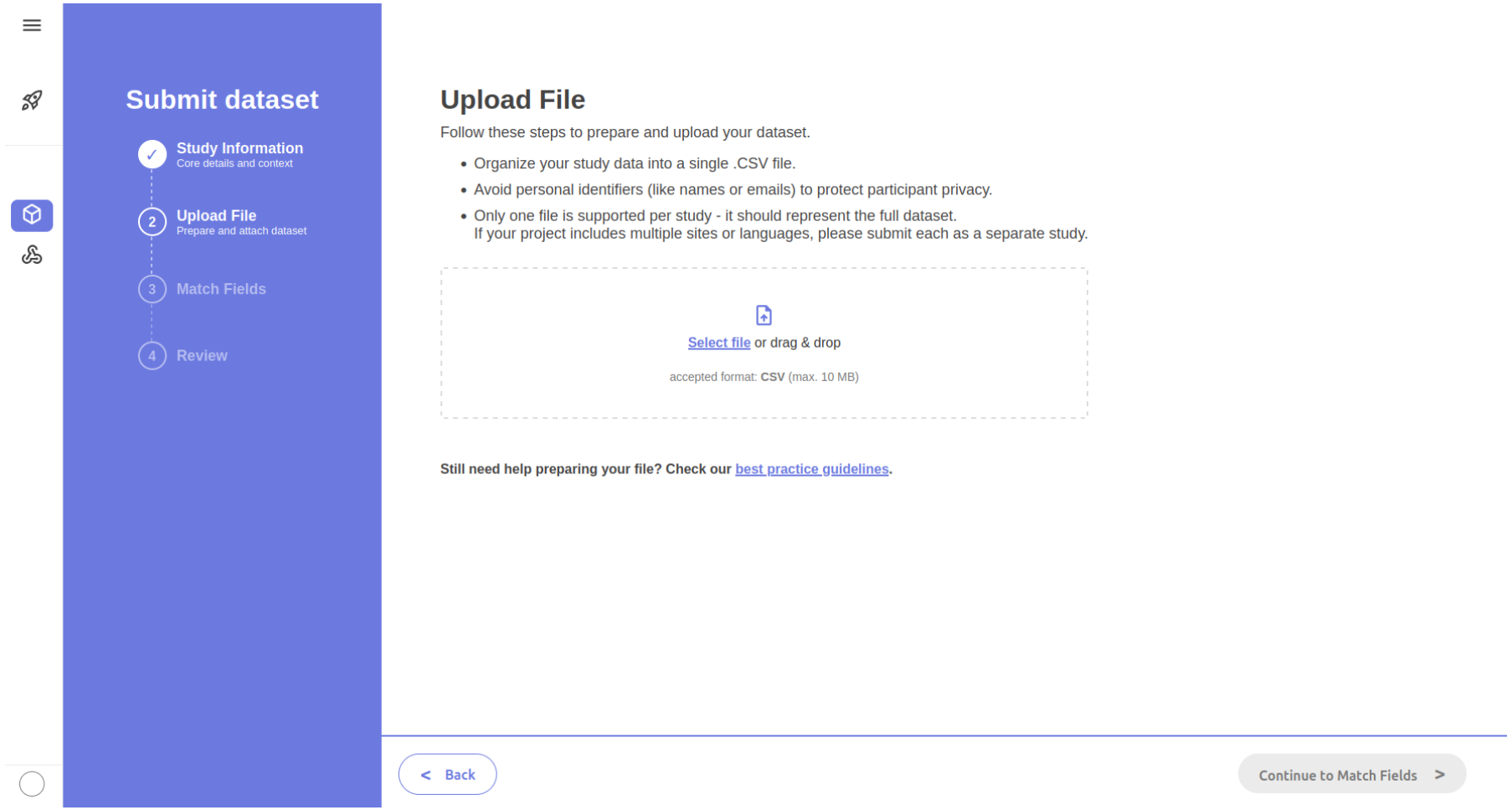

Figura 3.9: Interface da página do carregamento do CSV

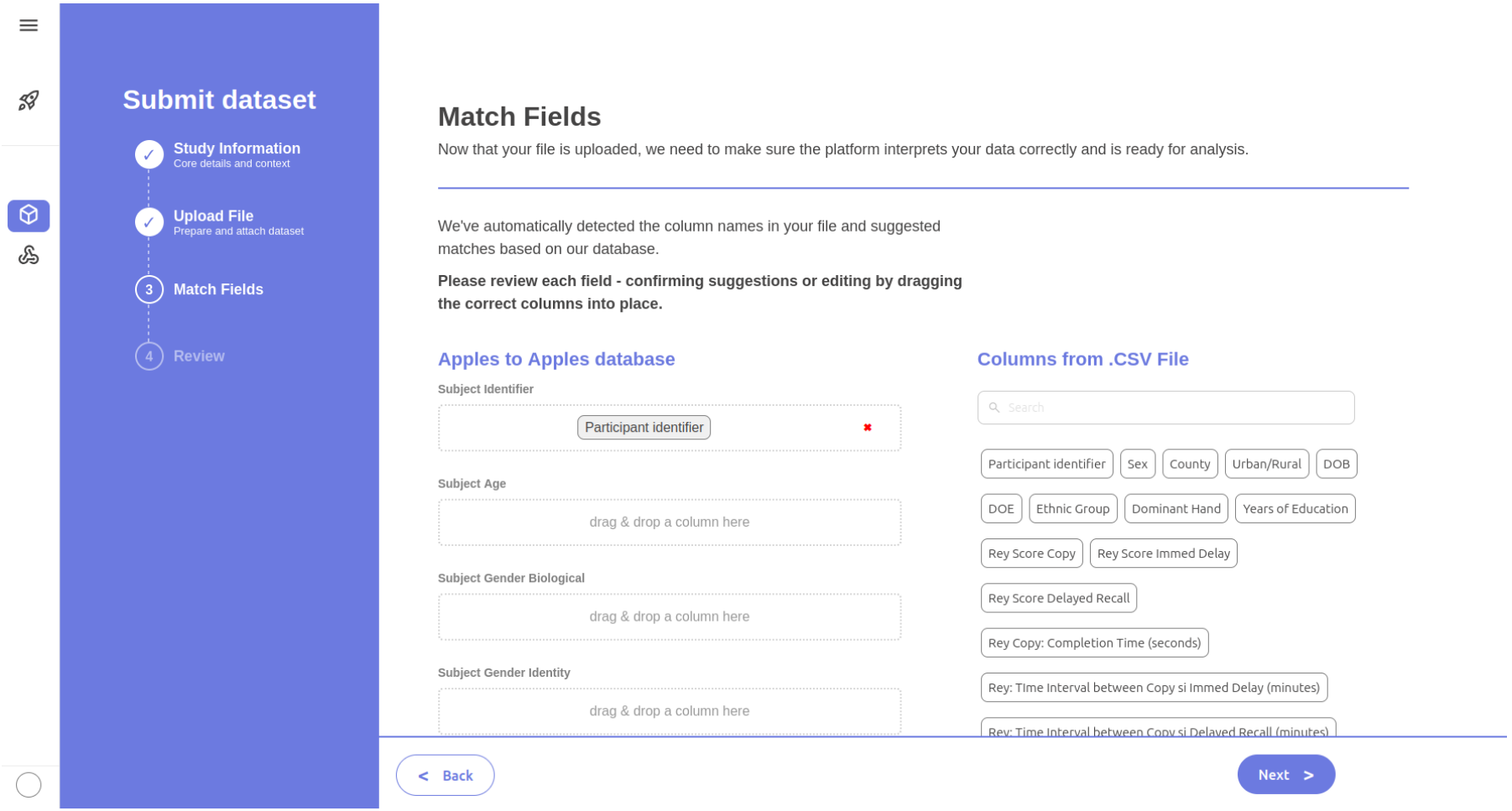

Figura 3.10: Interface da página do mapeamento dos campo

Comparação Normativa

Como mencionado anteriormente, a comparação normativa em neuropsicologia envolve comparar o desempenho de um indivíduo num teste com um grupo 'norma' pré-estabelecido de indivíduos saudáveis. Se o desempenho de um indivíduo se desviar significativamente do grupo norma, podemos então aferi-lo como atípico [Zadelaar et al., 2017]. A comparação normativa ajuda os clínicos e os investigadores a determinar se o desempenho de um indivíduo se enquadra na faixa esperada para um grupo de referência definido, tendo em consideração fatores demográficos e contextuais relevantes.

Os conjuntos de dados normativos tradicionais são frequentemente limitados em tamanho ou diversidade, tornando a comparação mais difícil para algumas populações. Esta aplicação irá realizar comparações normativas, utilizando idade, idioma, país/município, género e nível de educação para comparar os dados de um indivíduo com o conjunto de dados disponível.

Na aplicação a comparação normativa é uma funcionalidade importante que permite ao clínico comparar os dados do seu utente com os dados da base de dados do sistema. O resultado será a classificação da pontuação desse sujeito, ou seja, ajudará a determinar se os resultados do sujeito se enquadram nos intervalos normativos esperados.

Metodologia da implementação da solução

O processo comparativo começa com a inserção da informação do indivíduo a ser comparado. O clínico deve inserir os dados demográficos do indivíduo, incluindo a idade, o género, o nível de educação, o país, a etnicidade, a língua e o estado socioeconómico. No entanto, nem todos estes parâmetros são de caráter obrigatório; o estado socioeconómico e a etnicidade são opcionais. De seguida, o clínico gera um identificador para o indivíduo, que no futuro será utilizado para repetir o processo de comparação sem a necessidade de voltar a preencher os seus dados demográficos. Por fim, o utilizador seleciona os testes que pretende utilizar para comparar o seu utente com a população de referência. À medida que seleciona cada teste, são também introduzidos os *scores* obtidos pelo indivíduo em cada um dos seus parâmetros.

Após a seleção dos testes e a inserção dos valores, a aplicação realiza uma pesquisa à base de dados de acordo com as características selecionadas e recolhe as pontuações relevantes. O resultado da pesquisa é convertido num *dataframe*, onde as colunas correspondem às variáveis demográficas e aos parâmetros dos testes, e as linhas representam os indivíduos presentes na base de dados. A partir deste *dataframe*, é feita a normalização das variáveis categóricas, divido-o em vários *dataframes*, um para cada teste selecionado. A normalização efetuada nesta etapa utiliza as mesmas técnicas já mencionadas na Secção 3.3 da limpeza dos dados. Estas técnicas consistem na conversão do texto para letras minúsculas, no mapeamento dos géneros,

normalização baseada na similaridade textual e na correção dos países, com auxílio das bibliotecas *rapidfuzz* e *pycountry*.

De seguida, para cada *dataframe* de teste, é iniciado o processo de comparação normativa em cada um dos seus parâmetros, ou seja, é avaliado o desempenho do indivíduo, parâmetro a parâmetro para cada teste. Começa-se por definir o *dataframe* correspondente para cada parâmetro, que inclui como atributos as variáveis demográficas e os respetivos valores do parâmetro. Nesta fase, as únicas variáveis consideradas na comparação são a idade e a língua. Após a sua construção e da estruturação das variáveis, é realizada uma pequena limpeza ao conjunto de dados, que consiste na remoção das instâncias que apresentam valores *NaN*, tanto nas variáveis demográficas, como nas quantitativas.

Com os dados já normalizados, foi calculado o vetor das médias e a matriz de covariância das variáveis selecionadas (incluindo variáveis demográficas e quantitativas). Para assegurar a estabilidade numérica durante a inversão da matriz de covariância, foi aplicada uma regularização diagonal (foi acrescentado um pequeno valor à diagonal principal). O modelo foi criado através de uma distribuição normal multivariada, ou seja, usando uma função da biblioteca *NumPy* (`numpy.random.multivariate_normal`) que gera amostras aleatórias de uma distribuição normal multivariada, a fim de criar um modelo para os dados. Embora este modelo tenha sido instanciado, apenas foi utilizado para modelar a incapacidade do indivíduo numa fase exploratória, no entanto, não contribuiu para a determinação do nível de incapacitação exibido ao utilizador.

Durante a comparação individual para cada parâmetro, os dados do indivíduo são pré-processados, com base nas mesmas técnicas utilizadas para limpar os dados, e organizados num vetor com a mesma ordem das variáveis do conjunto. Quando existem valores em falta ou categorias desconhecidas, é utilizado o valor médio normativo dessa variável.

Por fim, são calculadas algumas medidas estatísticas (desvio padrão) com os valores do indivíduo e do conjunto normalizados. São calculados os `z-scores` individuais para cada variável, com inversão do sinal nos parâmetros temporais (onde valores mais elevados correspondem a um pior desempenho). O percentil é estimado com base na distribuição normal do `z-score` principal e gerada uma interpretação clínica automática, incluindo a indicação de um possível desempenho anómalo. Os resultados são enviados para o *frontend*, juntamente com um gráfico da curva da distribuição normal padrão de cada parâmetro, no qual é representado o respetivo valor de `z-score`. Este gráfico facilita a interpretação e compreensão do resultado obtido na comparação de cada parâmetro.

Desta forma, o sistema possibilita uma comparação normativa robusta, ao integrar simultaneamente diversas variáveis demográficas e parâmetros do teste numa única métrica de desvio clínico.

Embora o conjunto de dados atual seja limitado em tamanho e diversidade, a arqui-

tetura implementada é escalável e consistente, permitindo a atualização da comparação normativa com amostras mais amplas, sem necessidade de alterar a estrutura e a interpretação estatística. Esta limitação atual origina, em alguns casos, erros estatísticos, nomeadamente no cálculo dos `z-scores`, podendo estes possuir valores absurdos. No entanto, com a crescente adição de dados, este problema será mitigado.

Interface

Esta secção apresenta a interface de cada uma das etapas realizadas durante o processo da comparação. A Figura 3.11 ilustra a interface da primeira página/etapa, onde são inseridas as informações demográficas do indivíduo a comparar. A Figura 3.12 ilustra a interface da página que gera um identificador único para o indivíduo. As Figuras 3.13 e 3.14 ilustram as interfaces da página da seleção dos testes a utilizar para a comparação e da inserção dos resultados do indivíduo para cada parâmetro desses testes. Por fim, a Figura 3.15 ilustra a interface dos resultados da comparação.

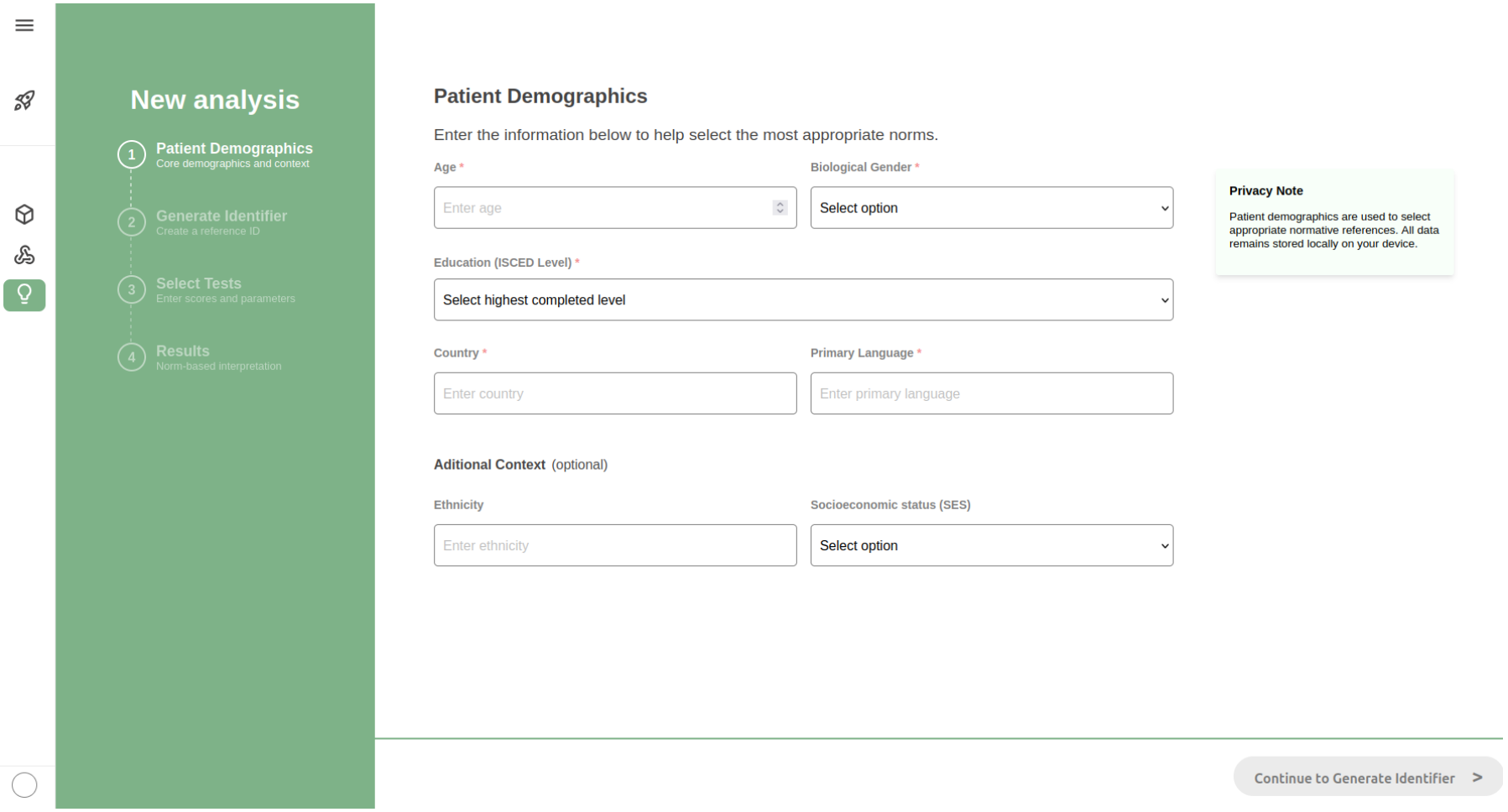

Figura 3.11: Interface da inserção dos dados demográficos do indivíduo

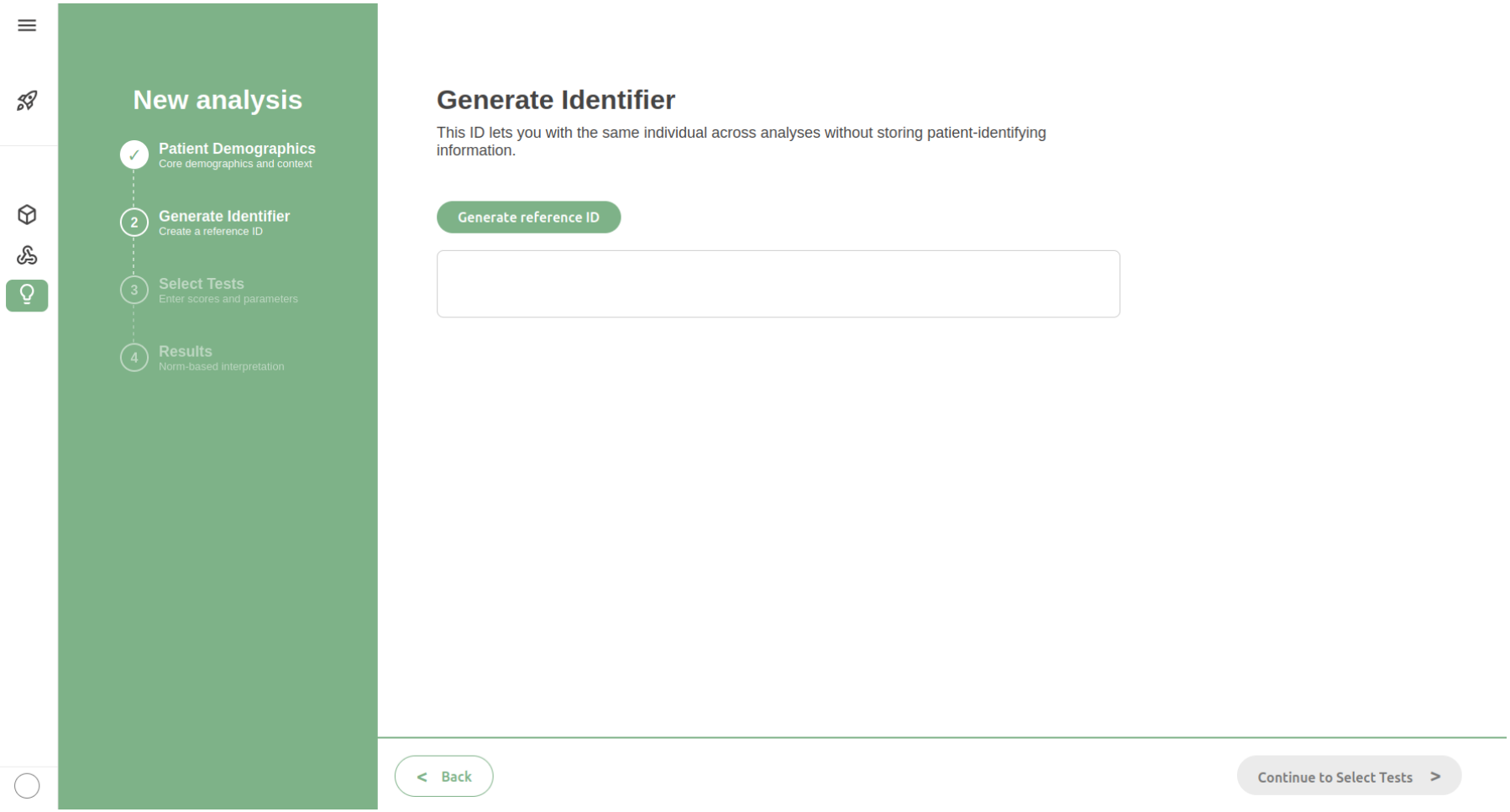

Figura 3.12: Interface da página que gera o identificador único para o indivíduo

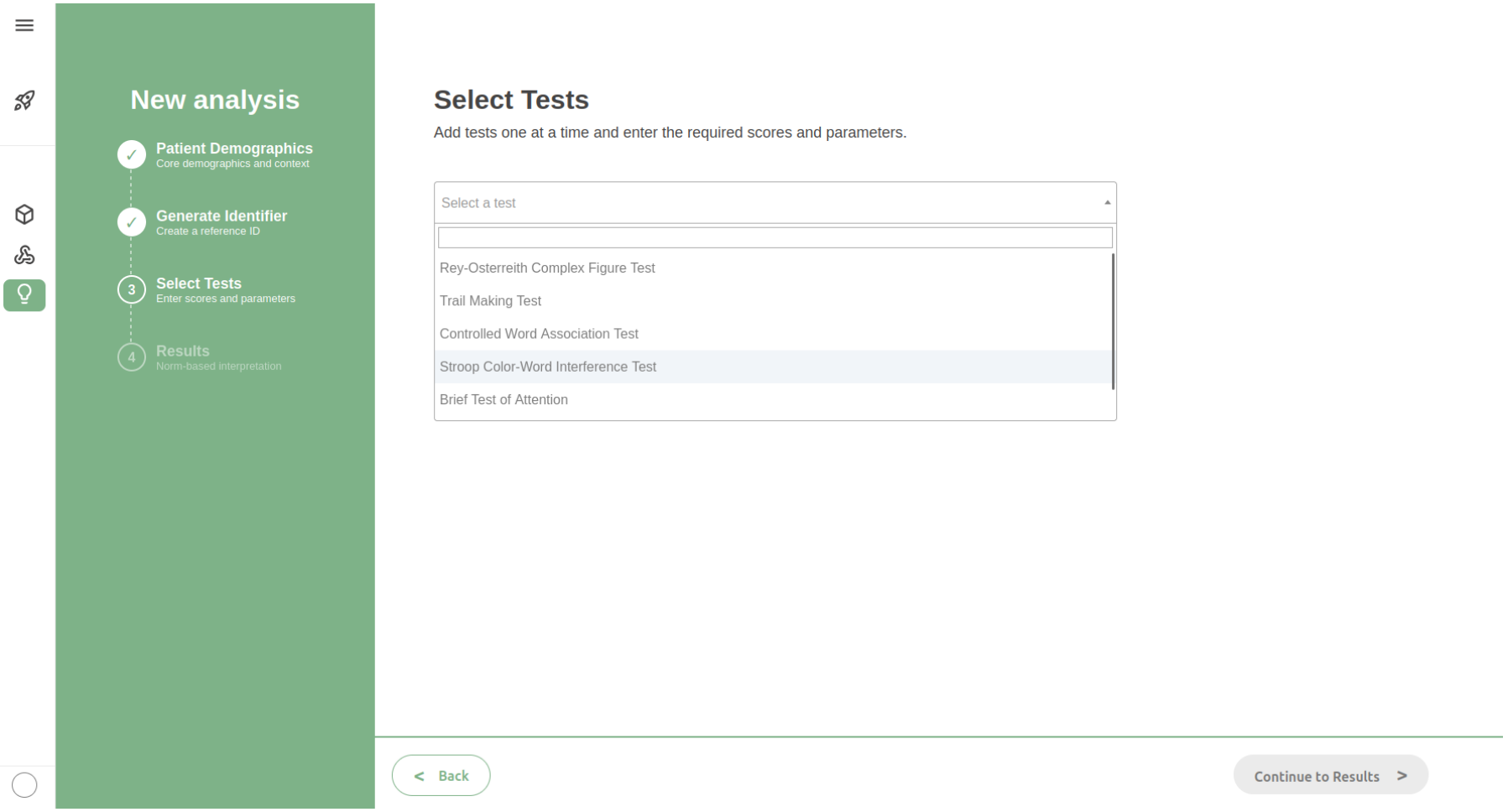

Figura 3.13: Interface da seleção dos teste a utilizar na comparação

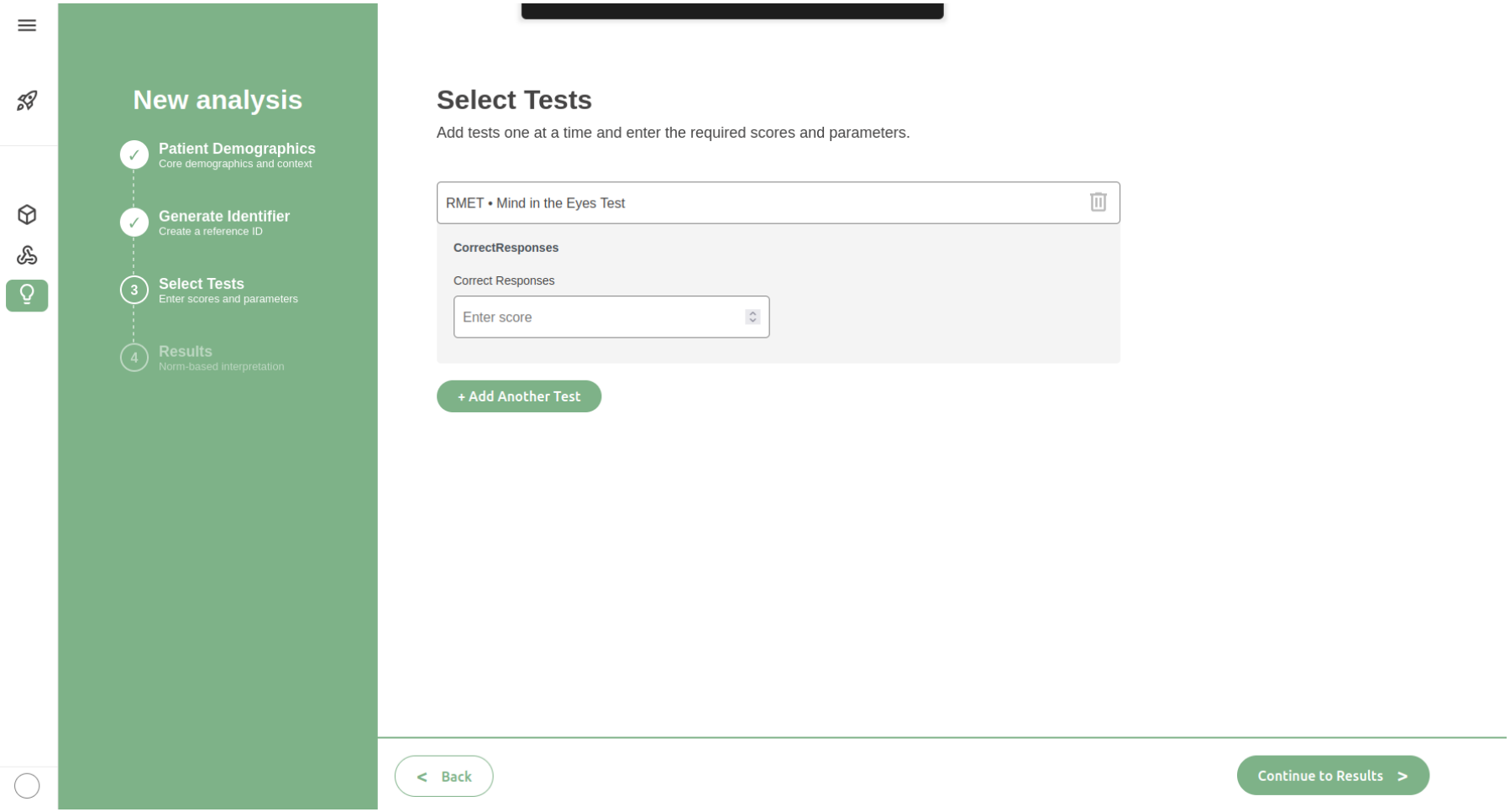

Figura 3.14: Interface da inserção dos resultados do indivíduo para cada parâmetro dos testes

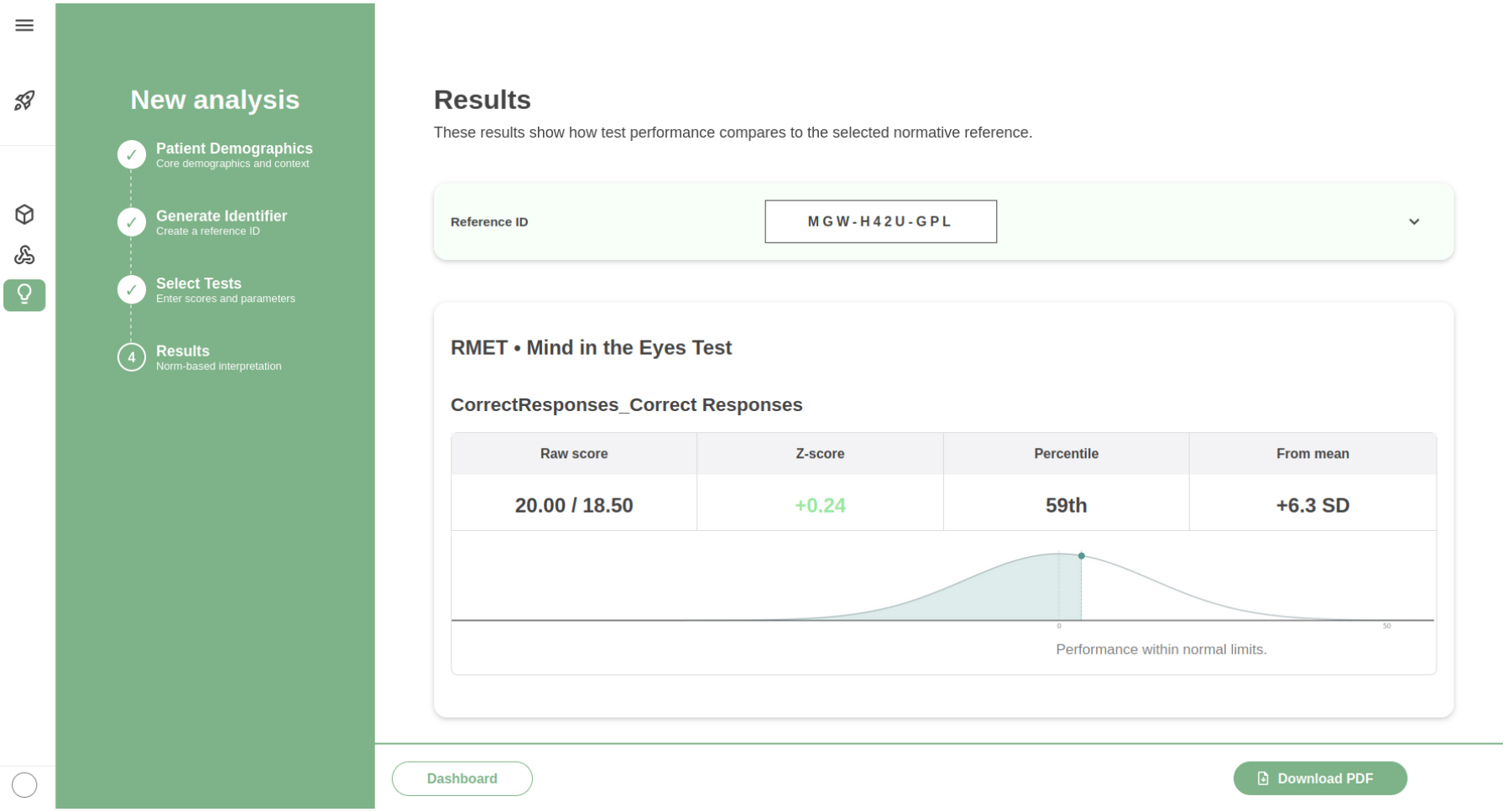

Figura 3.15: Interface da página dos resultados obtidos na comparação

Página de Pesquisa e Página Principal

A aplicação possui uma interface para pesquisar e filtrar dados armazenados. Os utilizadores podem realizar pesquisas com base em vários atributos, tais como: diferença de idade, país, idioma, teste, género e outros. Os resultados são apresentados de forma paginada, o que permite o acesso rápido e eficiente a dados específicos. Esta funcionalidade é essencial para investigadores e clínicos, que podem selecionar subconjuntos de testes neuropsicológicos para análise ou comparação.

A página principal corresponde à primeira página que os utilizadores têm acesso. Esta página corresponde ao *dashboard* para cada perfil de utilizador, ou seja, esta página varia consoante o perfil do utilizador autenticado, apresentando uma interface personalizada e apenas as funcionalidades a que este tem acesso, a fim de facilitar a navegação na plataforma Web. A solução proposta permite a criação de dois tipos de perfis: Clínico e Investigador. O Clínico é capaz de pesquisar resultados e realizar comparações normativas. E o Investigador consegue efetuar o *upload* dos seus estudos e realizar pesquisas de resultados. É portanto, extremamente essencial, permitindo aos utilizadores navegarem pelas várias páginas/funcionalidade da aplicação.

Interfaces

Nesta secção são apresentadas as interfaces da página principal para cada tipo de utilizador, uma vez que a interface da pesquisa ainda se encontra em desenvolvimento. A interface desta página está ilustrada na Figura 3.16 para o *role* do Clínico e na Figura 3.17 para o do Investigador.

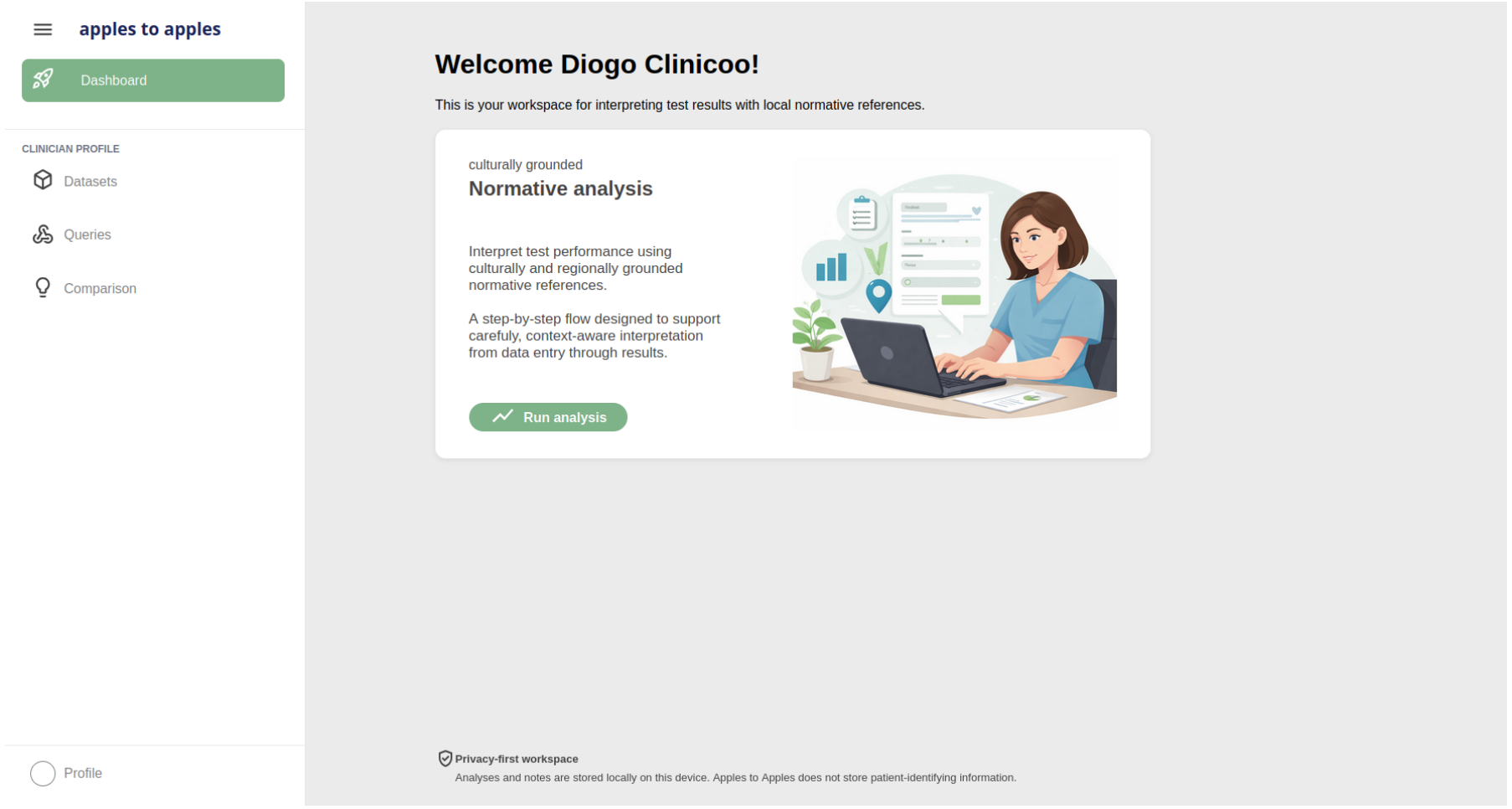

Figura 3.16: Interface da página principal para os clínicos

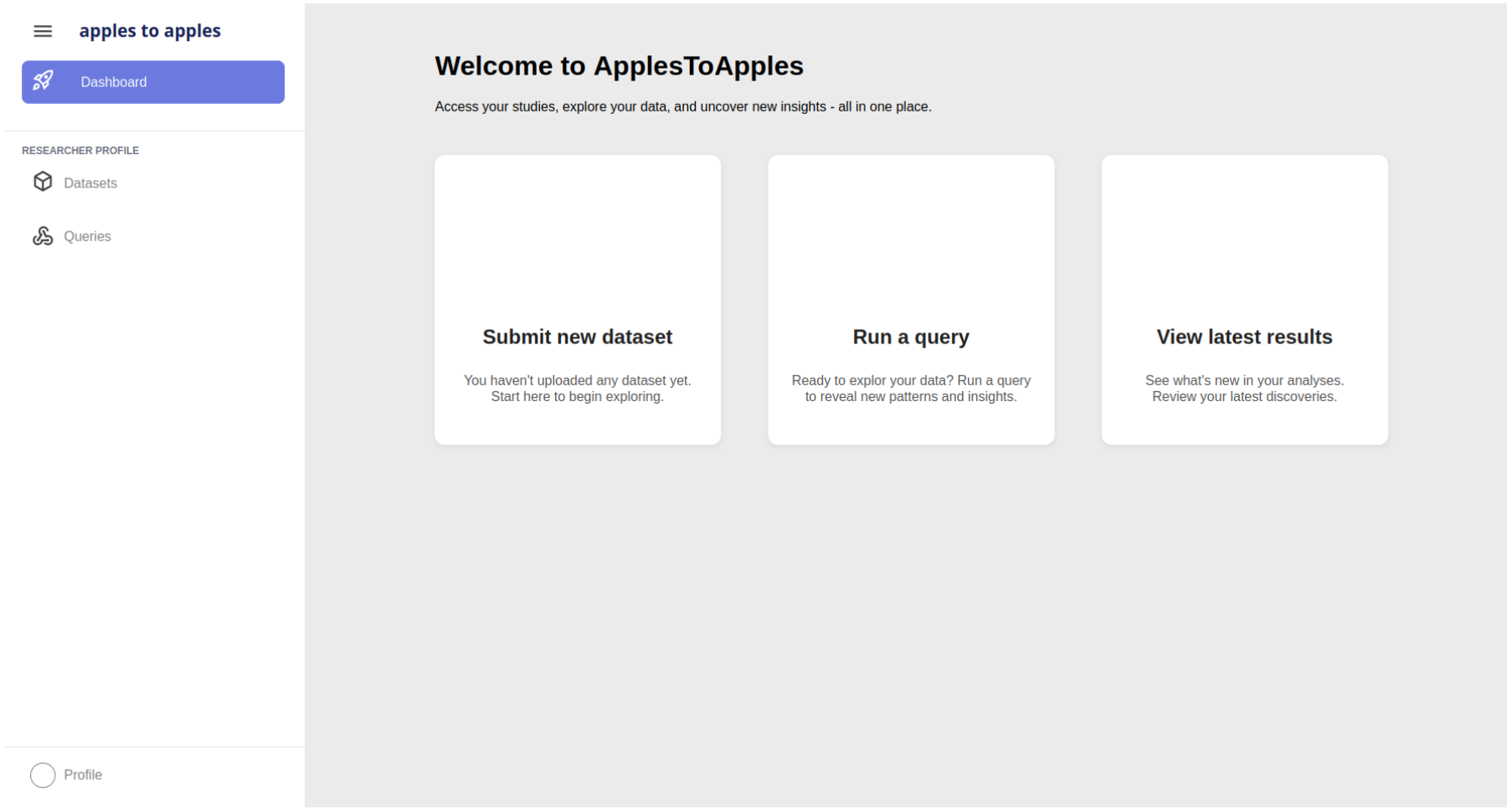

Figura 3.17: Interface da página principal para os investigadores

Autenticação e registo de utilizadores

A autenticação constitui um elemento essencial da aplicação, assegurada através das páginas iniciar sessão e registo de novos utilizadores. É ainda possível redefinir a palavra-passe de cada utilizador, sendo para tal necessário validar a conta através de um código enviado para o seu email. Durante a criação da conta é exigida a seleção do perfil, a inserção de informações necessárias, como nome do utilizador, email e dados específicos para cada perfil (por exemplo, para o investigador, é necessário inserir os dados como: afiliação e identificador único do investigador) e a validação do email, através do código enviado.

Interface

Nesta secção são apresentadas as interfaces das páginas de início de sessão, ilustrada pela Figura 3.18 e registo de novos utilizadores, ilustrada pela Figura 3.19.

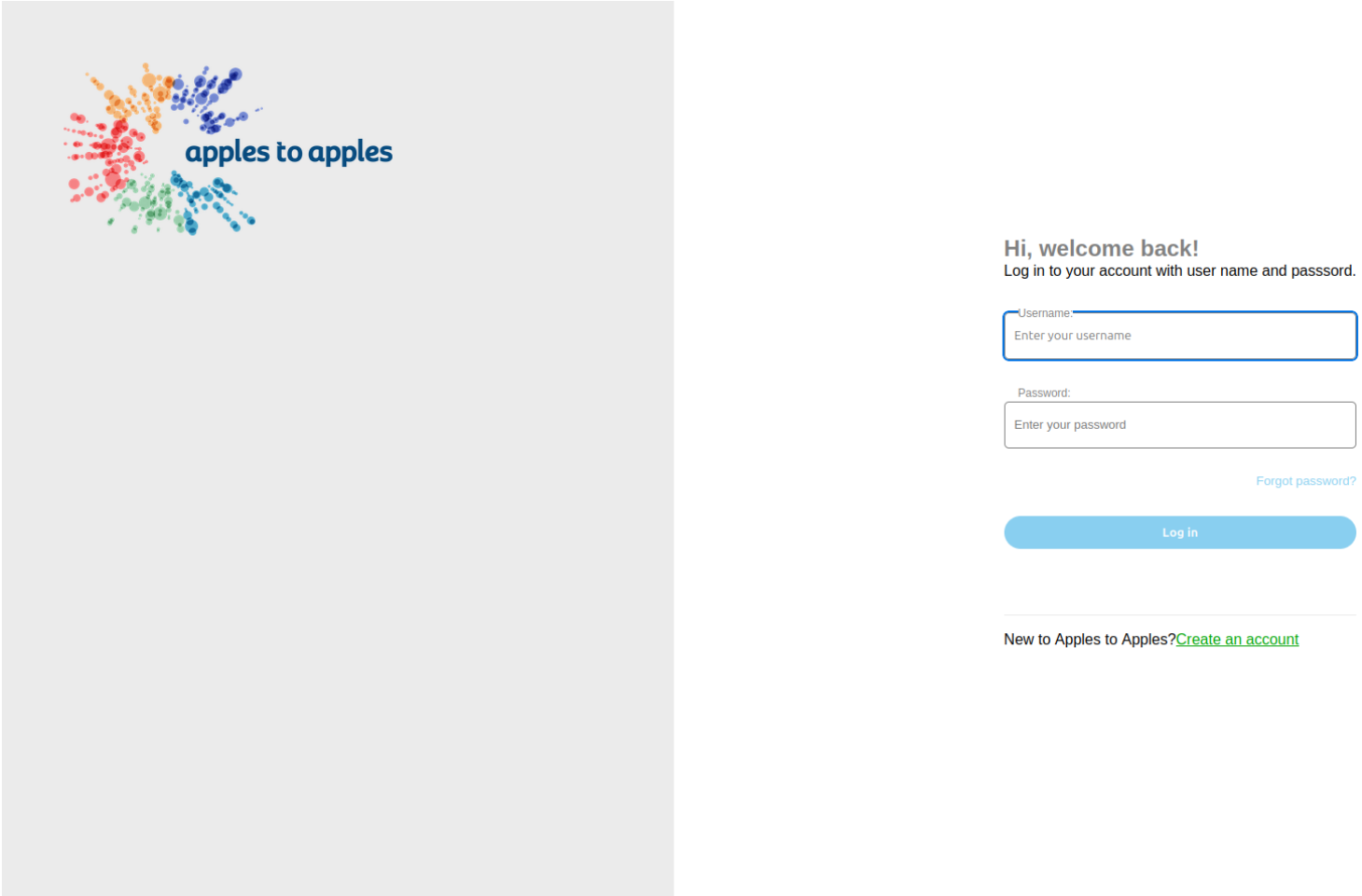

Figura 3.18: Interface da página de iniciar sessão

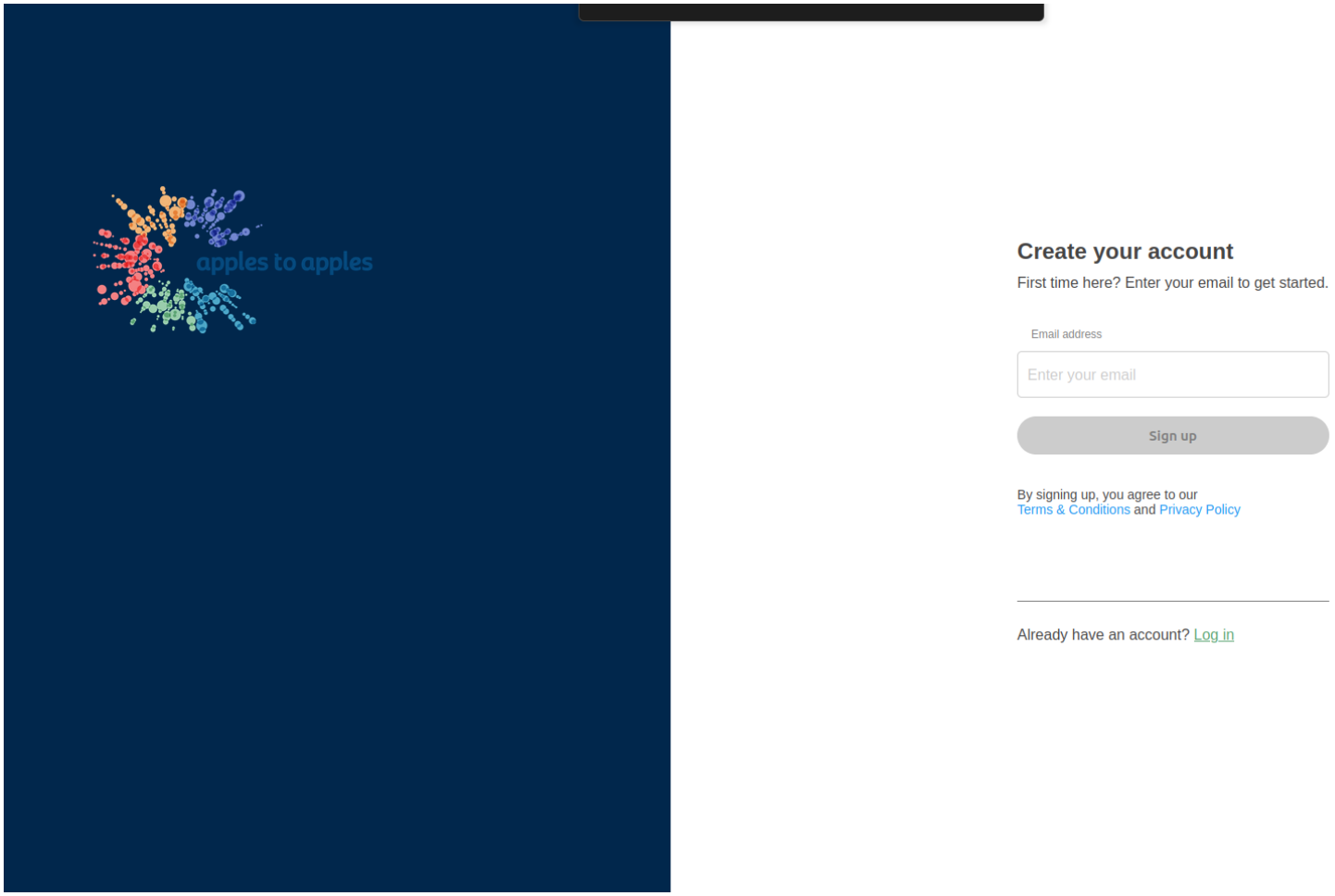

Figura 3.19: Interface da página de registo de utilizadores

Capítulo 4

Conclusões e trabalho futuro

Esta dissertação examina as abordagens existentes na gestão de testes neuropsicológicos (recolha, armazenamento e análise). Apresenta também o *design* e a implementação de uma aplicação baseada na Web, concebida para aprimorar a recolha, o armazenamento e a análise desses dados. Esta plataforma aborda desafios fundamentais nos testes neuropsicológicos, particularmente a necessidade de uma estrutura de dados flexível e escalável. Também permite comparações normativas mais adaptáveis culturalmente. Ao permitir que os utilizadores carreguem dados de estudos com segurança, realizem pesquisas direcionadas e comparem pontuações de testes individuais com uma base de dados normativa em expansão dinâmica, a aplicação oferece uma ferramenta prática e robusta para fins clínicos e de investigação.

Do ponto de vista da engenharia de *software*, este projeto demonstra a integração bem-sucedida da modelação dinâmica de dados, fluxos de trabalho seguros baseados em funções e análises numa aplicação Web unificada. Ao contrário das ferramentas neuropsicológicas existentes, este sistema suporta um esquema modular capaz de incorporar novos testes e parâmetros, ao mesmo tempo que realiza comparações normativas em tempo real.

O *design* da base de dados exigiu uma maior atenção para lidar com uma ampla variedade de testes e parâmetros, mantendo a consistência e o rigor analítico. Permite ainda a atualização dos testes existentes e adicionar novos, conforme necessário. Devido à escolha da base de dados, também requer a adição de informações sobre condições, categorias e parâmetros para cada teste.

Os principais desafios no desenvolvimento giraram em torno do *design* da base de dados, especificamente na criação de uma estrutura que fosse escalável o suficiente para gerir a complexidade inerente dos dados. Inicialmente a comparação normativa também foi um obstáculo, uma vez que, implicava a análise da amostra de referência, o cálculo dos `z-scores` e a padronização rigorosa das variáveis para assegurar comparabilidade estatística entre o indivíduo e o grupo normativo. No entanto, o maior desafio na comparação normativa residiu na sua definição conceptual e metodológica. Este desafio manifestou-se em várias dimensões: a compreensão do próprio

procedimento, a seleção dos métodos estatísticos mais adequados e a identificação das variáveis relevantes. Revelou-se igualmente desafiante integrar simultaneamente múltiplas variáveis demográficas e parâmetros dos testes numa única métrica quantitativa e interpretar adequadamente os resultados obtidos.

Após a superação destes desafios, foi possível desenvolver uma solução funcional que simplifica a gestão dos dados neuropsicológicos complexos e permite, simultaneamente, comparações normativas significativas. Ao facilitar uma recolha e análise de dados mais ampla e inclusiva, a aplicação tem o potencial de aumentar a precisão e a relevância das avaliações neuropsicológicas em diversas populações.

A arquitetura da aplicação suporta melhorias e aprimoramentos futuros. Desta forma, o trabalho futuro irá consistir em:

- Validação e qualidade do sistema: O trabalho futuro irá centrar-se na validação da aplicação, ou seja, em testar todas as funcionalidades à procura de erros, a fim de os corrigir. Isto inclui a validação em relação às avaliações clínicas para garantir a relevância da análise e a expansão do conjunto de dados para melhorar a precisão da comparação. Será ainda necessário realizar uma limpeza no código;
- Interface e usabilidade: O desenvolvimento futuro irá ainda centrar-se na melhoria da interface Web da aplicação (*frontend*), de acordo com as suas especificações, incluindo a implementação do *design* da página de pesquisa;
- Novas funcionalidades: Desenvolvimento de novas funcionalidades, como a capacidade do utilizador armazenar temporariamente os dados do seu paciente, com o objetivo de realizar futuras comparações sem a necessidade de re-inserir os seus dados, melhorando a experiência de utilização. Outras funcionalidades serão a capacidade de transferir os resultados da comparação para um ficheiro PDF e permitir aos clínicos/investigadores visualizarem o histórico das suas comparações/inserções;
- Autenticação e segurança: A autenticação também será atualizada com o envio de um código para o email do utilizador a partir do email da aplicação, para o validar;
- Testes em ambiente real: Por fim, aplicação deverá ser testada exaustivamente em ambiente real, utilizando dados reais de indivíduos e envolvendo utilizadores reais. Isto simulará o fluxo normal de *uploads* de dados, pesquisas e comparações normativas, sendo crucial para identificar e corrigir quaisquer problemas, erros, validações ou potenciais falhas, garantindo uma produção robusta e sem erros.

Bibliografia

Abdellaoui, S., Abbaoui, W., Meziane, L., Bhiri, B. E., and Ziti, S. (2025). Sql and nosql databases: A comparative study with perspectives on ia-based migration approach. *Engineering Proceedings*, 112(1).

Aggarwal, S. (2014). *Flask framework cookbook*. Packt Publishing Ltd.

Analytical Methods Committee AMCTB No. 93 (2020). To p or not to p: the use of p-values in analytical science. *Analytical Methods*, 12(6):872–874.

Andrade, C. (2021). Z scores, standard scores, and composite test scores explained. *Indian Journal of Psychological Medicine*, 43(6):555–557.

Arampatzi, X., Margioti, E. S., Messinis, L., Yannakoulia, M., Hadjigeorgiou, G., Dardiotis, E., Sakka, P., Scarmeas, N., and Kosmidis, M. H. (2024). Development of robust normative data for the neuropsychological assessment of greek older adults. *Journal of the International Neuropsychological Society*, 30(6):594602.

Ashendorf, L., Jefferson, A. L., O'Connor, M. K., Chaisson, C., Green, R. C., and Stern, R. A. (2008). Trail making test errors in normal aging, mild cognitive impairment, and dementia. *Archives of Clinical Neuropsychology*, 23(2):129–137.

Ayhan, Y., Akbulut, B. B., Şişman, A. H., and Velibaşoğlu, B. (2022). Psinorm: A fast, efficient and free open-source software for interpreting, reporting and archiving neuropsychological test results. *Turk Psikiyatri Dergisi*, 33(4):255–262. English, Turkish.

Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *Journal of Multivariate Analysis*, 88(1):190–206.

Batra, S., Sachdeva, S., and Bhalla, S. (2018). Entity attribute value style modeling approach for archetype based data. *Information*, 9(1).

Blagus, R., Leskoek, B., Ortega, F. B., Tomkinson, G., and Jurak, G. (2025). Recommendations for reporting regression-based norms and the development of free--access tools to implement them in practice. *PLOS ONE*, 20(6):1–12.

Brandt, C. A., Morse, R., Matthews, K., Sun, K., Deshpande, A. M., Gadagkar, R., Cohen, D. B., Miller, P. L., and Nadkarni, P. M. (2002). Metadata-driven creation of data marts from an eav-modeled clinical research database. *International Journal of Medical Informatics*, 65(3):225–241.

Bryant, A. and Freilich, B. and Papova, A. (2025). Excel with excel. https://theaacn.org/disruptive-technology-initiative/excel-with-excel/. Consultado em: 2025-10-13.

Chervinsky, A. B. (2007). Neuropsychological assessment platform (npap) and method. https://patentimages.storage.googleapis.com/66/da/26/fe2b14763bb0a5/US20070123757A1.pdf.

Cleveland Clinic (2023). Neuropsychological testing and assessment. https://my.clevelandclinic.org/health/diagnostics/4893-neuropsychological-testing-and-assessment. Consultado em: 2025-10-13.

CodeNet (2025). Bases de dados relacionais. https://codenet.pt/bases-de-dados/bases-de-dados-relacionais/. Consultado em: 2025-10-13.

CRAN (2025). Normdata: Derivation of regression-based normative data. https://cran.r-project.org/web/packages/NormData/index.html. Consultado em: 2025-10-13.

de Vent, N. R. (2021). *Advanced Neuropsychological Diagnostics Infrastructure: Improving Neuropsychology*. Phd thesis, University of Amsterdam. Tese entregue a 21 de Janeiro de 2021.

de Vent, N. R., Agelink van Rentergem, J. A., Schmand, B. A., Murre, J. M. J., , A. C., and Huizenga, H. M. (2016). Advanced neuropsychological diagnostics infrastructure (andi): A normative database created from control datasets. *Frontiers in Psychology*, Volume 7 - 2016.

Defense and Veterans Brain Injury Center (2009). Automated neuropsychological assessment metrics (anam). Brochure, Proponency Office for Rehabilitation & Reintegration, Version 3: 11 August 2009. Consultado em: 2025-10-13.

Defense Health Agency (2024). Automated neuropsychological assessment metrics database. https://www.health.mil/Reference-Center/Forms/2024/07/01/Automated-Neuropsychological-Assessment-Metrics-Database. Privacy Impact Assessment form. Consultado em: 2025-10-13.

delCacho Tena, A., Christ, B. R., Arango-Lasprilla, J. C., Perrin, P. B., Rivera, D., and Olabarrieta-Landa, L. (2023). Normative data estimation in neuropsychological tests: A systematic review. *Archives of Clinical Neuropsychology*, 39(3):383–398.

Dinu, V. and Nadkarni, P. (2007). Guidelines for the effective use of entityattributevalue modeling for biomedical databases. *International Journal of Medical Informatics*, 76(11):769–779.

DiPasquale, Z. T. (2024). *Revisiting the Global Measures of Overall Neuropsychological Revisiting the Global Measures of Overall Neuropsychological Function and the Impact of Race-based Normative Data in Clinical and Forensic Populations*. Dissertation, Philadelphia College of Osteopathic Medicine.

Django (2025). Why django? https://www.djangoproject.com/start/overview/. Consultado em: 2025-10-13.

Django Software Foundation (2026). Django: The web framework for perfectionists with deadlines. https://www.djangoproject.com/. Consultado em: 28 de março de 2026.

Dominik Liebler (2020). Entity-attribute-value (eav). https://designpatternsphp.readthedocs.io/pt-br/latest/More/EAV/README.html. Consultado em: 2025-10-13.

Elliot McClenaghan (2024). Mann-whitney u test: Assumptions and example. https://www.technologynetworks.com/informatics/articles/mann-whitney-u-test-assumptions-and-example-363425. Consultado em: 2025-10-13.

Emerson, R. W. (2023). Mann-whitney u test and t-test. *Journal of Visual Impairment & Blindness*, 117(1):99–100.

Espírito Santo, H., Lemos, L., Fernandes, D., Cardoso, D., Neves, C. S., Caldas, L., Pascoal, V., Marques, M., Guadalupe, S., and Daniel, F. (2015). Teste de stroop. *I&D CINEICC*.

EUI(European University Institute) (2025). Web of science. https://www.eui.eu/Research/Library/ResearchGuides/Economics/WebofScience. Consultado em: 2025-10-13.

fawwazmts (2024). Z-score and modified z-score. https://medium.com/@fawwazmts/z-score-and-modified-z-score-f689296e4d3a. Consultado em: 10 de junho de 2026.

FUIOR, F. (2021). Introduction in python frameworks for web development. *Romanian Journal of Information Technology and Automatic Control*, 31(3):97–108.

Gallagher, J., Mamikonyan, E., Xie, S. X., et al. (2023). Validating virtual administration of neuropsychological testing in parkinson disease: a pilot study. *Scientific Reports*, 13:16243.

Gary, S., Lenhard, W., and Lenhard, A. (2021). Modelling norm scores with the cnorm package in r. *Psych*, 3(3):501–521.

Ghimire, D. (2020). Comparative study on Python web frameworks: Flask and Django. Bachelor's thesis, Metropolia University of Applied Sciences.

Goette, W. F., Carlew, A. R., Schaffert, J., Mokhtari, B. K., and Cullum, C. M. (2023). Validation of a bayesian diagnostic and inferential model for evidence-based neuropsychological practice. *Journal of the International Neuropsychological Society*, 29(2):182192.

GoogleCloud (2025a). O que é um banco de dados nosql? https://cloud.google.com/discover/what-is-nosql?hl=pt-BR. Consultado em: 2025-10-13.

GoogleCloud (2025b). O que é um banco de dados relacional? https://cloud.google.com/learn/what-is-a-relational-database. Consultado em: 2025-10-13.

Gordon, E., Cooper, N., Rennie, C., Hermens, D., and Williams, L. M. (2005). Integrative neuroscience: the role of a standardized database. *Clinical EEG and Neuroscience*, 36(2):64–75.

Grasman, R. P., Huizenga, H. M., and Geurts, H. M. (2010). Departure from normality in multivariate normative comparison: The cramér alternative for hotelling's t2. *Neuropsychologia*, 48(5):1510–1516.

Han, D. Y., Hoelzle, J. B., Dennis, B. C., and Hoffmann, M. (2011). A brief review of cognitive assessment in neurotoxicology. *Neurologic Clinics*, 29(3):581–590. Frontiers in Clinical Neurotoxicology.

Harvey, P. D. (2012). Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience*, 14(1):91–99.

Hassan, M. A. (2021). Relational and nosql databases: The appropriate database model choice. In *2021 22nd International Arab Conference on Information Technology (ACIT)*, pages 1–6.

Health Center Program (2016). Uniform data system (uds). https://www.cms.gov/files/document/sgm-clearinghouse-uds.pdf. Consultado em: 2025-10-13.

Howieson, D. (2019). Current limitations of neuropsychological tests and assessment procedures. *The Clinical Neuropsychologist*, 33(2):200–208. PMID: 30608020.

Huizenga, H. M., Smeding, H., Grasman, R. P., and Schmand, B. (2007). Multivariate normative comparisons. *Neuropsychologia*, 45(11):2534–2542.

IBM (2025a). Analysis of variance (anova). https://www.ibm.com/docs/en/cognos-analytics/11.2.x?topic=tests-analysis-variance-anova. Consultado em: 2025-10-13.

IBM (2025b). What is a nosql database? https://www.ibm.com/think/topics/nosql-databases. Consultado em: 2025-10-13.

IBM (2025c). What is a relational database? https://www.ibm.com/think/topics/relational-databases#:~:text=What%20is%20a%20relational%20database?%20*%20A,a%20primary%20key%20or%20a%20foreign%20key. Consultado em: 2025-10-13.

IBM (2026). O que é um esquema de banco de dados? https://www.ibm.com/br-pt/think/topics/database-schema. Consultado em: 2026-03-18.

Innocenti, F., Candel, M. J. J. M., Tan, F. E. S., and van Breukelen, G. J. P. (2024). Sample size calculation and optimal design for multivariate regression-based norming. *Journal of Educational and Behavioral Statistics*, 49(5):817–847.

Iverson, G. L. (2011). *T Scores*, pages 2459–2460. Springer New York, New York, NY.

Jaqueline Terres Borges (2025). Testes cognitivos e neuropsicológicos: veja o que são e os principais. https://www.psitto.com.br/blog/testes-cognitivos-neuropsicologicos/#:~:text=Testes%20neuropsicol%C3%B3gicos%20mais%20utilizados%20Os%20testes%20neuropsicol%C3%B3gicos,avaliar%20flexibilidade%20cognitiva%20e%20velocidade%20de%20processamento. Consultado em: 2025-10-13.

Katie Marvin (2012). Trail making test (tmt). https://strokengine.ca/en/assessments/trail-making-test-tmt/. Consultado em: 2025-10-13.

Katie McLaughlin (2017). An introduction to the django orm. https://opensource.com/article/17/11/django-orm. Consultado em: 2025-10-13.

Kenton, W. (2025). What is analysis of variance (anova)? https://www.investopedia.com/terms/a/anova.asp#toc-what-is-analysis-of-variance-anova. Consultado em: 2025-10-13.

Kessels, R. P. and Hendriks, M. P. (2023). Neuropsychological assessment. In Friedman, H. S. and Markey, C. H., editors, *Encyclopedia of Mental Health (Third Edition)*, pages 622–628. Academic Press, Oxford, third edition.

Kessels, R. P. C. (2019). Improving precision in neuropsychological assessment: Bridging the gap between classic paper-and-pencil tests and paradigms from cognitive neuroscience. *The Clinical Neuropsychologist*, 33(2):357–368. PMID: 30394172.

Kim, H.-Y. (2019). Statistical notes for clinical researchers: the independent samples t-test. *Restorative Dentistry & Endodontics*, 44(3):e26. DOI: https://doi.org/10.5395/rde.2019.44.e26.

Krishna, G. and Pritibala, V. P. (2025). Python: An inclusive review of its user-friendly nature and applications. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, 3:1872–1875.

Kuntzelman, K. M., Williams, J. M., Lim, P. C., Samal, A., Rao, P. K., and Johnson, M. R. (2021). Deep-learning-based multivariate pattern analysis (dmvpa): A tutorial and a toolbox. *Frontiers in Human Neuroscience*, Volume 15 - 2021.

Lavrencic, L. M., Bennett, H., Daylight, G., Draper, B., Cumming, R., Mack, H., Garvey, G., Lasschuit, D., Hill, T. Y., Chalkley, S., Delbaere, K., Broe, G. A., and Radford, K. (2019). Cognitive test norms and comparison between healthy ageing, mild cognitive impairment, and dementia: A populationbased study of older aboriginal australians. *Australian Journal of Psychology*, 71(3):249–260.

Lähde, N., Basnyat, P., Raitanen, J., Kämppi, L., Lehtimäki, K., Rosti-Otajärvi, E., and Peltola, J. (2024). Complex executive functions assessed by the trail making test (tmt) part b improve more than those assessed by the tmt part a or

digit span backward task during vagus nerve stimulation in patients with drug-resistant epilepsy. *Frontiers in Psychiatry*, 15.

Maggio, M. G., Giambó, F. M., Barbera, M., Pasquale, P. D., Bruno, F., Calderone, A., Quartarone, A., Rizzo, A., and Calabró, R. S. (2025). Moving toward the digitalization of neuropsychological tests: An exploratory study on usability and operator perception. *Digital Health*, 11:20552076251334449.

Mark, H. and Workman, J. (2018). Chapter 75 - the statistics of spectral searches. In *Chemometrics in Spectroscopy (Second Edition)*, pages 507–511. Academic Press, second edition edition.

Martin, C. (2025). Independent samples t-test definition & guide. https://julius.ai/articles/independent-samples-t-test-definition-and-guide. Consultado em: 2025-10-13.

Matheus Santos (2024). Guia prático: Como usar o trail making test na clínica e na pesquisa. https://www.icc.clinic/guia-pratico-como-usar-o-trail-making-test-na-clinica-e-na-pesquisa. Consultado em: 2025-10-13.

McKee, A. J. (2025). T-score | definition. https://docmckee.com/cj/docs-research-glossary/t-score-definition/#:~:text=Limitations%20of%20T%2DScores%20While%20T%2Dscores%20are%20powerful,confuse%20T%2Dscores%20with%20percentages%20or%20raw%20scores. Consultado em: 2025-10-13.

MDN (2025). Introdução ao django. https://developer.mozilla.org/pt-BR/docs/Learn_web_development/Extensions/Server-side/Django/Introduction. Consultado em: 2025-10-13.

Milvus (2025). What are the limitations of relational databases? https://milvus.io/ai-quick-reference/what-are-the-limitations-of-relational-databases. Consultado em: 2025-10-13.

Mirzaei, G. and Adeli, H. (2022). Machine learning techniques for diagnosis of alzheimer disease, mild cognitive disorder, and other types of dementia. *Biomedical Signal Processing and Control*, 72:103293.

Mishra, P., Singh, U., Pandey, C. M., Mishra, P., and Pandey, G. (2019). Application of student's t-test, analysis of variance, and covariance. *Annals of Cardiac Anaesthesia*, 22(4).

Mokhtar, M. A. A. M., Yusoff, N. S., and Liang, C. Z. (2021). Robust hotellings t2 statistic based on m-estimator. *Journal of Physics: Conference Series*, 1988(1):012116.

Morrissey, S., Gillings, R., and Hornberger, M. (2024). Feasibility and reliability of online vs in-person cognitive testing in healthy older people. *Plos one*, 19(8):e0309006.

National Alzheimer's Coordinating Center (NACC) (2025). Uniform data set (uds). https://naccdata.org/data-collection/forms-documentation/uds-4. Consultado em: 2025-10-13.

Neuroimaging Research for Veterans Center (2025). Tutorials, databases, and more. https://sites.bu.edu/vabhs_neuroimaging/support/analysis-support/other-resources/. Consultado em: 2025-10-13.

NeurOn Neuropsychology Online (2025). Online cognitive testing platform. https://neuropsychology.online/. Consultado em: 2025-10-13.

NeuronUP (2025). Avaliação neuropsicológica em doenças raras: Desafios e estratégias para um diagnóstico preciso. https://neuronup.com/br/noticias-de-estimulacao-cognitiva/doencas-raras/avaliacao-neuropsicologica-em-doencas-raras-desafios-e-estrategias-para-um-diagnostico-preciso/. Consultado em: 2025-10-13.

Oracle (2021a). What is a relational database? (rdbms)? https://www.oracle.com/europe/database/what-is-a-relational-database/. Consultado em: 2025-10-13.

Oracle (2021b). What is nosql? https://www.oracle.com/europe/database/nosql/what-is-nosql/. Consultado em: 2025-10-13.

OVHcloud (2025). Base de dados relacional vs. não relacional. https://www.ovhcloud.com/pt/learn/relational-vs-non-relational-databases/. Consultado em: 2025-10-13.

Pallets Projects (2026). Flask. https://flask.palletsprojects.com/en/stable/. Consultado em: 28 de março de 2026.

Pandey, R. M. (2015). Commonly used t-tests in medical research. *Journal of the Practice of Cardiovascular Sciences*, 1(2):185–188. DOI: https://doi.org/10.4103/2395-5414.166321.

Pannmie (2023). Diving into non-parametric tests: The mann-whitney u test. https://www.kaggle.com/discussions/general/419726. Consultado em: 2025-10--13.

Patryk, S. (2024). Modern web technologies: frameworks, advantages, disadvantages and optimal applications. *Computer Science and Mathematical Modelling*, 19:25–34.

Peterson, K. A., Leddy, A., and Hornberger, M. (2025). Reliability of online, remote neuropsychological assessment in people with and without subjective cognitive decline. *PLOS Digital Health*, 4(4):e0000682.

Physiotutors (2025). Taxa de erros de tipo 1 | estatíticas. https://www.physiotutors.com/pt/wiki/type-1-error-rate-control/. Consultado em: 2025-10-13.

Poznanski, R. R. (2006). Editorial: Outcomes from the "integrative" brain resource international database. *Journal of Integrative Neuroscience*, 05(01):v–vi.

PubMed (2025). About pubmed. https://pubmed.ncbi.nlm.nih.gov/about/. Consultado em: 2025-10-13.

Quest (2025). What is postgresql and how does it compare to other database management systems? https://www.quest.com/learn/what-is-postgresql.aspx. Consultado em: 2025-10-13.

Robinson, A. D. and Finley, J.-C. A. (2025). Utilizing clinical assessment databases to enhance clinical training and research in neuropsychology: benefits, challenges, and practical strategies. *Journal of Clinical and Experimental Neuropsychology*, 0(0):1–10. PMID: 40350573.

Rockholt, M. M., Wu, R. R., Zhu, E., Perez, R., Martinez, H., Hui, J. J., Commeh, E. B., Denoon, R. B., Bruno, G., Saba, B. V., Waren, D., O'Brien, C., Aggarwal, V. K., Rozell, J. C., Furgiuele, D., Macaulay, W., Schwarzkopf, R., Schulze, E. T., Osorio, R. S., Doan, L. V., and Wang, J. (2025). Application of the uniform data set version 3 tele-adapted test battery (t-cog) for remote cognitive assessment preoperatively in older adults. *Frontiers in Aging Neuroscience*, Volume 16 - 2024.

Sahoo, S. and Grover, S. (2022). Technology-based neurocognitive assessment of the elderly: a mini review. *Consortium Psychiatricum*, 3(1):37–44.

Salkind, N. J. (2007). Factor scores. In *Encyclopedia of Measurement and Statistics*, pages 346–348. Sage Publications, Inc., Thousand Oaks, California. Consultado em: 2026-03-17.

Sbordone, R. J. (2001). Limitations of neuropsychological testing to predict the cognitive and behavioral functioning of persons with brain injury in real-world settings. *NeuroRehabilitation*, 16(4):199–201. PMID: 11790904.

Scarpina, F. and Tagini, S. (2017). The stroop color and word test. *Frontiers in Psychology*, Volume 8 - 2017.

Schaefer, L. A., Thakur, T., and Meager, M. R. (2023). Neuropsychological assessment. In StatPearls Publishing, editor, *StatPearls [Internet]*. StatPearls Publishing, Treasure Island (FL). Updated 2023 May 16. Available from: https://www.ncbi.nlm.nih.gov/books/NBK513310/. Consultado em: 2025-06-26.

Schretlen, D. J., Testa, S. M., and Pearlson, G. D. (2025). Calibrated neuropsychological normative system [cnns]. https://paa.com.au/product/cnns/?v=f77d8ed2dc48. Consultado em: 2025-10-13.

Scolary (2019). Pubpsych. https://scolary.com/tools/pubpsych. Consultado em: 2025-10-13.

Shadowfourati (2023). The use of neuropsychological assessment in legal and forensic settings. https://www.reflectneuro.com/the-use-of-neuropsychological-assessment-in-legal-and-forensic-settings/. Consultado em: 2025-10--13.

Shirk, S. D., Mitchell, M. B., Shaughnessy, L. W., Sherman, J. C., Locascio, J. J., Weintraub, S., and Atri, A. (2011). A web-based normative calculator for the uniform data set (uds) neuropsychological test battery. *Alzheimer's Research & Therapy*, 3(6):32.

Singh, M., Verma, A., Parasher, A., Chauhan, N., and Budhiraja, G. (2019). Implementation of database using python flask framework. *International Journal of Engineering and Computer Science*, 8(12):24890–24893.

Spreen, O. and Risser, A. H. (1998). 4 - assessment of aphasia. In Taylor Sarno, M., editor, *Acquired Aphasia (Third Edition)*, pages 71–156. Academic Press, San Diego, third edition edition.

Stephanie Glen (2025a). Hotellings t-squared: Simple definition. https://www.statisticshowto.com/hotellings-t-squared/. Consultado em: 2025-10-13.

Stephanie Glen (2025b). Mahalanobis distance: Simple definition, examples. https://www.statisticshowto.com/mahalanobis-distance/#:~:text=A%20different%20version%20of%20the%20formula%20uses,%5Bxi%20%E2%80%93%20x%CC%84)T%20C%2D1(xi%20%E2%80%93%20x%CC%84)%5D0.5.%20Where:. Consultado em: 2025-10-13.

Sulastri, A., Utami, M., Jongsma, M., and van Luijtelaar, G. (2018). The indonesian boston naming test: Normative data among healthy adults and effects of age and education on naming ability. *International Journal of Science and Research (IJSR)*, 8:134–139.

testbook (2025). Anova analysis | one-way, two-way, steps, uses, examples. https://testbook.com/ugc-net-commerce/anova-analysis-of-variance. Consultado em: 2025-10-13.

Thayer, G. and Robokos, D. D. (2025). Interpreting results from a neuropsychological testing report: An introductory guide. https://childtherapysupport.com/blog/interpreting-results-from-a-neuropsychological-testing-report-an-introductory-guide. Consultado em: 2025-06-26.

The PostgreSQL Global Development Group (2026). Postgresql: The world's most advanced open source relational database. https://www.postgresql.org/. Consultado em: 28 de março de 2026.

Timmerman, M., Voncken, L., and Albers, C. (2019). A tutorial on regression-based norming of psychological tests with gamlss. *OfPrePrints*.

Tripathi, R., Kumar, K., Bharath, S., Marimuthu, P., and Varghese, M. (2014). Age, education and gender effects on neuropsychological functions in healthy indian older adults. *Dementia & Neuropsychologia*, 8(2):148–154.

Tsiakiri, A., Christidi, F., Tsiptsios, D., Vlotinou, P., Kitmeridou, S., Bebeletsi, P., Kokkotis, C., Serdari, A., Tsamakis, K., Aggelousis, N., and Vadikolias, K. (2024). Processing speed and attentional shift/mental flexibility in patients with stroke: A comprehensive review on the trail making test in stroke studies. *Neurology International*, 16(1):210–225.

Urban, J., Scherrer, V., Strobel, A., and Preckel, F. (2025). Continuous norming approaches: A systematic review and real data example. *Assessment*, 32(5):654–674. PMID: 39066602.

van Rentergem, J. A. A., Murre, J. M. J., and Huizenga, H. M. (2017). Multivariate normative comparisons using an aggregated database. *PLOS ONE*, 12(3):e0173218.

VistaLifeSciences (2025). Anam faq. https://vistalifesciences.com/anam-faqs#:~:text=ANAM%20%2D%20The%20Automated%20Neuropsychological%20Assessment%20Metrics,ANAM%20is%20available%20exclusively%20from%20Vista%20LifeSciences.%https://www.oit.va.gov/services/trm/ToolPage.aspx?tid=16942. Consultado em: 2025-10-13.

Wadhwa, R. R. and Marappa-Ganeshan, R. (2023). T test. In *StatPearls [Internet]*. StatPearls Publishing. Disponível em: https://www.ncbi.nlm.nih.gov/books/NBK553048/.

Wahyuningrum, S., Sulastri, A., and Sanjaya, R. (2019). Information system databases for neuropsychology tests: case study in boston naming test. *SISFORMA*, 6:28.

Wahyuningrum, S. E., van Luijtelaar, G., and Sulastri, A. (2023). An online platform and a dynamic database for neuropsychological assessment in indonesia. *Applied Neuropsychology: Adult*, 30(3):330–339. PMID: 34256659.

Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N. R., Chui, H., Cummings, J., DeCarli, C., Foster, N. L., Galasko, D., Peskind, E., Dietrich, W., Beekly, D. L., Kukull, W. A., and Morris, J. C. (2009). The alzheimer's disease centers' uniform data set (uds): the neuropsychologic test battery. *Alzheimer Disease and Associated Disorders*, 23(2):91–101.

Wärn, E., Andersson, L., and Berginström, N. (2025). Remote neuropsychological testing as an alternative to traditional methodsa convergent validity study. *Archives of Clinical Neuropsychology*, 40(6):1123–1132.

Zadelaar, J. N., van Rentergem, J. A. A., and Huizenga, H. M. (2017). Univariate comparisons given aggregated normative data. *The Clinical Neuropsychologist*, 31(6-7):1155–1172. PMID: 28679311.

Zhang, X., Lv, L., Min, G., Wang, Q., Zhao, Y., and Li, Y. (2021). Overview of the complex figure test and its clinical application in neuropsychiatric disorders, including copying and recall. *Frontiers in Neurology*, 12:680474.