

Matilde Serra **Fiabilidade intra-avaliador de
uma bateria de testes para
avaliação da aptidão física em
indivíduos adultos saudáveis**

Dissertação de Mestrado em Fisioterapia
Relatório de Projeto de Investigação

Orientadora: Professora Doutora Rita
Fernandes
Coorientador: Professor Diogo Moço

Dezembro 2023

Relatório de Investigação apresentado para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Fisioterapia, área de especialização em Fisioterapia em Condições Músculo-esqueléticas realizada sob a orientação científica da Professora Doutora Rita Fernandes e do Professor Diogo Moço.

[DECLARAÇÕES]

Declaro que este Relatório de Projeto de Investigação é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

O candidato, _____

Local, de de

Declaro que este Relatório de Projeto de Investigação se encontra em condições de ser apresentada a provas públicas.

O(A) orientador(a), _____

Local, de de

Agradecimentos

Quero agradecer aos meus pais, pelo suporte e amor incondicional,

À minha avó pela motivação e força para a conclusão deste trabalho,

À minha família e amigos,

À minha colega Bruna, pela partilha de conhecimento, angústias e conquistas,

Aos professores Rita Fernandes e Diogo Moço pela orientação ao longo do processo,

A todos os participantes que se voluntariaram e abdicaram do seu tempo para participar neste estudo, e a todos aqueles que de alguma forma contribuíram para a realização deste projeto,

Um enorme obrigado!

RESUMO

Fiabilidade de uma bateria de testes para avaliação da aptidão física em indivíduos adultos saudáveis

Matilde Serra, Diogo Moço, Rita Fernandes

Palavras-Chave: Aptidão Física, Indivíduos Saudáveis, Fiabilidade intra-avaliador, Erro padrão da medida, Diferença mínima detetável

Introdução: As condições músculo-esqueléticas crónicas e persistentes, têm um enorme impacto global e no indivíduo, contribuindo de forma substancial para a incapacidade e morbilidade. Estas têm um impacto persistente na aptidão física dos indivíduos. É de extrema importância, ter instrumentos fiáveis para avaliar as várias componentes envolvidas na aptidão física destes indivíduos de forma a orientar planos de treino e intervenção. A aptidão física é composta por quatro componentes que podem ser avaliadas por testes de campo simples, capacidade cardiorrespiratória: *6-Minute Walk Test*; controlo neuromuscular: *Aberrant Movement Pattern* e *Prone Instability Test*; força muscular: *Trunk Flexor Test*, *Side Bridge*, *Biering Sørensen Test* e *60s Sit-to-Stand*; e flexibilidade: *Sit-and-Reach Test*, *Schober Test* e *Modified Thomas Test*. Para a correta interpretação destes instrumentos é necessário conhecer as suas propriedades psicométricas tanto em populações clínicas, como na população saudável. No entanto, existe carência de dados na literatura, acerca da fiabilidade destes testes em indivíduos saudáveis. Estabelecer conclusões sobre a fiabilidade destes testes, ajudará a interpretação fiável na prática em população clínicas. **Objetivo:** Avaliar a fiabilidade intra-avaliador de uma bateria de testes que avalia as várias componentes da aptidão física na população portuguesa saudável. **Metodologia:** 21 participantes foram submetidos a dois momentos de avaliação separados por 4.45 ± 1.70 dias, onde foi aplicada a bateria de testes em estudo. Para aferir à cerca da fiabilidade dos testes, para os com variável numérica foram calculados os CCI, e respetivos 95% IC, EPM e EPM%, assim como intervalos de LDC, DMD e DMD%; para os testes com variável categórica, foram calculados o coeficiente de k de Cohen, e % de concordância. Forma descritas frequências para todos os testes. **Resultados:** Para os testes *6MWT*, *Trunk Flexor Test*, *Side Bridge*, *Biering Sørensen Test*, *60s Sit to Stand*, *Sit-and-Reach Test*, *Schober Test* e *Schober Test* para a extensão, foram obtidos resultados de fiabilidade excelentes (CCI>0.94) e valores de EPM% baixos (EPM%<10.82%). Enquanto, para os testes *Aberrant Movement Pattern*, *Prone Instability Test* e *Modified Thomas Test* à esquerda fiabilidade moderada ($k>0.61$); e fiabilidade fraca ($k=0.54$) para o *Modified Thomas Test* à direita. Foram obtidos valores de concordância altos para todos os testes de variável categórica (%concordância>80.9%). **Conclusões:** Nos testes que se traduzem em variável numérica, o SRT, SB, STE e STF apresentam fiabilidade excelente e baixo erro de medição associado, enquanto os testes *6MWT*, *BST* e *60STS*, apesar de excelente fiabilidade, apresentam algum erro de medição. O TFT é o teste que apresenta maior erro associado, ainda que fiabilidade excelente. Nos testes que se traduzem em variáveis categóricas, o AMP, PIT e MTT do lado esquerdo apresentam fiabilidade moderada, mas IC amplos e apresentam baixo erro associado. Por fim o MTT do lado direito apresenta fiabilidade fraca e com IC amplo, ainda que o erro relativamente baixo.

ABSTRACT

Reliability of a test battery for assessing physical fitness in healthy adult individuals

Matilde Serra, Diogo Moço, Rita Fernandes

Keywords: Physical Fitness, Healthy Individuals, Intra-rater Reliability, Standard Error of Measurement, Minimal Detectable Difference.

Introduction: Chronic and persistent musculoskeletal conditions have a significant global impact, contributing substantially to disability and morbidity. These conditions have a persistent effect on the physical fitness of individuals. It is of utmost importance to have reliable instruments to assess various components involved in the physical fitness of these individuals to guide training and intervention plans. Physical fitness consists of four components that can be assessed by simple field tests: cardiorespiratory capacity (6-Minute Walk Test); neuromuscular control (Aberrant Movement Pattern and Prone Instability Test); muscular strength (Trunk Flexor Test, Side Bridge, Biering Sørensen Test, and 60s Sit-to-Stand); and flexibility (Sit-and-Reach Test, Schober Test, and Modified Thomas Test). For the correct interpretation of these instruments, knowledge of their psychometric properties is necessary in both clinical and healthy populations. However, there is a lack of data in literature regarding the reliability of these tests in healthy individuals. Establishing conclusions about the reliability of these tests will aid reliable interpretation in clinical populations.

Purpose: To assess the intra-rater reliability of a battery of tests evaluating various components of physical fitness in the healthy Portuguese population. **Methods:** 21 participants underwent two separate assessment sessions spaced 4.45 ± 1.70 days apart, where the study test battery was applied. To assess the reliability of the tests with numerical variables, ICC, and their 95% CI, SEM, and SEM% were calculated, as well as LDC, MDD, and MDD%; for tests with categorical variables, Cohen's k coefficient and the percentage of agreement were calculated. Frequencies were described for all tests. **Results:** For the 6MWT, Trunk Flexor Test, Side Bridge, Biering Sørensen Test, 60s Sit to Stand, Sit-and-Reach Test, Schober Test, and Schober Test for extension, excellent reliability results ($CCI > 0.94$) and low EPM% values ($EPM\% < 10.82\%$) were obtained. Meanwhile, for the Aberrant Movement Pattern, Prone Instability Test, and Modified Thomas Test on the left, moderate reliability ($k > 0.61$) was observed, and weak reliability ($k = 0.54$) was observed for the Modified Thomas Test on the right. High agreement values were obtained for all categorical variable tests ($\%agreement > 80.9\%$). **Conclusions:** In tests resulting in numerical variables, SRT, SB, STE, and STF exhibit excellent reliability and low associated measurement error, while the 6MWT, BST, and 60STS, despite excellent reliability, show some measurement error. The TFT is the test with the highest associated error, despite excellent reliability. In tests resulting in categorical variables, the AMP, PIT, and MTT on the left show moderate reliability, but with broad CI and low associated error. Finally, the MTT on the right shows weak reliability with a broad CI, although relatively low error.

Índice

1. Introdução.....	1
2. Metodologia	9
2.1 Desenho do Estudo.....	9
2.2 Cálculo da Amostra.....	9
2.3 Participantes	9
2.4 Recrutamento dos Participantes	9
2.5 Procedimentos	10
2.6 Outcomes e Instrumentos de Avaliação	11
2.6.1 Questionário Sociodemográfico e Ocupacional	11
2.6.2 Questionário Internacional de Atividade Física – IPAQ Versão Curta	11
2.6.3 Six-Minute Walk Test	12
2.6.4 Aberrant Movement Pattern.....	12
2.6.5 Prone Instability Test.....	13
2.6.6 Trunk Flexor Test.....	14
2.6.7 Side Bridge.....	14
2.6.8 Biering Sørensen Test	15
2.6.9 60s Sit to Stand.....	15
2.6.10 Sit-and-Reach Test	16
2.6.11 Schober Test.....	17
2.6.12 Modified Thomas Test.....	17
2.7 Processamento e Análise de Dados.....	18
3. Resultados	20
3.1 Caracterização da Amostra.....	20
3.2 Fiabilidade e Erro de Medição	22
4. Discussão	28
5. Conclusão.....	37
Bibliografia	38
APÊNDICES	46
Apêndice 1 – Panfleto Redes Sociais	47
Apêndice 2 – Formulário Online	48
Apêndice 3 – Carta Explicativa do Estudo	49
Apêndice 4 – Consentimento Informado	54
Apêndice 5 – Registo de Avaliações.	56
ANEXOS	58

Anexo 1 – Parecer da Comissão Especializada de Ética em Investigação.....	59
Anexo 2 – Questionário Sociodemográfico e Ocupacional.....	62
Anexo 3 – IPAQ Versão Curta.....	66

Lista de Abreviaturas

60STS – *60s Sit-to-Stand*

6MWT – *Six-Minute Walk Test*

AMP – *Aberrant Movement Pattern*

BST – *Biering Sørensen Test*

CCI – Coeficiente de Correlação Interclasse

CR – Coeficiente de Repetibilidade

DMD – Diferença Mínima Detetável

DPdif – Desvio Padrão das Diferenças

EPM – Erro Padrão de Medida

IPAQ – Questionário Internacional de Atividade Física

LDC – Limites de Concordância

MTT – *Modified Thomas Test*

PIT – *Prone Instability Test*

Q1 – Primeiro Quartil

Q2 – Segundo Quartil

Q3 – Terceiro Quartil

SB – *Side Bridge*

SBD – *Side Bridge* direita

SBE – *Side Bridge* esquerda

SRT – *Sit-and-Reach Test*

ST – *Schober Test*

STE – *Schober Test* Extensão

STF – *Schober Test* Flexão

TFT – *Trunk Flexor Test*

1. Introdução

As condições músculo-esqueléticas crónicas representam uma das maiores causas de incapacidade em todo o mundo. O aumento da esperança de vida e mudanças no estilo de vida, têm levado a um aumento da sua prevalência a nível mundial (March et al., 2014). Estas condições têm um impacto significativo no mundo e sistemas de saúde em geral, contribuindo de forma substancial para a incapacidade e morbilidade mundialmente (Sebbag et al., 2019).

Estas condições impactam significativamente os indivíduos, a sua qualidade de vida e bem-estar geral (Sebbag et al., 2019). Podem representar grandes impactos na funcionalidade, afetando as atividades da vida diária e participação na sociedade, para além de aumentarem o risco de comorbilidades (Sebbag et al., 2019; Briggs et al., 2018). Este impacto individual pode também ter grandes repercussões económicas, já que resulta em perdas de produtividade e aumento dos custos de saúde (Sebbag et al., 2019; Dagenais et al., 2008). Existe um paradigma similar na população portuguesa, sendo que este tipo de patologias constitui uma das principais razões de procura de cuidados médicos apresentando um custo significativo para os sistemas de saúde (DGS, 2018).

Um dos sintomas mais significativos associados a várias condições músculo-esqueléticas crónicas é a dor, que afeta cerca de 20% da população mundial (Blyth et al., 2019; Rice et al., 2019). Esta relação é especialmente notória ao observar-se que grande parte da população que sofre de dor crónica, apresenta igualmente condições músculo-esqueléticas como a principal causa dessa mesma dor (Vos et al., 2020). A dor crónica pode ter um impacto profundo e duradouro no bem-estar físico, emocional e social de um indivíduo (Moseley & Vlaeyen, 2015).

Para uma melhor compreensão deste impacto, é importante abordar outro aspeto comum neste tipo de condições. Após o término de um episódio sintomático, os indivíduos podem vir a experienciar uma recorrência ao longo do tempo. Para além do impacto individual e do ciclo de cronicidade associado às recorrências, estas impactam a gestão contínua das condições músculo-esqueléticas (Williams et al., 2018).

A dor e o movimento são fenómenos fundamentais para a vida do ser humano, é por isso de extrema importância perceber como funcionam e como interagem entre si (Merkle et al., 2020). Estes estão intrinsecamente relacionados, já que a dor pode causar várias adaptações neuromusculares e, por outro lado, o movimento é muitas vezes utilizado de forma a diminuir/minimizar a dor e restaurar a função (Merkle et al., 2020).

O indivíduo e a maneira como este funciona, altera-se na presença de dor, desde mudanças subtis na coordenação muscular até uma limitação completa da função. No entanto, os

mecanismos subjacentes a essas alterações estão longe de serem totalmente compreendidos (Hodges et al., 2015). Estas alterações, principalmente a longo prazo, podem resultar em grandes impactos na aptidão física do indivíduo, como diminuição da força muscular, flexibilidade, capacidade cardiorrespiratória, entre outras, perpetuando as limitações físicas (Gatchel et al. 2007; Rice et al., 2019). Alterações estas que também estão relacionadas com a capacidade de realizar atividades da vida diária, com o funcionamento geral do indivíduo e com a sua qualidade de vida (Nijs et al., 2010). Os efeitos a longo prazo da dor crónica na aptidão física dos indivíduos, tornam-se especialmente preocupantes, já que estes podem persistir mesmo após o desaparecimento dos sintomas (Nijs et al., 2010; Gatchel et al. 2007).

Tendo em conta o impacto prolongado da dor crónica na aptidão física, é fundamental que esta seja avaliada em indivíduos com dor crónica associada a uma condição músculo-esquelética, assim como em indivíduos que ultrapassaram um episódio recente (Geneen et al., 2017). Desta forma, as alterações na aptidão física são identificadas, orientando desta forma programas de prevenção de episódios futuros/recorrência adaptados aos utentes.

De forma a avaliar a aptidão física de um indivíduo, é necessário considerar as respetivas componentes: capacidade cardiorrespiratória, controlo neuromuscular, força muscular e flexibilidade (ACSM, 2018), que podem estar alteradas na presença de situações de dor/patologia de natureza músculo-esqueléticas, em particular as que são crónicas/recorrentes (Nijs et al., 2010). Cada uma destas componentes pode ser avaliada através de testes físicos. No entanto, para uma correta avaliação, é necessário que estes testes sejam válidos e fiáveis e que as suas propriedades psicométricas tenham sido estudadas nas populações clínicas em questão. A avaliação destas componentes requer uma posterior interpretação dos resultados, de forma a garantir uma perspetiva real sobre a aptidão física do indivíduo. Esta interpretação implica a comparação com a população adulta saudável, e por essa razão é relevante conhecerem-se valores de referência para a mesma, assim como as propriedades psicométricas dos testes que permitem a avaliação das referidas componentes de aptidão física. No entanto, a literatura atual apresenta fragilidades no que toca às propriedades psicométricas destes testes, quando considerada esta população.

A qualidade de um instrumento de medição é descrita através de três domínios: fiabilidade, validade e capacidade de resposta. A fiabilidade refere-se ao grau em que o instrumento está isento de erro de medição e até que ponto, as medições são as mesmas, para o mesmo indivíduo, no mesmo estado, mas em condições diferentes. Um instrumento fiável permite distinguir as características de cada indivíduo, apesar do erro associado à medida (de Vet et al., 2006a).

Dentro da fiabilidade existem três tipos de propriedades: fiabilidade, erro de medição e consistência interna (Mokkink et al., 2020; Scholtes et al., 2011). Como referido anteriormente, a fiabilidade estuda a aplicação do instrumento em condições diferentes: o indivíduo pode ser avaliado usando diferentes conjuntos de itens do mesmo instrumento (consistência interna); ao longo do tempo (fiabilidade teste-reteste); por diferentes avaliadores na mesma ocasião (fiabilidade inter-avaliador) ou pela mesma pessoa em ocasiões diferentes (fiabilidade intra-avaliador) (Scholtes et al., 2011). Para o seu cálculo, são utilizados coeficientes ou medidas de correlação. Relativamente à fiabilidade teste-reteste e inter e intra-avaliador, para variáveis categóricas é geralmente calculado o Coeficiente Kappa, enquanto para os variáveis contínuas é preferencialmente utilizado o Coeficiente de Correlação Interclasse (CCI), no entanto, ambos medem a homogeneidade das medidas (Mokkink et al., 2020). Para o Coeficiente de Kappa, valores entre 0.40 e 0.59 indicam fiabilidade fraca; entre 0.60 e 0.79 fiabilidade moderada, entre 0.80 e 0.90 fiabilidade forte; e valores acima de 0.90 indicam fiabilidade excelente (McHugh, 2012). Para o CCI, valores inferiores a 0.5 são indicativos de fiabilidade fraca; valores entre 0.5 e 0.75 de fiabilidade moderada; entre 0.75 e 0.9 fiabilidade boa e para valores superiores a 0.90 fiabilidade excelente (Koo & Li, 2016).

O erro de medição refere-se ao erro sistemático e aleatório associado à pontuação de um indivíduo quando executa um dado teste, que não é atribuído a verdadeiras alterações no domínio avaliado. Avalia assim quão próximos são os resultados de medições repetidas, através da estimativa do erro de medição (Mokkink et al., 2020). Esta não é uma medida relativa, por esta razão a variação entre sujeitos não afeta a medição, e a medição é realizada na escala própria do instrumento (Scholtes et al., 2011). A estatística preferencialmente utilizada para dados contínuos, para este domínio, é o Erro Padrão de Medida (EPM), representado pelo desvio padrão das medidas repetidas de um indivíduo e pelos Limites de Concordância (LDC) (Mokkink et al., 2020). Quanto menores estas medidas, mais capaz será o instrumento de distinguir alterações entre medidas repetidas, sendo assim maior a sua concordância e menor o erro associado (de Vet et al., 2006a).

Para além disto, é geralmente utilizada a Diferença Mínima Detetável (DMD), que representa a mudança mínima que deve ser ultrapassada de forma a garantir uma mudança real, ou seja, é a alteração mínima necessária para ter 95% de certeza de que a diferença entre o valor de uma medição inicial e a seguinte, realizada no mesmo indivíduo, não se deve ao erro de medição (de Vet et al., 2006a). Um instrumento de avaliação é mais fiável quando o seu EPM é inferior à DMD. Ao adicionar-se o cálculo da diferença mínima clinicamente importante, estas

estatísticas são também fundamentais para perceber mudanças clinicamente relevantes nos indivíduos (Scholtes et al., 2011). Para dados categóricos é utilizada a percentagem de concordância (Mokkink et al., 2020).

Por fim, a consistência interna difere das anteriores já que não envolve administração repetida do instrumento. Esta é obtida ao avaliar o grau de relação entre os itens individuais do instrumento. Esta estatística estabelece a correlação entre as respostas do indivíduo aos diferentes itens e sugere se esses medem o mesmo conceito (Scholtes et al., 2011). O conhecimento destas propriedades psicométricas, aquando do uso de instrumentos de medida, é fundamental para o seu bom uso e confiança nos resultados e na sua interpretação.

Como referido anteriormente, de forma a avaliar a aptidão física é necessário considerar as suas componentes: capacidade cardiorrespiratória, controlo neuromuscular, força muscular e flexibilidade (ACSM, 2018), que podem estar alteradas na presença de dor e/ou patologia de natureza músculo-esqueléticas, em particular nas condições crónicas ou persistentes (Nijs et al., 2010). A capacidade cardiorrespiratória parece estar alterada em indivíduos com condições músculo-esqueléticas, como a lombalgia, osteoartrose do joelho e anca, entre outros (Moseng et al., 2014). Para avaliar de forma simples e prática esta componente, são geralmente utilizados testes submáximos (Enright & Sherril, 1998). O *Six-Minute Walk Test* (6MWT) é um teste submáximo frequentemente utilizado, que avalia a capacidade cardiorrespiratória dos indivíduos (Enright & Sherril, 1998; Casanova et al., 2011). Este consiste em o indivíduo caminhar, o mais rápido possível num percurso previamente definido e standardizado durante 6 minutos, sendo medida a distância percorrida, que é utilizada para estimar a capacidade do indivíduo (ATS, 2002; Enright & Sherrill, 1998).

Este teste está validado para várias populações clínicas, assim como para a população saudável. Existem vários estudos relativos a valores e equações de referência para várias populações, no entanto, estas demonstram grande variabilidade no que se refere à sua capacidade preditiva (Casanova et al., 2011; Tveter et al., 2014; Enright & Sherril, 1998; Troosters et al., 1999; Gibbons et al., 2001; Camarri et al., 2006; Chetta et al., 2006). No que toca à população portuguesa existem dois estudos que descrevem valores de referência e proposta de equações preditivas para a população adulta saudável (Oliveira et al., 2019; Marques et al., 2020).

A fiabilidade teste-reteste do 6MWT foi estudada em populações com diversas condições músculo-esqueléticas, demonstrando boa fiabilidade (Tveter et al., 2014). O estudo de Arcuri et al. (2016) para a população saudável apresentou valores semelhantes, demonstrando uma boa fiabilidade teste-reteste (Koo & Li, 2016) (CCI=0.86, 95%IC 0.80 a 0.90). Apesar disto, no que toca

ao estudo das propriedades psicométricas deste teste para a população adulta saudável portuguesa, existem limitações na literatura, nomeadamente ausência de valores de fiabilidade inter-avaliador e de EPM e DMD.

A dor músculo-esquelética, também parece impactar negativamente o controlo neuromuscular dos indivíduos, sendo este o caso para variadas condições músculo-esqueléticas, nomeadamente na lombalgia (Sterling et al., 2001). Para a avaliação deste domínio em indivíduos com história de lombalgia, existem dois testes que podem ser utilizados, o *Aberrant Movement Patterns Test* (AMP) e o *Prone Instability Test* (PIT) (Hicks et al., 2005). O primeiro avalia a qualidade do movimento no movimento de flexão/extensão do tronco e a instabilidade segmentar lombar (Hicks et al., 2005). A fiabilidade inter-avaliador deste teste foi estudada em indivíduos com lombalgia crónica, apresentando, de uma forma geral, fiabilidade moderada (McHugh, 2012) (Rabin et al. (2013): $k=0.64$, 95%IC 0.32 a 0.90; Biely et al. (2014): $k=0.65$, 95%IC 0 a 1; Alyazedi et al. (2015): $k=0.79$, 95%IC 0.39 a 1.19; e Hicks et al. (2003): $k=0.60$, 95%IC 0.47 a 0.73). Resultados similares foram encontrados por Biely et al. (2014), para a população saudável ($k=0.72$, 95%IC 0.46 a 0.98). Foi também estudada a fiabilidade para cada componente isolada deste teste, para população clínica e saudável, tendo sido obtidos valores distintos entre estudos, que variaram desde fiabilidade mínima a forte (fiabilidade mínima: k entre 0.21 a 0.39 e fiabilidade forte: k de 0.80 a 0.90 (McHugh, 2012)) (Rabin et al., 2013; Biely et al., 2014; Alyazedi et al., 2015; Hicks et al., 2003). Esta variação pode dever-se ao facto de que em muitos casos as definições e critérios de cada padrão diferem de autor para autor, influenciando os valores de fiabilidade e a sua comparação. Para a população saudável portuguesa, é necessária mais investigação sobre as propriedades psicométricas deste teste, nomeadamente a fiabilidade intra-avaliador.

O PIT é uma ferramenta de diagnóstico clínico, utilizada principalmente para avaliar a capacidade dos músculos do tronco para manter a instabilidade segmentar lombar em indivíduos com lombalgia (Hicks et al., 2005). Na população com lombalgia a fiabilidade inter-avaliador foi estudada por Alyazedi et al. (2015), revelando fiabilidade moderada (McHugh, 2012) ($k=0.71$, 95%IC 0.45 a 0.98). Enquanto Hicks et al. (2003), considerou esta fiabilidade moderada a excelente (McHugh, 2012) (k de 0.74, 95%IC 0.64 a 0.83, a 1.00, 95%IC 1.00 a 1.00). Não são conhecidas para a população saudável adulta as suas propriedades psicométricas.

A força muscular, mais especificamente do tronco, pode estar afetada na presença de dor crónica e condições músculo-esqueléticas, como lombalgia e fibromialgia (Larsson et al., 2015; Kato et al., 2019). Existem vários testes utilizados para avaliar este domínio, nomeadamente o *Trunk Flexor Test* (TFT) para a avaliação da força resistência dos músculos flexores do tronco,

que tem como objetivo manter, pelo maior período possível, uma posição de flexão do tronco a 60° (Mcgill et al., 1999). Para a população saudável foi estudada a fiabilidade deste teste, sendo considerada de boa (Koo & Li, 2016) (CCI=0.71, 95%IC 0.38 a 0.91 (Durall et al., 2012)) a excelente (Koo & Li, 2016) (CCI=0.97 (Mcgill et al., 1999)) No estudo de Evans et al. (2007), foi também investigada a fiabilidade para atletas, tendo os autores concluído que a fiabilidade tanto intra como inter-avaliador é excelente (fiabilidade intra-avaliador: CCI=0.95; fiabilidade inter-avaliador (avaliada em dois momentos para dois avaliadores): CCI=0.97 – 0.98), e EPM, revelando algum erro de medição. No entanto, para a população adulta saudável, é necessária investigação adicional, acerca da fiabilidade inter-avaliador e da DMD.

Com o objetivo de avaliar a força resistência dos músculos flexores laterais do tronco, o *Side Bridge* (SB) é um teste simples e económico, sendo pedido ao indivíduo para manter uma posição de prancha lateral o maior tempo possível (Anderson et al., 2014; McGill et al., 1999). A sua fiabilidade intra e inter-avaliador e EPM foram estudadas em diversas populações. Para a população com lombalgia não específica a fiabilidade intra-avaliador foi excelente (SB direita (SBD) CCI=0.95, 95%IC 0.91 a 0.97; SB esquerda (SBE) CCI=0.96, 95%IC 0.92 a 0.98) (Ozcan Kahraman et al., 2016); para atletas foi considerada boa (SBD CCI=0.82 95%IC 0.63 a 0.92; SBE CCI=0.85 95%IC 0.69 a 0.93) (Evans et al., 2007); para um grupo de militares foi considerada boa (à esquerda) e excelente (à direita) (SBD CCI=0.91, 95%IC 0.78 a 0.96); SBE CCI=0.82, 95%IC 0.59 a 0.93) (Larsson et al., 2014); enquanto para a população saudável, a fiabilidade foi considerada boa à direita e excelente à esquerda (SBD CCI=0.78 e SBE CCI=0.91) (Greene et al., 2012). É também conhecida a DMD para militares (Larsson et al., 2014). É necessário estudar o valor de DMD deste teste para a população saudável portuguesa.

O *Biering Sørensen Test* (BST) é um teste validado para a avaliação da força resistência dos músculos extensores do tronco (Demoulin et al., 2006). Este consiste em manter o tronco horizontalizado o maior período possível, em decúbito ventral, com os membros inferiores fixos e o tronco fora da marquesa (Martínez-Romero et al., 2020). A fiabilidade teste-reteste deste teste foi estudada em várias populações, sendo considerada excelente em indivíduos com lombalgia (CCI=0.93 (Keller et al., 2001)). Para a população saudável, foi encontrada uma fiabilidade teste-reteste intra-avaliador boa (CCI=0.80 (Keller et al., 2001); CCI=0.78, 95%IC 0.62 a 0.88 (Juan-Recio et al., 2018); Martínez-Romero et al., 2020) e fiabilidade inter-avaliador excelente (CCI=0.94, 95%IC 0.84 a 0.98 (Martínez-Romero et al., 2020)). Para este teste são também conhecidos valores de EPM e de DMD, para população com lombalgia e também para a saudável (Latimer et

al., 1999; Simmonds et al., 1998; Ozcan Kahraman et al., 2016). Não existem conhecimento sobre as suas propriedades psicométricas na população portuguesa saudável.

Por fim, para a avaliação da força resistência da musculatura dos membros inferiores, o *60s Sit-to-Stand* (60STS) é um teste simples que consiste em realizar o movimento de levantar e sentar numa cadeira, durante 60 segundos, o mais rápido possível (Bohannon & Crouch, 2019; Strassmann et al., 2013). A sua fiabilidade teste-reteste, é conhecida para a população saudável com idades compreendidas entre os 55 e 70 anos, sendo considerada boa (CCI=0.80, 95%IC 0.65 a 0.89) (Ritchie et al., 2005). É também conhecida para indivíduos a realizar hemodialise (CCI=0.97, 95%IC 0.94 a 0.98) (Segura-Ortí et al., 2011) e indivíduos com fibrose quística (CCI=0.98, 95% IC 0.96 a 0.99) (Radtke et al., 2016), sendo considerada excelente para ambas as populações (Koo & Li, 2016). Apesar disto, não existe qualquer informação no que toca à fiabilidade inter-avaliador, EPM ou DMD na população adulta saudável. Quando pesquisados na literatura existente, podemos encontrar valores de referência para a população saudável (Strassmann et al., 2013), e para a população saudável portuguesa com mais de 60 anos (Quiterio et al., 2022).

Relativamente à flexibilidade, existem vários testes que podem ser utilizados para a sua avaliação, sendo o *Sit-and-Reach Test*, *Schober Test* e *Modified Thomas Test* frequentemente utilizados. O *Sit-and-Reach Test* (SRT), é um teste comumente utilizado e validado para a avaliação da flexibilidade dos músculos isquiotibiais (Gabbe et al., 2004; Mayorga-Vega et al., 2014). Este é um teste de aplicação simples, que requer que o indivíduo, na posição sentada com as pernas estiradas, estenda os braços à frente, tocando com as mãos juntas o mais longe possível, em direção a uma caixa com uma régua integrada. A maior distância é assinalada, avaliando desta forma a flexibilidade do indivíduo (Bozic et al., 2010). A fiabilidade inter-avaliador deste teste foi estudada na população adulta saudável, sendo considerada excelente (CCI=0.97, 95%IC 0.91 a 0.99 (Gabbe et al., 2004); CCI=0.98, 95%IC 0.96 a 0.99 (Ozcan Kahraman et al., 2016)) sendo também conhecidos valores de EPM, que indicam baixo erro de medição (Gabbe et al., 2004). Existe também conhecimento da DMD e EPM para a população de jovens adultos (Liu et al., 2022). É necessário estabelecer valores de DMD para este teste na população adulta saudável.

O *Schober Test* para flexão (STF), é utilizado para avaliar a amplitude de movimento da coluna lombar em flexão. Foi estudada a sua fiabilidade inter-avaliador para a população adulta saudável, sendo considerada boa (Coeficiente de fiabilidade de *Pearson* STF: $r=1.0$) (Fitzgerald et al., 1983); e também para a população com espondilite anquilosante ($r = 0.94$) (Rezvani et al.,

2012). No que toca ao *Schober Test* para extensão (STE) apenas é conhecida a sua fiabilidade inter-observador para a população adulta saudável, sendo considerada boa (Coeficiente de fiabilidade de *Pearson STE*: $r=0.88$) (Fitzgerald et al., 1983). Não são conhecidos na literatura, para nenhum dos dois testes, valores relativos ao erro de medição (EPM, DMD), assim como também não é conhecida a fiabilidade intra-avaliador.

O *Modified Thomas Test* (MTT), avalia a flexibilidade dos músculos flexores da anca através do uso de um goniómetro ou inclinómetro, apresentando fiabilidades semelhantes entre ambos os instrumentos (Clapis et al., 2008). Este teste tem sido amplamente estudado, no entanto para além da existência de algumas incoerências acerca da sua fiabilidade, os estudos são maioritariamente em atletas. Para esta população, este teste apresentou fiabilidade intra e inter-avaliador boa a excelente (k de Fleiss entre 0.95 e 1 (Cady et al., 2022); CCI=0.87 a 0.91 (Cejudo et al., 2015)) e são também conhecidos valores de EPM e DMD (Cady et al., 2022; Cejudo et al., 2015). Para a população adulta saudável, a fiabilidade inter-avaliador foi considerada excelente (Koo & Li, 2016) (CCI=0.92, 95%IC 0.79 a 0.97) enquanto a fiabilidade teste-reteste inter-avaliador foi considerada moderada 95 (Koo & Li, 2016) (CCI=0.63, 95%IC 0.20 a 0.86, a 0.75, 95%IC 0.41 a 0.95) (Gabbe et al., 2004). São também conhecidos valores de EPM para a população saudável (Gabbe et al., 2004). Assim, é necessária investigação acerca da fiabilidade intra-avaliador e também um maior estudo do erro de medição do instrumento, através da determinação da DMD, na população saudável.

Concluindo, devido à prevalência de condições musculoesqueléticas crónicas/persistentes e ao respetivo impacto na aptidão física dos utentes, existe necessidade de incluir esta componente na avaliação. De forma a interpretar corretamente os resultados dos testes que avaliam a aptidão física em diferentes populações clínicas de natureza musculoesquelética, é necessário conhecer as propriedades psicométricas desses mesmos testes. Desta forma, o objetivo deste estudo é determinar a fiabilidade intra-avaliador de uma bateria de testes para avaliação da aptidão física em indivíduos saudáveis.

2. Metodologia

2.1 Desenho do Estudo

Foi realizado um estudo prospetivo, do tipo teste-reteste, com um único avaliador.

2.2 Cálculo da Amostra

Para o cálculo do tamanho da amostra foi considerado o CCI como variável contínua para cada teste, com intervalo de confiança de 95%. Foi considerado 0.90 como valor ideal de CCI e 0.70 como valor mínimo (Keszei et al., 2010). Foi igualmente considerado um erro tipo II de 20%, o que implica uma potência do teste de 80%. De acordo com a fórmula descrita por Kraemer & Thiemann (1987), foi determinado um número mínimo de 18 participantes, para rejeitar a hipótese nula, que sugere um CCI entre testes, inferior a 0.70. Para colmatar possíveis desistências por parte dos participantes, foram considerados mais 3 participantes, perfazendo dessa forma uma amostra total de 21 participantes.

2.3 Participantes

Foi recrutada uma amostra não probabilística por conveniência de 21 indivíduos saudáveis (12 mulheres e 9 homens) a partir da comunidade e através de meios de comunicação digital.

Os participantes foram selecionados com base nos critérios de inclusão e exclusão pré-definidos. Foram considerados elegíveis para o estudo indivíduos saudáveis, com idades entre os 18 e 65 anos, que soubessem ler e falar em português e com telemóvel capaz de receber e enviar mensagens de texto. Foram excluídos indivíduos com diagnóstico médico de qualquer condição clínica (músculo-esquelética, neurológica, pulmonar, sistémica ou cardíaca) (Gabbe et al., 2004; Arcuri et al., 2016) ou presença de sintomas que pudessem afetar a realização da bateria de testes em estudo ou qualquer atividade da vida diária (Anderson et al., 2014), bem como a presença de outras condições de saúde (osteoporose, fratura, patologia inflamatória, radiculopatia, infeção acompanhada por febre e cancro), assim como mulheres grávidas (Anderson et al., 2014).

2.4 Recrutamento dos Participantes

O recrutamento foi realizado a partir da comunidade residente na região metropolitana de Lisboa e através da divulgação através de meios digitais. Foi publicado um panfleto de apresentação do estudo (Apêndice 1 – Panfleto Redes Sociais) nas redes sociais *Instagram* e *Facebook* dos investigadores. Os indivíduos interessados puderam aceder, através do *link* fornecido no panfleto, a um formulário com uma breve descrição do estudo (Apêndice 2 – Formulário Online).

Os indivíduos interessados partilharam o seu contacto telefónico e endereço de e-mail, de forma a serem contactados. Aquando do contacto por via telefónica, foi-lhes explicado o objetivo,

todos os procedimentos, as vantagens/riscos do estudo e retiradas eventuais dúvidas relativamente ao mesmo. Para além disso, foi feita uma breve triagem relativamente aos critérios de inclusão e exclusão. Foi clarificado que os indivíduos poderiam, a qualquer momento desistir, sem qualquer prejuízo para os mesmos; e que todas as questões relacionadas com o estudo seriam respondidas de forma clara. Foi explicado aos participantes que o seu contributo ao longo do estudo seria anónimo e que, em momento algum, seriam recolhidos dados que os identificassem, assegurando desta forma a confidencialidade. Durante o contacto foram também marcadas as duas sessões de avaliação, consoante a disponibilidade de ambas as partes. Por fim, o investigador informou os indivíduos que iriam receber um email com a carta explicativa do estudo (Apêndice 3 – Carta Explicativa do Estudo), de forma a consolidar todas as informações relativas ao mesmo.

A Comissão de Ética do Instituto Politécnico de Setúbal aprovou o estudo (Anexo 1 – Parecer da Comissão Especializada de Ética em Investigação). Todos os participantes assinaram um consentimento livre e informado. Todos os participantes contribuíram de forma anónima e voluntária, todos os dados foram codificados e não foram recolhidos quaisquer dados que pudessem identificar os indivíduos.

2.5 Procedimentos

No dia 27 de Fevereiro de 2023, foi realizado um treino de competências, no Instituto Politécnico de Setúbal, que teve como objetivo instruir o investigador de todos os procedimentos necessários para a aplicação dos questionários e testes da bateria em estudo. Este treino foi conduzido por um investigador experiente na aplicação da referida bateria.

De forma a testar a fiabilidade intra-avaliador, foram realizadas duas sessões de avaliação com cada participante, com cerca de 60 (T0) e 50 (T1) minutos de duração cada. Estas tiveram um intervalo entre si de 4.45 ± 1.70 dias (2 a 7 dias), minimizando assim o efeito de aprendizagem e evitando alterações no desempenho motor dos indivíduos (McDermott et al., 2010). Ambas as sessões de avaliação foram realizadas no *Work in Studio*, localizado no centro de Lisboa.

Na primeira sessão de avaliação (T0), foi solicitado aos participantes a assinatura de dois exemplares do formulário de consentimento informado (Apêndice 4 – Consentimento Informado), um dos exemplares para o investigador e outro para o participante. Foram verificados novamente os critérios de inclusão e exclusão. Posteriormente, os participantes foram convidados a responder a dois questionários em formato digital: o Questionário Sociodemográfico e Ocupacional, de forma a caracterizar o indivíduo (Anexo 2 – Questionário Sociodemográfico e Ocupacional); e o

Questionário Internacional de Atividade Física – IPAQ, na sua versão curta (Campaniço, 2016), com o objetivo de avaliar o nível de atividade física (Anexo 3 – IPAQ Versão Curta).

Após o preenchimento dos questionários, foi aplicada a bateria de testes para avaliação da aptidão física em estudo, constituída por onze testes, que foram aplicados pela seguinte ordem: 6 – *Minute Walk Test* (6MWT); *Aberrant Movement Pattern*; *Prone Instability Test*; *Trunk Flexor Test*; *Side Bridge* direita e esquerda; *Biering Sørensen Test*; *60s Sit to Stand*; *Sit-and-Reach Test*; *Schober Test* e *Schober Test* para a extensão; e *Modified Thomas Test*. Após o intervalo entre sessões, previamente referido, foi realizado o segundo momento de avaliação (T1), onde foi questionado aos participantes se tinha existido alguma alteração desde T0, que pudesse influenciar a inclusão no estudo, e foi novamente aplicada a bateria de testes. Os testes foram realizados no mesmo período do dia, nos dois momentos de avaliação. Os resultados dos testes de ambos momentos de avaliação foram registados no formulário de registo da avaliação da aptidão física (Apêndice 5 – Registo de Avaliações).

2.6 Outcomes e Instrumentos de Avaliação

2.6.1 Questionário Sociodemográfico e Ocupacional

Este questionário, tem como objetivo caracterizar o indivíduo de forma abrangente e em múltiplas dimensões. Para isso, inclui dados pessoais como idade, sexo, peso, altura, habilitações literárias, situação profissional e respetivas características e hábitos do indivíduo. Este questionário foi cedido pelo Projeto “MyBack – Efetividade e implementação de um programa de autogestão personalizado para prevenir recorrências e incapacidade e promover a saúde músculo-esquelética em utentes com lombalgia”, financiado por fundos nacionais através da Fundação para a Ciência e a Tecnologia com a referência PTDC/SALSER/7406/2020. (<https://myback.ips.pt/>), tendo sido adaptado à população em estudo (foram retiradas as secções relacionadas com a lombalgia).

Este questionário, em formato digital, foi preenchido pelo participante em T0.

2.6.2 Questionário Internacional de Atividade Física – IPAQ Versão Curta

O Questionário Internacional de Atividade Física foi desenvolvido para avaliar o nível de atividade física em adultos dos 15 aos 69 anos, tendo sido proposto pela Organização Mundial de Saúde em 1998. Baseia-se na perceção individual de cada indivíduo sobre os seus níveis de atividade física (IPAQ, 2021; Campaniço, 2016).

Este questionário avalia a atividade física através de quatro domínios: atividades de lazer, atividades domésticas, atividade física relacionada com o trabalho e relacionada com o transporte (IPAQ, 2021). É um questionário auto preenchido, que na sua forma curta, é constituído por nove

itens, que questionam sobre a atividade física geral do indivíduo. Estes itens estão estruturados de forma a fornecer pontuações individuais de cada atividade avaliada. De forma a calcular a sua pontuação utiliza-se um somatório da duração e frequência de cada tipo de atividade: atividades físicas de intensidade vigorosa e moderada, o tempo de caminhada e o tempo passado na posição sentada (IPAQ, 2021).

A fiabilidade deste questionário foi avaliada em vários países. Relativamente à população portuguesa, foi obtida uma boa fiabilidade teste-reteste ilustrada por um CCI de 0.77 (Craig et al., 2003). No estudo de Campaniço (2016) foi avaliada a validade da versão portuguesa do IPAQ, na população saudável portuguesa. Neste estudo, e para a versão curta do questionário, foi encontrada uma correlação (Coeficiente de Correlação de Spearman $r=0.33$) entre o questionário e a média de impulsos/minuto registados por acelerómetros. Desta forma, os resultados indicam aceitabilidade deste questionário na sua versão curta, para monitorizar a atividade física da população adulta saudável.

Este questionário, em formato digital, foi preenchido pelo participante em T0.

2.6.3 Six-Minute Walk Test

O 6MWT tem como objetivo avaliar a capacidade aeróbia do participante durante a marcha (Casanova et al., 2011; Enright & Sherrill, 1998). A sua fiabilidade teste-reteste foi estudada na população saudável (CCI=0.86, 95%IC 0.80 a 0.90) (Arcuri et al., 2016), no entanto, não são conhecidos os valores de EPM e DMD para esta mesma população.

Este teste foi realizado ao ar livre, num espaço com 30 metros, chão plano, sem curvas e distrações mínimas. Os participantes foram instruídos para caminhar, sem uso de auxiliares de marcha, o mais rápido possível por um período de 6 minutos, sem correr. O percurso foi realizado em forma de “8”, sendo contornado um cone em ambas as extremidades. Os participantes foram informados que poderiam descansar sempre que necessário, no entanto, deveriam recomeçar assim que possível, já que o relógio não iria parar. A distância percorrida foi contabilizada assim que o tempo de teste terminava (ATS, 2002; Enright, 2003).

2.6.4 Aberrant Movement Pattern

O teste AMP avalia a qualidade do movimento na flexão do tronco e instabilidade segmentar lombar (Hicks et al., 2005). O resultado do teste é considerado negativo caso não exista alteração do movimento. Caso esteja presente um padrão de movimento alterado/aberrante, o teste é considerado positivo, indicando instabilidade lombar segmentar. Os padrões podem ser classificados nas seguintes categorias: Arco de dor em flexão – dor durante o movimento e não no limite da amplitude; Arco de dor em extensão – dor no movimento de regresso à posição inicial

e não no limite da amplitude; “*Gower Sign*” – utilização das mãos, de forma sequencial nas coxas, para regressar à posição inicial, diminuindo carga na coluna lombar; “*Instability Catch Sign*” – quando existe qualquer aceleração ou desaceleração repentina do movimento do tronco, presença de movimento que ocorra fora do plano primário de movimento, como flexão lateral ou rotação, ou presença de tremor momentâneo, tremor observado nos músculos paravertebrais ou quando o movimento pedido é interrompido; Alteração do ritmo lombo-pélvico – quando durante a flexão, o movimento da anca é maior que o da coluna lombar durante o primeiro terço do movimento e/ou o movimento da coluna lombar é maior que o movimento da anca durante o último terço do movimento. Também positivo quando, no retorno à posição inicial o movimento da coluna lombar é maior do que o movimento na anca durante o primeiro terço do movimento e/ou o movimento na anca é maior que o movimento da coluna lombar durante o último terço do movimento. O teste também é positivo quando, ao retornar da posição de flexão, o indivíduo realiza flexão dos joelhos e desloca a pélvis anteriormente. No final do teste é registado se este é negativo ou positivo, e no último caso qual dos padrões é observado (Rabin et al., 2013).

A fiabilidade deste teste foi estudada em indivíduos com lombalgia crónica apresentado fiabilidade moderada nos seguintes estudos Rabin et al. (2013) ($k=0.64$, 95%IC 0.32 a 0.90), Biely et al. (2014) ($k=0.65$, 95%IC 0 a 1), Alyazedi et al. (2015) ($k=0.79$, 95%IC 0.39 a 1.19) e Hicks et al. (2003) ($k=0.60$, 95%IC 0.47 a 0.73). Para a população saudável, o estudo de Biely et al. (2014) revelou resultados similares ($k=0.72$, 95%IC 0.46 a 0.98, %concordância=89%). São também conhecidas as fiabilidades associadas à identificação de padrões específicos, apresentando valores distintos entre estudos, possivelmente devido a variações na definição dos padrões, revelando fiabilidade de mínima a forte (Rabin et al., 2013; Biely et al., 2014; Alyazedi et al., 2015; Hicks et al., 2003). Não é conhecida a fiabilidade intra-avaliador para a população saudável.

2.6.5 Prone Instability Test

O *Prone Instability Test* é utilizado principalmente para identificar instabilidade segmentar lombar em indivíduos com lombalgia (Hicks et al., 2005). Para a população com lombalgia é conhecida a sua sensibilidade (0.71, 95%IC 0.51 a 0.83) e especificidade (0.57, 95%IC 0.39 a 0.78) (Fritz et al., 2005). Foi também estudada para esta população, a fiabilidade inter-avaliador por Alyazedi et al. (2015) ($k=0.71$, 95%IC 0.45 a 0.98; %concordância 0.90) e por Hicks et al. (2003) (k de 0.74, 95%IC 0.64 a 0.83, a 1.00, 95%IC 1.00 a 1.00; %concordância 87% - 100%). Não existe informação relativamente à população saudável.

Este teste iniciou-se em decúbito ventral com os membros inferiores fora da marquês e os pés a tocar no chão. O avaliador iniciou o teste ao realizar mobilização acessória das vértebras

lombares (póstero-anterior nos processos espinhosos), questionando existência de sintomatologia álgica durante a mobilização. De seguida foi pedido ao participante para retirar os pés do chão, elevando os membros inferiores (pôde agarrar-se na marquesa, para manter a posição). Nesta posição, é realizada novamente mobilização acessória, nas vértebras que tinham inicialmente despoletado dor. Este teste é positivo quando existe dor na primeira fase, mas desaparece quando o indivíduo eleva os membros inferiores, indicando presença de instabilidade lombo-pélvica. O teste é negativo caso não exista dor em nenhum ponto do teste, ou a dor não tenha melhorado (Ferrari et al., 2015).

2.6.6 Trunk Flexor Test

O TFT tem como objetivo avaliar a força resistência dos músculos flexores do tronco (Mcgill et al., 1999). A sua fiabilidade teste-reteste foi estudada no estudo de Durall et al. (2012) (CCI=0.71, 95%IC 0.38 a 0.91), e no de de McGill et al. (1999) com apenas 5 participantes (CCI=0.97) para a população adulta saudável. No estudo de Evans et al. (2007), foram estudadas as propriedades psicométricas para atletas (fiabilidade intra-avaliador: CCI=0.95, EPM=59.1s; fiabilidade inter-avaliador (avaliada em dois momentos para dois avaliadores): CCI=0.97 – 0.98, EPM=50.2s – 60.2s). Não são conhecidos para a população saudável a fiabilidade inter-avaliador e a DMD.

O teste foi iniciado na posição de sentado, com o tronco apoiado numa estrutura com inclinação de 60 graus (uso de goniómetro recomendado), e as ancas e joelhos fletidos a 90 graus e os pés presos à marquesa. Os braços do participante estavam cruzados sobre o peito. O teste foi iniciado ao puxar a estrutura de apoio, 10 centímetros para trás, deixando o tronco sem qualquer apoio. Foi pedido ao participante que mantivesse a posição o máximo de tempo possível (Mcgill et al., 1999). O tempo foi registado em segundos, terminando quando o tronco perdia a amplitude de 60 graus, ou quando o participante saía da posição de teste. Na versão original os pés são fixados à marquesa através de uma cinta, no entanto, neste estudo a fixação foi feita pelo avaliador que segurou os pés do participante. Esta alteração é válida e fiável quando comparada com o método original (CCI=0.89 – 0.95) (Reiman et al., 2012).

2.6.7 Side Bridge

Este é um teste simples e económico utilizado para avaliar a força de resistência dos músculos flexores laterais do tronco (Anderson et al., 2014; McGill et al., 1999). Para a população adulta saudável é conhecida a fiabilidade teste-reteste (SBD CCI=0.78 e SBE CCI=0.91) (Greene et al., 2012). São também conhecidos valores de EPM para militares (SBD EPM=15s (13.2%) e SBE EPM=18s (16.8%)) (Larsson et al., 2014), e para atletas (SBD EPM=14.9s e SBE EPM=17.4s)

(Evans et al., 2007); e o valor de DMD para militares (SBD DMD=42s e SBE DMD=50s). Para a população saudável não é conhecida a DMD.

Para realizar este teste, o participante iniciou em decúbito lateral, com os membros inferiores em extensão, com o pé de cima posicionado a frente do pé de baixo. Foi pedido que o indivíduo levantasse a bacia do chão mantendo o corpo alinhado. O suporte do peso foi realizado pelo cotovelo e pés, o membro superior não utilizado permaneceu cruzado ao peito com a mão a tocar no ombro contrário. O tempo foi registrado em segundos, e o teste terminou quando a anca infra-lateral do participante tocou no chão, ou quando se perdeu a posição de prancha lateral. Este teste foi realizado para ambos os lados (McGill et al., 1999).

2.6.8 Biering Sørensen Test

O BST é um teste validado para a avaliação da força de resistência dos músculos extensores do tronco (Demoulin et al., 2006). Para a população saudável, a sua fiabilidade teste-reteste foi estudada por Keller et al. (2001) (CCI = 0.80) e por Juan-Recio et al. (2018) (CCI=0.78, 95%IC 0.62 a 0.88). Foi ainda estudada na meta-análise de Martínez-Romero et al. (2020) que reportou fiabilidade inter-avaliador elevada (CCI=0.94, 95%IC 0.84 a 0.98) a partir de evidência classificada como moderada, fiabilidade intra-avaliador (intra-sessão) também elevada (CCI=0.88, 95%IC 0.83 a 0.92) a partir de evidência de qualidade baixa; assim como fiabilidade intra-avaliador elevada (inter-sessão) (CCI=0.88, 95%IC 0.80 a 0.92) a partir de evidência de qualidade moderada. Os valores de EPM e DMD foram também estudados para esta população por Latimer et al. (1999) (EPM=17.4s), Simmonds et al., (1998) (teste-reteste: EPM=14.3s; inter-avaliador: EPM=1.2s) e por Ozcan Kahraman et al. (2015) (EPM=8.71s; DMD=24.14s).

Este teste foi iniciado em decúbito ventral, com os membros inferiores fixos à marquesa pelo avaliador, e o tronco fora da mesma (Reiman et al., 2012). Antes do início do teste, o indivíduo pôde apoiar os seus braços numa cadeira. Posteriormente foi-lhe pedido que horizontalizasse o tronco e alinhasse em relação aos membros inferiores, cruzando os braços sobre o peito, mantendo esta posição o máximo de tempo possível. O tempo foi contabilizado em segundos, até o participante sair da posição de teste (Demoulin et al., 2006).

2.6.9 60s Sit to Stand

O teste *60s Sit-to-Stand*, é um teste utilizado para avaliar a força de resistência da musculatura dos membros inferiores (Strassmann et al., 2013). A sua fiabilidade teste-reteste foi estudada, para uma população saudável com idades compreendidas entre 55 e 70 anos, por Ritchie et al. (2005), demonstrando uma boa fiabilidade (CCI=0.80, 95%IC 0.65 a 0.89). Esta também foi estudada para outras populações, revelando-se excelente (CCI=0.97, 95%IC 0.94 a

0.98) para indivíduos a realizar hemodiálise; e também excelente (CCI= 0.98, 95% IC 0.96 a 0.99) para indivíduos com fibrose quística. Foram também para este teste descritos valores de referência por Strassmann et al. (2013), categorizados em sexo e idade, idades compreendidas entre os 20 e os 79 anos e respetivamente por percentil. Não existem informações acerca da fiabilidade inter-avaliador, EPM ou DMD na população adulta saudável.

Inicialmente o participante encontra-se sentado numa marquesa ou cadeira sem braços com uma altura de 45 a 48 cm. Deve estar sentado na ponta da marquesa/cadeira, com os pés totalmente apoiados no chão, com os braços cruzados ao peito, e pernas afastadas na marquesa/cadeira. Foi então pedido ao participante que realizasse o movimento de levantar e sentar numa cadeira, durante 60 segundos, sem ajuda dos braços, o mais rápido que conseguisse. Sem que fosse necessário descansar, o participante podia fazê-lo, devendo retomas o movimento assim que possível, sem o tempo parar. Foram contabilizadas pelo avaliador todas as repetições realizadas no tempo previsto (Bohannon & Crouch, 2019).

2.6.10 Sit-and-Reach Test

O SRT é um teste comumente utilizado para a avaliação da flexibilidade dos músculos isquiotibiais (Gabbe et al., 2004; Mayorga-Vega et al., 2014). A fiabilidade inter-avaliador e EPM foram estudados por Gabbe et al. (2004), numa população saudável com idade média de 31.6 anos, 95%IC 27.0 a 36.2, sendo considerada excelente (CCI=0.97, 95%IC 0.91 a 0.99; EPM=2cm). Foi também investigada, neste estudo, a fiabilidade teste-reteste com dois avaliadores distintos, sendo o CCI de 0.98, 95%IC 0.94 a 0.99, a 0.99, 95%IC 0.98 a 1.00 (EPM=1cm). Indo de encontro a resultados de outros estudos efetuados como o de Ozcan Kahraman et al. (2016) ou o de Bozic et al. (2010), assim como os da meta-análise de Cuenca-Garcia et al. (2022). O estudo de Liu et al. (2022), apresenta, para jovens adultos saudáveis (22.35 ± 2.80 anos), valores de EPM e DMD (inter-avaliador: EPM=1.23 cm, 95%IC 2.36 a 7.20 e DMD=3.42 cm; teste-reteste: EPM= 1.62 cm, 95%IC 2.07 a 8.44 e DMD=4.50 cm). Não é conhecido o valor da DMD para a população adulta saudável.

Para este teste é necessária a utilização de uma “*Sit-and-Reach Box*” ou similar. Neste caso foi utilizada uma caixa com uma régua posicionada na horizontal, com os 26cm alinhados com o início da caixa. O participante iniciou o teste sentado no chão, com os joelhos em extensão, e as tibiotársicas em posição neutra, com os pés apoiados na caixa. Foi pedido ao participante que, com uma mão em cima da outra, deslizasse, lentamente e o mais longe que conseguisse, as mãos sobre a régua, posicionada em cima da caixa, mantendo a posição durante dois segundos. O

participante teve duas tentativas, sendo contabilizada a melhor, em centímetros. (Bozic et al., 2010; Gabbe et al., 2004).

2.6.11 Schober Test

O STF e o STE são utilizados para avaliar a amplitude de movimento da coluna lombar em flexão e em extensão, respetivamente. A sua fiabilidade inter-avaliador foi estudada para uma população saudável dos 20 aos 82 anos (Coeficiente de fiabilidade de *Pearson* STF: $r=1.0$; STE: $r=0.88$) no estudo de Fitzgerald et al. (1983). No que toca aos estudos relativos a fiabilidade intra-avaliador, apenas é conhecida para o STF para a população com espondilite anquilosante ($r = 0.94$) (Rezvani et al., 2012). Não são conhecidos os valores de DMD, EPM ou fiabilidade intra-avaliador para a população saudável.

O STF iniciou com o participante em posição ortostática e com os pés alinhados com os ombros. Foram realizadas duas marcações na pele com um marcador, a primeira localizada na coluna e alinhada com a porção mais superior da crista ilíaca, palpada pelo investigador. A segunda marcação foi efetuada 10 centímetros acima da primeira. Foi pedido ao participante que realizasse a máxima flexão da coluna lombar. No final do movimento, foi medida e registada a distância entre as duas marcações, esse valor foi posteriormente subtraído por 10, obtendo-se dessa forma o valor final. O teste foi repetido três vezes, sendo registado o valor com a maior amplitude.

Foi também realizado o ST para a extensão, avaliando dessa forma a amplitude de movimento da coluna lombar em extensão. Para este teste foi utilizado um goniómetro. O participante começou em posição ortostática, com os pés posicionados à largura dos ombros. O fulcro do goniómetro foi colocado na porção mais superior da crista ilíaca e alinhado com a linha axilar média. O braço móvel foi orientado com a linha axilar média e o fixo foi posicionado perpendicular ao chão, e o participante realizou a extensão máxima da coluna lombar, sempre com os joelhos em extensão. O movimento foi corrigido sempre que existiram compensações. Foi registado o valor da amplitude em graus.

2.6.12 Modified Thomas Test

Este teste avalia a flexibilidade dos músculos flexores da anca. A sua fiabilidade foi estudada em atletas de rugby, no estudo de Cady et al. (2022), apresentando valores de k de Fleiss entre 0.95 e 1; e no estudo de Cejudo et al. (2015), que avaliou jogadores de andebol e futsal, demonstrando uma boa fiabilidade intra-avaliador, entre 2 pares de sessões (ICC=0.87 a 0.91). Este estudo calculou também o valor do Erro Padrão da Média (1.3°), e a Diferença Mínima Detetável (4.7° 95%IC). Para a população saudável, com uma idade média de 31.6 anos, 95%IC

27.0 a 37.2, o estudo de Gabbe et al. (2004) concluiu que, a fiabilidade inter-avaliador para este teste é de CCI=0.92, 95%IC 0.79 a 0.97 e EPM de 3°; e uma fiabilidade teste reteste de CCI=0.63, 95%IC 0.20 a 0.86, a 0.75, 95%IC 0.41 a 0.95 e EPM de 5°.

O participante iniciou o teste na posição de sentado, na ponta da marquesa, levando um joelho em direção ao peito e agarrando a respetiva coxa. Deitou-se lentamente sobre a marquesa, na posição de decúbito dorsal. A coxa do membro contralateral não estava suportada, e a coluna esteve sempre em contacto com a marquesa, o avaliador palpou a região lombar de forma a perceber se existia lordose. Na posição final de teste, o avaliador mede a amplitude da anca com auxílio de um goniómetro, colocando o fulcro no grande trocânter, braço fixo alinhado com a linha média da pélvis e braço livre lateral ao fémur (Clapis et al., 2008). O teste foi considerado negativo caso o ângulo fosse 0 graus, ou seja, o participante conseguia manter a coxa horizontalizada (Harvey, 1998; Wakefield et al., 2015).

2.7 Processamento e Análise de Dados

Após os momentos de T0 e T1 de todos os participantes, os dados foram introduzidos no programa IBM SPSS *Statistics* ©, versão 29, contendo apenas o código de cada participante, sem qualquer informação que os pudesse identificar. Os registos em papel foram arquivados num local seguro ao encargo da equipa de investigadores.

De forma a caracterizar a população foi utilizada estatística descritiva: foram calculadas as medidas de tendência central (média) e de dispersão (desvio padrão e percentis) para as variáveis numéricas, assim como a distribuição (%) para as variáveis nominais e ordinais (variáveis qualitativas). De forma a calcular a fiabilidade teste-reteste, no que toca às variáveis numéricas – 6MWT, TFT, SB, BST, 60STS, SRT, STF e STE – foi calculado o CCI, com um Intervalo de Confiança (IC) de 95%. Para tal, e devido às características do estudo, foi utilizado o modelo estatístico *two-way mixed effects*, isto é, CCI (3,1) (Shrout & Fleiss, 1979; de Vet et al., 2006). Como é pretendido o cálculo da fiabilidade intra-avaliador, foi utilizado o CCI consistência, já que este desconsidera a variação entre avaliadores (de Vet et al., 2006). Foi considerado um valor mínimo aceitável de CCI de 0.70. Desta forma o CCI foi calculado através da seguinte fórmula (de Vet et al., 2006a):

$$CCI_{consistência} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{residual}^2}$$

Nesta fórmula, σ corresponde à variância e é um termo estatístico que diz respeito à variabilidade. σ_p corresponde à variância dos participantes em estudo e $\sigma_{residual}$ corresponde à variância presente na interação entre participantes e investigador (de Vet et al., 2006).

De forma a avaliar o erro de medição de cada teste, foi calculado, através do programa *Microsoft Excel*, o EPM com a seguinte fórmula (de Vet et al., 2006a):

$$EPM_{consistência} = \frac{DP_{dif}}{\sqrt{2}}$$

Foram calculados os valores médios das diferenças entre medições e o Desvio Padrão das diferenças entre medições (DPdiff). A partir destes valores, foi calculada também a DMD através da seguinte fórmula (de Vet et al., 2006b):

$$DMD = 1.96 \times \sqrt{2} \times EPM$$

Foram também calculados os 95% dos Limites de Concordância e o Coeficiente de Repetibilidade.

De forma a calcular a fiabilidade teste-reteste, no que toca as variáveis categóricas – AMP, PIT e MTT – foi utilizado o cálculo o coeficiente de *Kappa* de *Cohen* (de Vet et al., 2011). Foi considerado que valores de 0.80 a 0.90 equivalem a concordância forte e valores acima de 0.90 concordância elevada (McHugh, 2012).

O cálculo destas componentes foi realizado através do programa IBM SPSS *Statistics* ©, versão 29 e o nível de significância considerado foi de $p < 0.05$. Por fim, os diagramas de *Bland & Altman* foram realizados usando o *MedCalc* statistical software versão 19.2.6 (*MedCalc Software* bv, Ostend, Belgium; <https://www.medcalc.org>; 2020). Por fim, os valores absolutos e de percentagem do EPM e da DMD em relação ao respetivo valor médio foram calculados no *Microsoft Excel* 2007 (*Microsoft Corp.*, Redmond, WA).

3. Resultados

3.1 Caracterização da Amostra

Este estudo incluiu uma amostra de 21 participantes (12 mulheres e 9 homens; 41.95 ± 15.11 anos; 68.86 ± 12.62 kg; 170.19 ± 10.37 cm) cujas características sociodemográficas, ocupacionais e de atividade física se encontram descritas na Tabela 1.

Um dos participantes do estudo, aquando do preenchimento questionário sociodemográfico e ocupacional, não respondeu à secção referente aos atributos da atividade profissional. Por essa razão, as percentagens dessa secção são relativas a 20 participantes.

Tabela 1 - Características sociodemográficas, ocupacionais e nível de atividade física dos participantes

	Média \pm Desvio Padrão	%
Características sociodemográficas e ocupacionais		
Idade (anos)	41.95 ± 15.11	
Sexo		
Masculino		42.9%
Feminino		57.1%
Massa (kg)	68.86 ± 12.62	
Altura (cm)	170.19 ± 10.37	
Habilitações literárias		
Ensino básico completo (9ºano de escolaridade)		9.5%
Ensino Secundário ou equivalente completo (12ºano de escolaridade)		33.3%
Ensino superior completo		57.1%
Situação profissional atual		
Trabalhador(a) e/ou estudante		95.2%
Reformado		4.8%
Atributos da atividade profissional		
Predominantemente em pé		40%*
Predominantemente sentado		60%*
Predominantemente na mesma posição		45%*
Predominantemente agachado, ajoelhado e/ou subir/descer escadas		20%*
Trabalho manual com cargas inferiores a 11kg levantar/carregar/mover)		25%*
Trabalho manual com repetição dos mesmos movimentos		0%
Levantar/mover pessoas		15%*
Horas de sono/dia	7 ± 1.52	
Horas em Posição Sentada/dia	4.67 ± 2.54	
Hábitos tabágicos		
Nunca fui fumador(a)		42.9%
Sou ex-fumador(a)		14.3%
Sou fumador(a) ocasional		19.0%
Sou fumador(a) regular		23.8%
Características de atividade física		
Quantos dias fez pelo menos 30 minutos de	2.71 ± 2.08	

atividade física (última semana)		
IPAQ		
Nível de Atividade Física Baixo		38%
Nível de Atividade Física Moderado		57%
Nível de Atividade Física Elevado		5%

*n=20

Com o objetivo de melhor caracterizar a amostra em questão, foi analisada a frequência dos resultados (T0 e T1) de cada teste de aptidão física que se traduz numa variável numérica (Tabela 2). Os dados foram ordenados de forma crescente e, posteriormente, pelos três quartis. Cerca de 25% dos dados são inferiores ou iguais ao primeiro quartil (Q1), cerca de 25% são superiores ou iguais ao terceiro quartil (Q3) e aproximadamente 50%, situam-se entre Q1 e Q3. O segundo quartil (Q2) representa a mediana.

No que toca ao teste 6MWT, a mediana corresponde a 683 metros, sendo que 50% dos valores se situam entre 565.37 e 751.62 metros. Relativamente ao TFT, a mediana é 64.50 segundos, e 50% dos valores se encontram entre 43 e 159.75 segundos. No teste SB, a mediana quando o teste é realizado à direita é de 30.5 segundos, enquanto à esquerda é 23.5 segundos. Metade dos valores obtidos situam-se se entre 17 e 44.5 segundos do lado direito e 15 e 41.75 segundos do lado esquerdo. No teste 60STS, a mediana corresponde a 41 repetições e, 50% dos participantes, obtiveram resultados entre as 35 e 47.25 repetições.

No que toca ao BST, o Q2 é de 56 segundos, e 50% dos resultados encontram-se entre os 37 e 103 segundos. No SRT, o Q2 é de 30.05 centímetros, e 50% dos valores estão entre os 26.85 e 37.30 centímetros. No STF o Q2 é de 3.35 centímetros, e 50% dos valores estão entre 3 e 4.12 centímetros. Por fim, no que toca ao STE, o Q2 é de 29.95° e 50% dos resultados está entre 24.82 e 39.72°.

Tabela 2 – Quartis da frequência de resultados

TESTES	Percentis			TESTES	Percentis		
	25%	50%	75%		25%	50%	75%
6 Minute Walk Test (metros)	565.37	683.00	751.62	Biering-Sørensen Test (segundos)	37.00	56.00	103.00
Trunk Flexor Test (segundos)	43.00	64.50	159.75	Sit-and-Reach Test (centímetros)	26.85	30.05	37.30
Side Bridge Direita (segundos)	17.00	30.50	44.50	Schober Test Flexão (centímetros)	3.00	3.35	4.12
Side Bridge Esquerda (segundos)	15.00	23.50	41.75	Schober Test Extensão (graus)	24.82	29.95	39.72
60s Sit to Stand (repetições)	35.00	41.00	47.25				

Foram também calculadas as frequências dos resultados (T0 e T1) dos testes cujos resultados se traduzem em variáveis categóricas (Tabela 3). Todos os testes apresentaram uma

maioria de resultados negativos, sendo que os testes AMP e PIT, apresentaram mais de 80% dos resultados negativos, enquanto o MTT direita e esquerda apresentou uma frequência de 71.4% e 69% de resultados negativos, respetivamente.

Tabela 3 – Frequência de resultados variáveis categóricas (T0 e T1)

TESTES	Frequência	
	Positivo	Negativo
<i>Aberrant Movement Pattern</i>	11.9%	88.1%
<i>Prone Instability test</i>	14.3%	85.7%
<i>Modified Thomas test</i> direita	28.6%	71.4%
<i>Modified Thomas test</i> esquerda	31.0%	69.0%

3.2 Fiabilidade e Erro de Medição

A fiabilidade e o erro de medição dos testes, assim como todas as medidas a esta associadas, encontram-se descritas na Tabela 4, 5 e 6.

Relativamente à fiabilidade dos testes cujo resultado corresponde a variáveis numéricas, os valores de CCI obtidos foram superiores a 0.90 para todos os testes (Tabela 6). Para os testes com variáveis categóricas, nenhum dos valores de k foi superior a 0.80 (Tabela 4). Estes variaram de 0.54, 95%IC 0.15 a 0.93, a 0.77, 95%IC 0.35 a 1.19, para o MTT à direita e para o AMP, respetivamente. Os restantes valores de k obtidos foram de 0.61, 95%IC 0.12 a 1.1, para o PIT e 0.67, 95%IC 0.32 a 1.01, para o MTT à esquerda.

Tabela 4 – Valores de fiabilidade e concordância dos testes *Aberrant Movement Pattern*, *Prone Instability test* e *Modified Thomas test* direita e esquerda

TESTES	K	95%IC	% Concordância	p
Controlo Neuromuscular				
<i>Aberrant Movement Pattern</i>	0.77	0.35 a 1.19	95.2%	<0.001
<i>Prone Instability test</i>	0.61	0.12 a 1.1	90.5%	0.005
Flexibilidade				
<i>Modified Thomas test</i> direita	0.54	0.15 a 0.93	80.9%	0.011
<i>Modified Thomas test</i> esquerda	0.67	0.32 a 1.01	85.7%	0.002

K, k de Cohen; **95% IC**, intervalo de confiança para o k; **% concordância**, percentagem de concordância entre os dois momentos de avaliação; **p**, nível de significância da diferença dos valores de k dos dois momentos de avaliação

No que toca ao erro de medição, para o 6MWT o valor do EPM foi de 12.13 metros, traduzindo-se numa %EPM de 1.93%. Relativamente aos testes de força muscular, no TFT o EPM foi de 10.24 segundos (%EPM 10.82%); nos SBD e SBE os valores de EPM foram de 2.47 segundos (%EPM 6.14%) e 2.09 segundos (%EPM 5.80%), respetivamente. Observou-se um EPM de 6.16 segundos (%EPM 8.71%) no BST, e de 3.29 repetições (%EPM 7.92%) no 60STS. Por fim nos testes de flexibilidade, foram observados para o ST de flexão e extensão valores de EPM de 0.24 centímetros (%EPM 6.63%) e 1.08 graus (%EPM 3.49%), respetivamente; e para o SRT, 0.94 centímetros (%EPM 3.00%) (Tabela 6).

Tabela 5 – Valores relativos à média, dos testes *6MWT*, *Trunk Flexor Test*, *Side Bridge* direita e esquerda, *60s Sit to Stand*, *Biering-Sørensen Test*, *Sit-and-Reach Test* e *Schober Test* flexão e extensão

TESTES	Média ± DP	95%IC	Média ± DP	95%IC	Min e Max		Média ± DP	Min e Max
	T0		T1		T0	T1	T0 e T1	
Capacidade Cardiorrespiratória								
6 Minute Walk Test (metros)	655.1 ± 138.23	592.22 a 718.07	653.1 ± 137.91	590.32 a 715.873	303.50 a 869.00	300.50 a 875.50	654.12 ± 136.38	300.50 a 875.50
Força Muscular								
Trunk Flexor Test (segundos)	93.57 ± 74.01	59.88 a 127.26	95.71 ± 72.68	62.63 a 128.80	18.00 a 260.00	23.00 a 244.00	94.64 ± 72.46	18.00 a 260.00
Side Bridge Direita (segundos)	39.62 ± 40.32	21.26 a 57.97	40.86 ± 38.83	23.18 a 58.53	7.00 a 182.00	8.00 a 174.00	40.24 ± 39.10	7.00 a 182.00
Side Bridge Esquerda (segundos)	36.33 ± 35.05	20.38 a 52.29	35.76 ± 33.75	20.40 a 51.12	9.00 a 159.00	8.00 a 154.00	36.05 ± 33.98	8.00 a 159.00
60s Sit to Stand (repetições)	41.33 ± 9.80	36.87 a 45.80	41.81 ± 9.40	37.53 a 46.09	25.00 a 61.00	23.00 a 60.00	41.57 ± 9.49	23.00 a 61.00
Biering-Sørensen Test (segundos)	69.48 ± 48.36	47.46 a 91.49	71.86 ± 46.67	50.61 a 93.10	13.00 a 168.00	18.00 a 162.00	70.67 ± 46.95	13.00 a 168.00
Flexibilidade								
Sit-and-Reach Test (centímetros)	31.47 ± 7.50	28.06 a 34.88	31.49 ± 7.84	27.93 a 35.06	16.50 a 43.00	15.30 a 43.10	31.48 ± 7.58	15.30 a 43.10
Schober Test Flexão (centímetros)	3.67 ± 1.06	3.18 a 4.15	3.64 ± 0.84	3.26 a 4.02	2.50 a 7.00	2.20 a 5.90	3.65 ± 0.95	2.20 a 7.00
Schober Test Extensão (graus)	30.97 ± 8.59	27.06 a 34.88	30.81 ± 8.12	27.11 a 34.50	16.50 a 44.30	16.30 a 44.0	30.89 ± 8.25	16.30 a 44.30

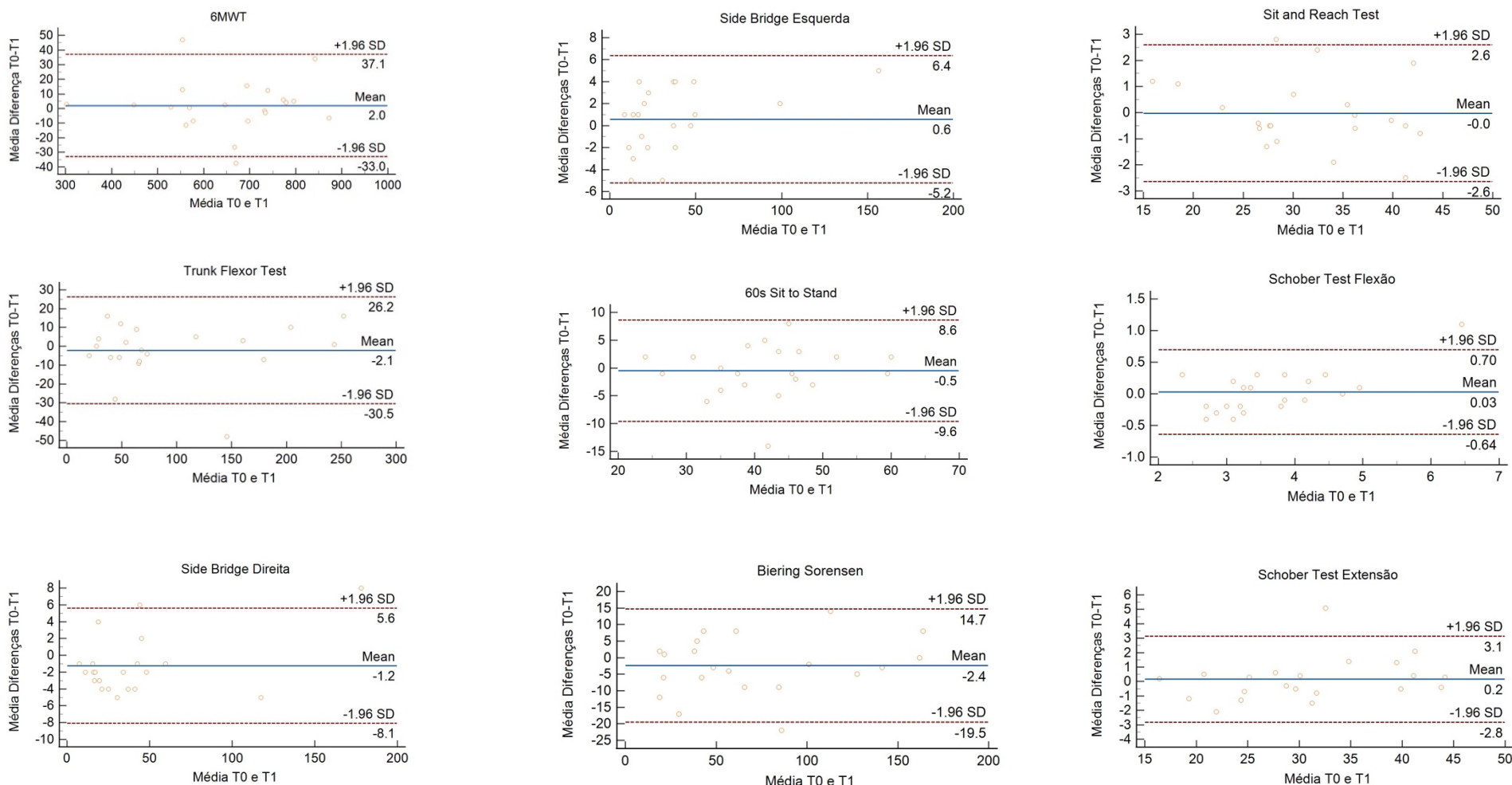
Média, média das medições e **DP**, desvio padrão; **95% IC**, intervalo de confiança para as médias; **Min a Max**, valores mínimos e máximo da média; **T0**, primeiro momento de avaliação; **T1**, segundo momento de avaliação;

Tabela 6 – Valores de fiabilidade e erro de medição, dos testes *6MWT*, *Trunk Flexor Test*, *Side Bridge* direita e esquerda, *60s Sit to Stand*, *Biering-Sørensen Test*, *Sit-and-Reach Test* e *Schober Test* flexão e extensão

TESTES	CCI	95%IC	Dif	95% IC	p	DPdif	95% LDC	EPM	%EPM	DMD	%DMD	CR	95% IC
Capacidade Cardiorrespiratória													
6 Minute Walk Test (metros)	0.99	0.990 a 0.998	2.05	-6.08 a 10.18	0.531	±17.86	-32.96 a 37.06	12.63	1.93%	35.01	5.35%	34.40	26.47 a 49.16
Força Muscular													
Trunk Flexor Test (segundos)	0.99	0.98 a 0.99	-2.14	-8.73 a 4.45	0.852	±14.48	-30.53 a 26.24	10.24	10.82%	28.38	29.99%	28.02	21.56 a 40.04
Side Bridge Direita (segundos)	0.99	0.995 a 0.999	-1.23	-2.83 a 0.35	0.069	±3.49	-8.08 a 5.60	2.47	6.14%	6.84	17.01%	7.11	5.47 a 10.15
Side Bridge Esquerda (segundos)	0.99	0.995 a 0.999	0.57	-0.77 a 1.92	0.407	±2.96	-5.23 a 6.37	2.09	5.80%	5.80	16.09%	5.77	4.44 a 8.25
60s Sit to Stand (repetições)	0.94	0.85 a 0.97	-0.48	-2.59 a 1.64	0.881	±4.65	-9.60 a 8.65	3.29	7.92%	9.12	21.94%	8.95	6.89 a 12.79
Biering-Sørensen Test (segundos)	0.99	0.97 a 0.99	-2.38	-6.34 a 1.58	0.225	±8.71	-19.45 a 14.69	6.16	8.71%	17.07	24.16%	17.30	13.3 a 24.72
Flexibilidade													
Sit-and-Reach Test (centímetros)	0.99	0.98 a 0.99	-0.02	-0.63 a 0.58	0.651	±1.33	-2.64 a 2.59	0.94	3.00%	2.61	8.31%	2.55	1.96 a 3.65
Schober Test Flexão (centímetros)	0.97	0.92 a 0.99	0.03	-0.13 a 0.18	0.970	±0.34	-0.64 a 0.70	0.24	6.63%	0.67	18.37%	0.65	0.51 a 0.94
Schober Test Extensão (graus)	0.99	0.98 a 0.99	0.16	-0.54 a 0.85	0.903	±1.52	-2.83 a 3.14	1.08	3.49%	2.98	9.66%	2.93	2.25 a 4.19

CCI, coeficiente de correlação intra-classe; **95% IC**, intervalo de confiança para o CCI; **Dif**, média das diferenças entre o primeiro e o segundo momento; **95% IC**, intervalo de confiança para a diferença; **p**, nível de significância da diferença entre médias; **DPdif**, desvio padrão das diferenças; **EPM**, erro padrão de medida; **%EPM**, Percentagem do erro padrão de medida em relação ao respetivo valor médio; **LDC**, 95% dos limites de concordância; **DMD**, diferença mínima detetável; **DMD%**, Percentagem da diferença mínima detetável em relação ao respetivo valor médio; **CR**, Coeficiente de Repetibilidade; **95% IC**, intervalo de confiança CR

Figura 1 – Diagramas de Bland-Altman dos testes 6MWT, Trunk Flexor Test, Side Bridge direita e esquerda, 60s Sit to Stand, Biering-Sørensen Test, Sit-and-Reach Test e Schober Test flexão e extensão.



Mean, média das diferenças entre o primeiro e o segundo momento de avaliação; **SD**, 95% dos limites de concordância

Foi calculada para estes testes a DMD, que para o 6MWT foi de 35.01m 5.35% (DMD% 5.35%); para o TFT 28.38 s (DMD% 29.99%); para o SBD e SBE 6.84 s (DMD% 17.01%) e 5.80 s (DMD% 16.09%), respetivamente; para o BST a DMD foi de 17.07 s (DMD% 24.16%); e para o 60STS 9.12 repetições (DMD% 21.94%). Para a os testes relacionados com a flexibilidade os valores da DMD foram de 2.61 cm (DMD% 8.31%) para o SRT e 0.67 cm (DMD% 18.37%) e 2.98° (DMD% 9.66%) para o STF e STE respetivamente. Quando comparados os valores obtidos em cada teste nos dois momentos de avaliação, não foram encontradas diferenças estatisticamente significativas (Tabela 6).

Relativamente aos diagramas de *Bland Altman* (Fig. 1), é possível verificar que, a linha da média das diferenças é diferente de zero em todos os gráficos (exceto no gráfico do SRT, em que é 0), no entanto quando comparados com a magnitude das médias, tal não parece significar a presença de viés relevante. Para além disso, observando o valor p (H_0 : Média=0) dos diagramas, não existe nenhum valor $p < 0.05$. Por esta razão, nenhum dos testes parece apresentar erro sistemático absoluto – o quanto as diferenças entre T0 e T1 se afastam do zero (Atkinson & Nevill, 1998).

Quando observamos o diagrama do 6MWT (Fig. 1), a dispersão dos resultados não parece estar relacionada com a magnitude das medições, no entanto os LDC correspondem a um intervalo amplo (-32.96 a 37.06). O Coeficiente de Repetibilidade, para este teste é equivalente a 34.40, 95%IC 26.47 a 49.16.

Os testes de Força Muscular, apresentam uma dispersão de resultados normal, sem aparente impacto da magnitude de medições (Fig. 1). Os LDC correspondem a intervalos amplos para os testes TFT, -30.53 a 26.24; BST, -19.45 a 14.69; 60STS, -9.60 a 8.65. E intervalos de LDC restritos para SBD, -8.08 a 5.60; SBE, -5.23 a 6.37. E os Coeficientes de Repetibilidade para estes testes são 28.02, 95%IC 21.56 a 40.04 para o TFT; 7.11, 95%IC 5.47 a 10.15 para o SBDta; 5.77, 95%IC 4.44 a 8.25 para o SBEsq; 8.95, 95%IC 6.89 a 12.79 para o 60sSTS e 17.30, 95%IC 13.3 a 24.72 para o BST.

Por fim, relativamente aos testes de flexibilidade, e no que diz respeito ao SRT e ao STE, a dispersão dos resultados não parece estar relacionada com a magnitude das medições (Fig. 1). No entanto, parece existir um comportamento diferente dos resultados no STF. O padrão observado no diagrama deste teste, parece indicar a presença de erro proporcional. Os seus LDC são correspondentes a intervalos pouco amplos entre -2.64 a 2.59 para o SRT; -2.83 a 3.14 para o STF; -0.64 a 0.70 para o STE. Relativamente ao Coeficiente de Repetibilidade, para o SRT o CR

é de 2.55, 95%IC 1.96 a 3.65, para o STF 0.65, 95%IC 0.51 a 0.94 e para o STE 2.93, 95%IC 2.25 a 4.19.

Todos os diagramas, apresentam *outliers*, correspondendo a indivíduos distintos, que ultrapassam os LDC, superior e/ou inferiormente, exceto o teste SBESq, onde não são observados quaisquer *outliers*. Após análise dos *outliers*, foi identificado um indivíduo que constituía um *outlier* para o 6MWT, 60STS e STF, no entanto não foram encontradas características específicas que o distinguíssem dos restantes participantes.

Relativamente ao erro de medição dos testes com variáveis categóricas (Tabela 4), os testes que avaliam o controlo neuromuscular apresentaram uma percentagem de concordância superior a 90% (AMP – 95.2% e PIT – 90.5%); enquanto o MTT à direita e esquerda apresentou 80.9% e 85.7% de concordância respetivamente.

4. Discussão

O presente estudo contribui para definir valores de fiabilidade intra-avaliador, erro de medição e DMD para uma bateria de testes de avaliação da aptidão física, de forma a colmatar a carência de informação sobre as suas propriedades psicométricas na população adulta saudável portuguesa.

A amostra é constituída por 21 participantes com idade média de 41.95 ± 15.11 anos. O tamanho de amostra não diferiu de forma significativa de outros estudos semelhantes, acerca da fiabilidade de testes de terreno (Gabbe et al., 2004; Reiman et al., 2012), que apresentam uma variação entre 15 e 28 participantes.

A bateria em estudo, é constituída por 11 testes físicos, que avaliam a aptidão física dos indivíduos. Em conjunto, estes testes avaliam de forma abrangente todas as componentes da aptidão física relacionadas com a saúde músculo-esquelética definidas pela ACSM (2018). Estes são testes de aplicação simples e prática, frequentemente utilizados e foram previamente estudados em diversas populações com condições músculo-esqueléticas crónicas ou persistentes, tornando esta bateria abrangente para a avaliação de indivíduos com diversas condições do foro músculo-esquelético.

Os testes que se traduzem em variáveis numéricas, demonstraram de uma forma geral fiabilidade excelente ($CCI > 0.94$) (Koo & Li, 2016) e valores de %EPM aparentemente baixos ($\%EPM < 10.82\%$). A maioria dos testes também apresentou valores de %DMD aparentemente baixos ($\%DMD < 18.37\%$), no entanto para o TFT, 60STS e BST, estes parecem ser altos ($21.94\% < \%DMD < 29.99\%$).

Para a avaliação da capacidade cardiorrespiratória, o teste em estudo foi o 6MWT (Enright & Sherrill, 1998). Observando os resultados obtidos e a média de T0 e T1, 654.12 ± 36.8 metros, podemos comparar aos dois estudos realizados na população saudável portuguesa (Oliveira et al., 2019; Marques et al., 2020). À semelhança deste estudo, ambos realizaram o teste com a distância de 30 metros e em indivíduos saudáveis. No estudo de Oliveira et al. (2019) podemos verificar que as médias dos dois momentos de avaliação correspondem a 604.8 ± 68.6 metros (T0) e 627.8 ± 73.6 metros (T1), para uma população com idade média de 39.0 ± 13.1 anos. No estudo de Marques et al. (2020), a média é ligeiramente inferior, 524 ± 208 metros, sendo que a população apresenta uma idade média de 53 ± 25 anos. Tendo em conta a idade, e a sua relação com o desempenho neste testes (Enright & Sherrill, 1998), é expectável que os resultados se aproximem mais dos de Oliveira et al. (2019), já que a amostra do nosso estudo tem uma idade média de 41.95 ± 15.11 anos.

Este teste apresentou uma fiabilidade intra-avaliador excelente (Koo & Li, 2016) (CCI=0.99, 95%IC 0.990 a 0.998) com um intervalo de confiança restrito. Estes valores revelam uma fiabilidade ligeiramente superior aos obtidos por Arcuri et al. (2016) na população saudável brasileira, que revelaram boa fiabilidade (Koo & Li, 2016) (CCI=0.86, 95%IC 0.80 a 0.90). Este teste apresentou um valor de %EPM (1.93%) aparentemente baixo, demonstrando um erro mínimo associado a este teste. O mesmo acontece com a %DMD (5.35%), que também parece ser baixa, o que significa que o 6MWT é um teste com elevado potencial para revelar mudanças reais na capacidade cardiorrespiratória em indivíduos saudáveis, quando submetidos a 2 avaliações.

Ao analisar o diagrama de *Bland Altman*, podemos verificar que a média das diferenças é de 2.05, 95%IC -6.08 a 10.18. Apesar de o valor ser próximo de zero, indicando uma diferença mínima entre medições, a amplitude do IC indica que é preciso alguma cautela na interpretação da diferença de médias entre duas avaliações consecutivas. O intervalo dos LDC (-32.96 a 37.06) parece ser amplo, com uma excursão de 70.02 m, sendo significativamente superior ao desvio padrão da média das medições (± 36.8 metros). Indicando uma variabilidade substancial intra e inter individual entre os dois momentos de avaliação.

Existem dois *outliers*, que quando identificados não parecem ter características distintas dos outros participantes, ou semelhantes entre si. Podemos verificar que apesar se apresentarem fora dos LDC, as medições de T0 e T1 de ambos (T0: 578.0 m T1: 531.0 m e T0: 653.5 m T1: 680.0 m) estão relativamente próximas. Podemos considerar que como as restantes medições apresentam uma grande homogeneidade, estes dois *outliers*, apesar de estarem fora dos LDC, continuam a apresentar uma diferença entre medições baixa.

Para este teste o CR foi 34.40m, indicando que, nas mesmas condições, existe 95% de probabilidade que a diferença entre as duas medições seja inferior a 34.40m. Quando comparado com o desvio padrão da média (± 36.8 metros) este parece sugerir uma baixa variabilidade de resultados.

No que toca aos testes de força muscular, para o TFT, os resultados obtidos foram em média 94.64 ± 72.46 s (min 18s e máx 260s). O estudo americano de Anderson et al. (2014) avaliou uma amostra de 116 indivíduos com uma idade média de 28.8 ± 9.64 anos (idades compreendidas entre 19 e 55 anos) e apresentou uma média de 125.2 ± 70.2 s para indivíduos que não atletas de competição. Existe uma diferença de mais de 10 anos nas médias de idades entre os dois estudos que de alguma forma poderá justificar os resultados obtidos neste estudo e no de Anderson et al. (2014): 94.64 ± 72.46 s e 125.2 ± 70.2 s respetivamente).

O TFT apresentou uma fiabilidade intra-avaliador excelente (Koo & Li, 2016) com um intervalo de confiança restrito (CCI=0.99, 95%IC 0.98 a 0.99). Este resultado é semelhante ao encontrado por McGill et al. (1999) na mesma população (CCI=0.97) e por Evans et al. (2007) para atletas (CCI=0.95), o que concorre para podermos afirmar que se trata de um teste fiável. Estes valores, são significativamente superiores aos de Durall et al. (2012) (CCI=0.71, 95%IC 0.38 a 0.91).

Relativamente aos valores de %EPM (10.82%) verifica-se um erro menor comparativamente ao obtido no estudo de Evans et al. (2007) em atletas (%EPM=16.3%). Ainda que, este seja o teste da bateria com um maior EPM, este parece ser um valor baixo, indicando um baixo erro associado à medição. Já a %DMD (29.99%) parece ser um valor consideravelmente alto, revelando que, apesar deste teste ter alta fiabilidade e um baixo erro associado, é preciso ter cautela na sua capacidade em identificar mudanças reais nos indivíduos.

Quando observado o diagrama de Bland Altman do TFT, a média é de -2.14 s, 95%IC -8.73 a 4.45, demonstrando proximidade em relação ao zero, o que sugere ausência de erro sistemático, no entanto, o seu IC parece ser amplo, indicando alguma incerteza do real valor da média. O intervalo de LDC parece ser amplo (95%LDC -30.53 a 26.24), indicando uma grande variabilidade intra e inter individual. Esta variabilidade é corroborada pelo DP da média ($\pm 72.46s$) que também parece ser alto. Para este teste o CR foi 28.02s, indicando que, nas mesmas condições, existe 95% de probabilidade que a diferença entre as duas medições, seja no máximo de 28.02s, o que tendo em conta a magnitude das medições é uma diferença razoavelmente alta, principalmente considerando o valor de DMD (28.38s).

O teste SB, que avalia a força muscular dos flexores laterais do tronco (Anderson et al., 2014; McGill et al., 1999), apresentou resultados consistentes para ambos os lados. No que diz respeito à fiabilidade intra-avaliador, o CCI de 0.99, 95%IC 0.99 a 0.99 para ambos os lados, sugere um elevado nível de consistência entre medições. Quando comparados com estudos anteriores, a fiabilidade apresentada por McGill et al. (1999) (SBD e SBE CCI=0.99) que avaliou 5 indivíduos saudáveis, é a que se aproxima mais dos resultados do presente estudo. Este estudo revelou valores de EPM para o SB, aparentemente baixos (SBD %EPM=6.14% e SBE %EPM=5.80%), indicando uma elevada precisão dos resultados. Estes valores são significativamente inferiores aos encontrados nos estudos de Larsson et al. (2004) (SBD %EPM=13.2% e SBE %EPM=16.8%) e Evans et al. (2007) (SBD %EPM=17.6% e SBE %EPM=19.8%) em militares e atletas, respetivamente. A %DMD foi de 17.01% para o SBD e

16.09% para o SBE, indicando um aparente potencial para revelar mudanças reais nos indivíduos, ou seja, mudanças que não sejam atribuídas ao erro do próprio teste.

Relativamente à informação obtida a partir dos diagramas de *Bland Altman*, a média das diferenças parece ser reduzida para ambos os lados e com IC restritos que passam pelo valor de zero (SBD -1.23s, 95%IC -2.38 a 0.35 e SBE 0.57, 95%IC -0.77 a 1.92), indicando ausência de erro sistemático. Os intervalos de LDC (SBD 95%LDC -8.08 a 5.60 e SBE 95%LDC -5.23 a 6.37) parecem ser restritos, indicando concordância nas medições repetidas. Para este teste o CR foi 7.11s para a direita e 5.77s para a esquerda, indicando que, nas mesmas condições, existe 95% de probabilidade que a diferença entre duas medições em dois momentos diferentes no tempo, seja inferior a 7.11s e 5.77s. Estes valores parecem indicar uma moderada repetibilidade já que se aproximam dos valores de DMD (SBD 6.84s e SBE 5.80s).

O BST, que avalia a força de resistência dos músculos extensores do tronco (Demoulin et al., 2006), apresentou neste estudo um valor de CCI excelente (Koo & Li, 2016) (0.99, 95%IC 0.97 a 0.99), com um IC restrito. Indicando uma fiabilidade excelente, superior à encontrada por Martínez-Romero et al. (2020), CCI=0.88, 95%IC 0.80 a 0.92. O EPM para este teste foi de 6.16s (%EPM=8.71%), sugerindo um baixo erro de medição, inferior ao encontrado por Latimer et al. (1999) (EPM=17.4s) e por Ozcan Kahraman et al. (2015) (EPM=8.71s). A %DMD (24.14%) parece ser algo significativa, o que poderá afetar o potencial deste teste para identificar diferenças reais quando aplicado a indivíduos saudáveis.

O digrama de Bland-Altman que diz respeito a este teste demonstra uma média das diferenças próxima de zero, -2.38, 95%IC -6.34 a 1.58, com um IC relativamente restrito, indicando ausência de erro sistemático. Relativamente aos LDC (95%LDC -19.45 a 14.69), este intervalo parece ser amplo, demonstrando uma possível variação de 34.14s, que parece muito superior ao CR (17.30s). Demonstrando que, ainda que o valor médio das diferenças demonstre ausência de erro, o amplo intervalo de LDC indica uma substancial variabilidade entre os dois momentos de avaliação. Para este teste o CR foi 17.30s indicando, que nas mesmas condições, existe 95% de probabilidade que a diferença entre as duas medições, se situe no máximo a este valor, sendo uma diferença razoavelmente baixa, tendo em conta a DMD e a magnitude das medições.

Por fim, o 60STS, que avalia a musculatura dos membros inferiores Bohannon & Crouch, (2019), obteve, no que toca a fiabilidade, um valor de CCI excelente (Koo & Li, 2016) (0.94, 95%IC 0.85 a 0.97) indicando alta fiabilidade, com um intervalo de confiança restrito. Ainda que não existam estudos sobre a fiabilidade na população saudável em todas as idades, o estudo de Ritchie et al. (2005), com uma população saudável, mas com idades entre os 55 e 70 anos,

apresenta também valores altos de fiabilidade (CCI=0.80). Este teste apresentou um valor de EPM (3.29 repetições; EPM% 7.92%) aparentemente baixo, indicando um baixo erro do instrumento. No entanto a sua %DMD é de 21.94%, indicando aparentemente uma grande margem de erro, potencialmente limitando a sensibilidade do teste para detetar diferenças reais nos indivíduos.

Quando analisamos o diagrama de Bland-Altman podemos verificar que a média das diferenças está muito próxima do zero (-0.48, 95%IC -2.59 a 1.64), e com um IC que passa por 0 indicando que esta diferença não é estatisticamente diferente de zero. No entanto, ao observar os LDC (-9.60 a 8.65), podemos verificar que se trata de um intervalo amplo, abrangendo 18.25 repetições, valor que é significativamente superior ao desvio padrão da média (± 9.49 repetições). Estes valores indicam-nos que, enquanto a média das diferenças demonstra ausência de erro sistemático, o amplo intervalo de LDC indica uma substancial variabilidade entre os dois momentos de avaliação. Para este teste o CR foi 8.95 repetições, indicando que, nas mesmas condições, existe 95% de probabilidade que a diferença entre as duas medições, se situe no máximo a 8.95 repetições, o que parece uma diferença baixa, tendo em conta a média das medições.

No que toca aos testes de avaliação da componente da flexibilidade, o SRT apresentou, à semelhança do documentado pelo estudo de Gabbe et al. (2004) em indivíduos saudáveis (CCI=0.97) e no de Liu et al. (2022) para jovens adultos saudáveis (CCI=0.99), um valor de fiabilidade inter-avaliador excelente (Koo & Li, 2016), com um IC restrito, CCI: 0.99, 95%IC 0.98 a 0.99. No presente estudo foi estudado o EPM, tendo o SRT obtido um valor de 0.94cm (%EPM=3%), inferior ao encontrado por Gabbe et al. (2004) (EPM=2cm) na mesma população. Relativamente à DMD este estudo apresentou um valor de 2.61cm (%DMD=8.31%), no estudo de Liu et al. (2022) para uma população de jovens adultos, este valor foi de 3.42cm. Estes valores demonstram uma grande confiabilidade, parecendo ser um teste fiável, com um baixo erro de medição e que demonstra sensibilidade às mudanças reais do indivíduo.

Quando interpretado o diagrama de Bland-Altman, este teste mantém uma ótima prestação, apresentando um intervalo aparentemente restrito de LDC (-2.64 a 2.59) e uma média de diferenças entre T0 e T1 de -0.02 ± 1.33 cm, 95%IC -0.63 a 0.58, demonstrando uma grande precisão entre medições e ausência de erro sistemático. Para este teste o CR foi 2.55cm, indicando que, nas mesmas condições, existe 95% de probabilidade que a diferença entre as duas medições, se situe no máximo a este valor, sendo uma diferença baixa, tendo em conta a média das medições.

Por fim, para avaliar a amplitude de movimento da coluna lombar em flexão e em extensão, os testes STF e STE. Ambos apresentam valores de referência para a população saudável

americana, para o STF os valores variam entre as faixas etárias de 20-29 anos (3.7 cm) e 60-69 anos (2.4 cm), sendo o valor mais alto atribuído aos indivíduos entre os 30 e 39 anos (3.9 cm) (Fitzgerald et al., 1983). Observando os valores obtidos no presente estudo (25% 3.00 cm; 50% 3.35 cm; 75% 4.12), estes parecem ser ligeiramente superiores. Para o STE no estudo de Fitzgerald et al. (1983) os valores variam de 41.2° para idades entre os 20-29 anos e 17.4° para os 60-69 anos, parecendo estar próximos dos observados no presente estudo (25% 24.82°; 50% 29.95°; 75% 39.72°). Devemos ter em consideração que a amostra de Fitzgerald et al. (1983) se compõe maioritariamente de homens (97.7% de homens), enquanto a amostra do presente estudo é mais equilibrada relativamente ao sexo (42.9% de homens), o que pode impactar a comparação.

Ambos apresentaram uma excelente fiabilidade (Koo & Li, 2016) com ICs restritos (STF CCI=0.97, 95%IC 0.92 a 0.99; STE CCI=0.99, 95%IC 0.98 a 0.99). Estes valores vão ao encontro dos obtidos no estudo de Fitzgerald et al. (1983) para uma população saudável (4 mulheres e 168 homens) com idades compreendidas entre os 20 e os 82 anos. O erro de medição parece ser baixo para ambos os testes (STF %EPM=6.63%, STE %EPM=3.49%). E os valores de %DMD foram de 18.37% para o STF e 9.66% para o STE. Estes testes parecem apresentar um baixo erro de medição, e o STE apresenta uma DMD significativamente mais baixa que a do STF. Estes resultados indicam que, sob as mesmas condições, para o STF é necessária uma maior mudança entre avaliações consecutivas para se ter confiança que esta é real e não devida a erro de medição, quando comparado com o STE.

Ambos apresentam intervalos de LDC restritos (STF 95%LDC=-0.64 a 0.70; STE 95%LDC=-2.83 a 3.14), e médias das diferenças entre T0 e T1 muito próximas do zero (STF Dif= 0.03 ±0.34 cm, 95%IC -0.13 a 0.18; STE Dif=0.16 ±1.52°, 95%IC -0.54 a 0.85); indicando baixa variabilidade de medições entre sessões e ausência de erro sistemático. Ao observar o diagrama do STF e a sua distribuição de dados, parece existir, um comportamento diferente em faixas de valores diferentes. Para magnitudes de medição menores, existe uma sobrevalorização de T1, enquanto em magnitudes maiores, existe uma sobrevalorização de T0. Indicando existência de erro proporcional, o que pode indicar, que em determinadas faixas de valores, as medições podem não ser verdadeiras, impactando, desta forma a fiabilidade deste teste. Será importante testar o STF com uma maior amostra, de forma a ter uma visão mais geral sobre a distribuição dos dados. Para estes testes o CR foi de 0.65cm para o STF e 2.93° para o STE, estes parecem ser valores baixos, quando em conta a média das medições, corroborando a fiabilidade dos testes.

Os testes que se traduzem em variáveis categóricas, obtiveram de uma forma geral fiabilidade fraca ($k=0.54$) a moderada ($k>0.61$) (McHugh, 2012), com ICs relativamente amplos e valores de concordância altos (%concordância $>80.9\%$).

Relativamente aos testes que avaliam a componente de aptidão física controlo neuromuscular; o AMP, que avalia a qualidade do movimento na flexão e extensão do tronco e instabilidade segmentar lombar, foi negativo em 88.1% das medições. Os 11.9% de medições positivas, quando observadas, correspondem a dois participantes com idades acima dos 43 anos, que para além de apresentar altos níveis de stress, cansaço e sentimentos negativos associados ao trabalho, não realizam exercício físico, e ainda que não apresentem queixas álgicas, podem ser resultados expectáveis tendo em conta o contexto. A frequência de resultados foi muito diferente da encontrada no estudo de Biely et al. (2014), que no grupo saudável apresentou 80.0% de positivos.

Este teste apresentou fiabilidade moderada (McHugh, 2012), $k=0.77$, 95%IC 0.35 a 1.19, apresentando, no entanto, um amplo IC, demonstrando alguma incerteza na real fiabilidade, já que o limite superior representa uma fiabilidade perfeita e o inferior fiabilidade mínima (McHugh, 2012). Quando comparado com os resultados do estudo de Biely et al. (2014) para a mesma população ($k=0.72$), parecem apresentar resultados semelhantes. Corroborando o valor de k deste teste a percentagem de concordância foi de 95.2%, com um $p<0.001$, indicando que é improvável, que a alta concordância entre medições, seja devido a erro aleatório, dando mais confiança, na utilização deste instrumento.

O PIT, utilizado para a identificação de instabilidade segmentar da coluna lombar apresentou (Hicks et al., 2005), apresentou neste estudo fiabilidade moderada (McHugh, 2012), $k=0.61$, 95%IC 0.12 a 1.1, e uma %concordância de 90.5%. Estes resultados sugerem que apesar da fiabilidade moderada, o amplo IC demonstra alguma incerteza na estimativa da fiabilidade. No entanto, o valor da concordância demonstra uma grande consistência entre medições. Não existem, na literatura, estudos acerca da fiabilidade intra-avaliador deste teste, impossibilitando desta forma a comparação entre resultados.

Já na componente de flexibilidade, o MTT, que avalia a flexibilidade dos músculos flexores da anca, apresentou para o lado direito, fiabilidade fraca com um IC amplo ($k=0.54$, 95%IC 0.15 a 0.93) e para o lado esquerdo fiabilidade moderada também com IC amplo ($k=0.67$, 95%IC 0.32 a 1.01) e uma percentagem de concordância alta (%concordância $>80.9\%$) (McHugh, 2012). Ao compararmos estes resultados com os dos estudos de Cady et al. (2022) (k de Fleiss entre 0.95 e 1) e de Cejudo et al. (2015) (ICC=0.87 a 0.91) em atletas, observamos que os valores do presente

estudo apresentam um nível de fiabilidade significativamente mais baixo. Os resultados obtidos sugerem que este teste apesar de uma concordância aparentemente alta, apresenta fiabilidade moderada a fraca. Tendo em conta que o k de Cohen tem em consideração a concordância esperada por acaso, é possível que a elevada %concordância, não seja totalmente atribuível a uma real concordância entre medições (McHugh, 2012). Será importante em futuros estudos, perceber a origem desta variabilidade entre medições.

Analisando os resultados, para os testes que se traduzem em variáveis numéricas, existe uma clara superioridade do SRT, SB, STE e STF, demonstrando excelente fiabilidade e baixo erro associado, sendo os testes que apresentam resultados mais robustos. O 6MWT, BST e 60STS, apesar de apresentarem uma excelente fiabilidade, apresentam alguns resultados que sugerem alguma margem de erro, será importante em estudos futuros investigar a origem do erro de forma a tornar estes instrumentos mais confiáveis. Por fim o TFT apresenta o maior erro associado, o que pode impactar a sua utilização na prática clínica, ainda que seja um teste fiável.

Nos testes que se traduzem em variáveis categóricas, o AMP, PIT e MTT do lado esquerdo apresentaram fiabilidade moderada, no entanto IC amplos, pondo em causa a real fiabilidade destes testes, no entanto apresentaram baixo erro associado. Por fim o MTT do lado direito não apresenta uma fiabilidade aceitável (fraca e com IC amplo) ainda que o erro pareça ser relativamente baixo.

Tendo em conta o presente estudo será importante a investigação futura relativamente a fiabilidade inter-avaliador destes testes e, para os estudos que apresentam maior erro associado, investigar a origem desse erro. Desta forma ficando a conhecer de forma mais abrangente as propriedades psicométricas, e caso estas sejam adequadas, estabelecer valores de referência para a população adulta saudável portuguesa.

Contributos e Limitações do estudo

Este estudo contribuiu para ajudar a colmatar alguns dos *gaps* existentes na literatura, sobre as propriedades psicométricas, relativas à fiabilidade de onze testes de aptidão física. Para além de estabelecer valores de fiabilidade intra-avaliador, foram estabelecidos valores de EPM e DMD, de forma a ajudar a contextualizar e interpretar os resultados na prática clínica.

Destacamos alguns pontos fortes do presente estudo, que contribuíram para aumentar a segurança e interpretação dos resultados. Todas as medições em T0 e T1, para todos os participantes, foram realizadas da mesma forma e pela mesma ordem, no mesmo local adequado à recolha, e no mesmo período do dia, de forma a maximizar a fiabilidade dos resultados. Para

além disto, de forma a melhorar a qualidade das recolhas, o avaliador foi alvo de um treino prévio na implementação dos testes, com instruções claras tanto para avaliador como participantes, por um investigador experiente.

Como limitações do estudo, destacamos o tamanho da amostra, que apesar de ser calculado com base no CCI Kraemer & Thiemann (1987), poderá ser insuficiente para o cálculo dos LDC e interpretação dos diagramas de Bland-Altman, já que para este tipo de análise é recomendada uma maior amostra. Consideramos também que apesar de não ser o objetivo do estudo, o estudo da fiabilidade inter-avaliador seria importante. E por fim, consideramos também o facto de o avaliador não ser um elemento cego, já que os dois momentos de avaliação e o tratamento de dados é realizado pela mesma pessoa, ainda que as folhas de registo de T0 e T1 fossem diferentes de forma a diminuir o viés.

5. Conclusão

O presente estudo contribuiu para a definição de valores de fiabilidade intra-avaliador, EPM e DMD, para uma bateria de testes da aptidão física na população saudável portuguesa. Foram considerados fiáveis e com um mínimo erro de medição os testes SRT, SB, STE e STF, apresentando os resultados mais promissores. Os testes 6MWT, BST e 60STS apresentaram fiabilidades excelentes, no entanto com maior percentagem de erro de medição, necessitando de mais informação acerca da origem deste erro. Os testes AMP, PIT e MTT do lado esquerdo, apesar de um baixo erro de medição, apresentam fiabilidade moderada com IC amplos, indicando incerteza sobre a sua real fiabilidade, sendo também necessários mais estudos para definir a sua fiabilidade. Por fim o TFT ainda que fiável, apresenta um grande erro associado; e o MTT do lado direito ainda que apresente um erro de medição baixo, não parece ser um instrumento fiável.

Este estudo aumenta o conhecimento até agora limitado sobre as propriedades psicométricas dos testes individualmente, na população saudável portuguesa.

É fundamental frisar a importância da implementação destes testes e, por essa razão, a importância de conhecer profundamente as suas propriedades psicométricas, assim como valores de referência, de forma a aplicar e interpretar de forma fiável e correta. Assim podendo ter um impacto positivo, na reabilitação de indivíduos, cuja aptidão física foi impactada por dor e/ou condições músculo-esqueléticas crónicas ou persistentes.

Por estas razões recomendamos o desenvolvimento de estudos de qualidade e amostra representativa acerca das propriedades psicométricas destes testes, de forma a aprofundar o conhecimento sobre os testes inseridos nesta bateria.

Bibliografia

- Alyazedi, F. M., Lohman, E. B., Wesley Swen, R., & Bahjri, K. (2015). The inter-rater reliability of clinical tests that best predict the subclassification of lumbar segmental instability: structural, functional and combined instability. *The Journal of Manual & Manipulative Therapy*, 23(4), 197–204.
- American College of Sports Medicine. (2018). *ACSM's Guidelines for Exercise Testing and Prescription* (10th ed.). LWW. <https://doi.org/10.1017/CBO9781107415324.004>;
- Anderson, A., Hoffman, J., Johnson, B., Simonson, A., & Urquhart, L. (2014). Doctor of Physical Therapy Research Papers Physical Therapy 4-2014 Anna; and Urquhart. Doctor of Physical Therapy Research, 1–57.
- Arcuri, J. F., Borghi-Silva, A., Labadessa, I. G., Sentanin, A. C., Candolo, C., & Pires Di Lorenzo, V. A. (2016). Validity and Reliability of the 6-Minute Step Test in Healthy Individuals: A Cross-sectional Study. *Clinical Journal of Sport Medicine: Official Journal of the Canadian Academy*
- Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26(4), 217–238. *Sport Medicine*, 26(1), 69–75.
- ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. (2002). ATS statement: guidelines for the six-minute walk test. *American Journal of Respiratory and Critical Care Medicine*, 166(1), 111–117.
- Biely, S. A., Silfies, S. P., Smith, S. S., & Hicks, G. E. (2014). Clinical observation of standing trunk movements: what do the aberrant movement patterns tell us? *The Journal of Orthopaedic and Sports Physical Therapy*, 44(4), 262–272.
- Blyth, F. M., Briggs, A. M., Schneider, C. H., Hoy, D. G., & March, L. M. (2019). The Global Burden of Musculoskeletal Pain-Where to From Here? *American Journal of Public Health*, 109(1), 35–40.
- Bohannon, R. W., & Crouch, R. (2019). 1-Minute Sit-to-Stand Test: SYSTEMATIC REVIEW OF PROCEDURES, PERFORMANCE, AND CLINIMETRIC PROPERTIES. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 39(1), 2–8.
- Briggs, A. M., Woolf, A. D., Dreinhöfer, K., Homb, N., Hoy, D. G., Kopansky-Giles, D., Åkesson, K., & March, L. (2018). Reducing the global burden of musculoskeletal conditions. *Bulletin of the World Health Organization*, 96(5), 366–368.
- Cady, K., Powis, M., & Hopgood, K. (2022). Intrarater and interrater reliability of the modified Thomas Test. *Journal of Bodywork and Movement Therapies*, 29, 86–91.

- Cambarri, B., Eastwood, P. R., Cecins, N. M., Thompson, P. J., & Jenkins, S. (2006). Six minute walk distance in healthy subjects aged 55-75 years. *Respiratory Medicine*, 100(4), 658–665.
- Campaniço, Helena (2016). Validade simultânea do questionário internacional de actividade física através da medição objectiva da actividade física por actigrafia proporcional. [Tese de Mestrado, Faculdade de Morticidade Humana, Universidade de Lisboa]. <https://www.repository.utl.pt/handle/10400.5/11866>
- Casanova, C., Celli, B. R., Barria, P., Casas, A., Cote, C., de Torres, J. P., Jardim, J., Lopez, M. V., Marin, J. M., Montes de Oca, M., Pinto-Plata, V., Aguirre-Jaime, A., & Six Minute Walk Distance Project (ALAT). (2011). The 6-min walk distance in healthy subjects: reference standards from seven countries. *The European Respiratory Journal: Official Journal of the European Society for Clinical Respiratory Physiology*, 37(1), 150–156.
- Cejudo, A., Sainz de Baranda, P., Ayala, F., & Santonja, F. (2015). Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players. *Physical Therapy in Sport: Official Journal of the Association of Chartered Physiotherapists in Sports Medicine*, 16(2), 107–113.
- Chetta, A., Zanini, A., Pisi, G., Aiello, M., Tzani, P., Neri, M., & Olivieri, D. (2006). Reference values for the 6-min walk test in healthy subjects 20-50 years old. *Respiratory Medicine*, 100(9), 1573–1578.
- Clapis, P. A., Davis, S. M., & Davis, R. O. (2008). Reliability of inclinometer and goniometric measurements of hip extension flexibility using the modified Thomas test. *Physiotherapy Theory and Practice*, 24(2), 135–141.
- Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, 35(8), 1381–1395.
- Cuenca-Garcia, M., Marin-Jimenez, N., Perez-Bey, A., Sánchez-Oliva, D., Camiletti-Moiron, D., Alvarez-Gallardo, I. C., Ortega, F. B., & Castro-Piñero, J. (2022). Reliability of Field-Based Fitness Tests in Adults: A Systematic Review. *Sports Medicine*, 52(8), 1961–1979.
- Dagenais, S., Caro, J., & Haldeman, S. (2008). A systematic review of low back pain cost of illness studies in the United States and internationally. *The Spine Journal: Official Journal of the North American Spine Society*, 8(1), 8–20.
- del Pozo-Cruz, B., Mocholi, M. H., del Pozo-Cruz, J., Parraca, J. A., Adsuar, J. C., & Gusi, N. (2014). Reliability and validity of lumbar and abdominal trunk muscle endurance tests in

- office workers with nonspecific subacute low back pain. *Journal of Back and Musculoskeletal Rehabilitation*, 27(4), 399–408.
- Demoulin, C., Vanderthommen, M., Duysens, C., & Crielaard, J.-M. (2006). Spinal muscle evaluation using the Sorensen test: a critical appraisal of the literature. *Joint, Bone, Spine: Revue Du Rhumatisme*, 73(1), 43–50.
- de Vet, H. C., Terwee, C. B., Ostelo, R. W., Beckerman, H., Knol, D. L., & Bouter, L. M. (2006). Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes*, 4, 54.
- de Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59(10), 1033–1039.
- de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Practical guides to biostatistics and epidemiology: Measurement in medicine: A practical guide: A practical guide*. Cambridge University Press.
- Direção-Geral da Saúde, Institute for Health Metrics and Evaluation. Portugal: The Nation's Health 1990–2016: An overview of the Global Burden of Disease Study 2016 Results. Seattle, WA: IHME, 2018.
- Durall, C. J., Greene, P. F., & Kernozek, T. W. (2012). A comparison of two isometric tests of trunk flexor endurance. *Journal of Strength and Conditioning Research / National Strength & Conditioning Association*, 26(7), 1939–1944.
- Enright, P. L. (2003). The six-minute walk test. *Respiratory Care*, 48(8), 783–785.
- Enright, P. L., & Sherrill, D. L. (1998). Reference equations for the six-minute walk in healthy adults. *American Journal of Respiratory and Critical Care Medicine*, 158(5 Pt 1), 1384–1387.
- Evans, K., Refshauge, K. M., & Adams, R. (2007). Trunk muscle endurance tests: reliability, and gender differences in athletes. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, 10(6), 447–455.
- Ferrari, S., Manni, T., Bonetti, F., Villafañe, J. H., & Vanti, C. (2015). A literature review of clinical tests for lumbar instability in low back pain: validity and applicability in clinical practice. *Chiropractic & Manual Therapies*, 23, 14.
- Fitzgerald, G. K., Wynveen, K. J., Rheault, W., & Rothschild, B. (1983). Objective assessment with establishment of normal values for lumbar spinal range of motion. *Physical Therapy*, 63(11), 1776–1781.

- Fritz, J. M., Whitman, J. M., & Childs, J. D. (2005). Lumbar spine segmental mobility assessment: an examination of validity for determining intervention strategies in patients with low back pain. *Archives of Physical Medicine and Rehabilitation*, 86(9), 1745–1752.
- Gabbe, B. J., Bennell, K. L., Wajswelner, H.; Finch, C. F. (2004). Reliability of common lower extremity musculoskeletal screening tests. , 5(2), 90–97. doi:10.1016/j.ptsp.2004.01.003
- Gatchel, R. J., Peng, Y. B., Peters, M. L., Fuchs, P. N., & Turk, D. C. (2007). The biopsychosocial approach to chronic pain: scientific advances and future directions. *Psychological Bulletin*, 133(4), 581–624.
- Gibbons, W. J., Fruchter, N., Sloan, S., & Levy, R. D. (2001). Reference values for a multiple repetition 6-minute walk test in healthy adults older than 20 years. *Journal of Cardiopulmonary Rehabilitation*, 21(2), 87–93.
- Greene, P. F., Durall, C. J., & Kernozek, T. W. (2012). Intersession reliability and concurrent validity of isometric endurance tests for the lateral trunk muscles. *Journal of sport rehabilitation*, 21(2), 161–166. <https://doi.org/10.1123/jsr.21.2.161>
- Harvey, D. (1998). Assessment of the flexibility of elite athletes using the modified Thomas test. *British Journal of Sports Medicine*, 32(1), 68–70.
- Hicks, G. E., Fritz, J. M., Delitto, A., & McGill, S. M. (2005). Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives of Physical Medicine and Rehabilitation*, 86(9), 1753–1762.
- Hicks, G. E., Fritz, J. M., Delitto, A., & Mishock, J. (2003). Interrater reliability of clinical examination measures for identification of lumbar segmental instability. *Archives of Physical Medicine and Rehabilitation*, 84(12), 1858–1864.
- Hodges, P. W., & Smeets, R. J. (2015). Interaction between pain, movement, and physical activity: Short-term benefits, long-term consequences, and targets for treatment. *Clinical Journal of Pain*, 31(2), 97–107. <https://doi.org/10.1097/AJP.0000000000000098>
- IPAQ-Research-Committee. Guidelines for data processing and analysis of the international physical activity questionnaire (Ipaq)-short and long forms. (2021)
- Ito, T., Shirado, O., Suzuki, H., Takahashi, M., Kaneda, K., & Strax, T. E. (1996). Lumbar trunk muscle endurance testing: an inexpensive alternative to a machine for evaluation. *Archives of Physical Medicine and Rehabilitation*, 77(1), 75–79.
- Juan-Recio, C., López-Plaza, D., Murillo, D. B., Pilar García-Vaquero, M., & Vera-García, F. J. (2018). Reliability assessment and correlation analysis of 3 protocols to measure trunk

muscle strength and endurance. *Journal of Sports Sciences*.
<https://doi.org/10.1080/02640414.2017.1307439>

- Kato, S., Murakami, H., Demura, S., Yoshioka, K., Shinmura, K., Yokogawa, N., Igarashi, T., Yonezawa, N., Shimizu, T., & Tsuchiya, H. (2019). Abdominal trunk muscle weakness and its association with chronic low back pain and risk of falling in older women. *BMC Musculoskeletal Disorders*, 20(1), 273.
- Keller, A., Hellesnes, J., & Jens I. Brox, A. (2001). Reliability of the Isokinetic Trunk Extensor Test, Biering-Sørensen Test, and Åstrand Bicycle Test: Assessment of Intraclass Correlation Coefficient and Critical Difference in Patients With Chronic Low Back Pain and Healthy Individuals. *Spine*, 26(7), 771.
- Keszei, A. P., Novak, M., & Streiner, D. L. (2010). Introduction to health measurement scales. *Journal of Psychosomatic Research*, 68(4), 319–323.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*.
<https://doi.org/10.1016/j.jcm.2016.02.012>
- Kraemer, H.C. and Thieman, S. (1987) *How Many Subjects? Statistical Power Analysis in Research*. Sage Publications, Newbury Park.
- Larsson, A., Palstam, A., Löfgren, M., Ernberg, M., Bjersing, J., Bileviciute-Ljungar, I., Gerdle, B., Kosek, E., & Mannerkorpi, K. (2015). Resistance exercise improves muscle strength, health status and pain intensity in fibromyalgia--a randomized controlled trial. *Arthritis Research & Therapy*, 17(1), 161.
- Latimer, J., Maher, C. G., Refshauge, K., & Colaco, I. (1999). The reliability and validity of the Biering-Sorensen test in asymptomatic subjects and subjects reporting current or previous nonspecific low back pain. *Spine*, 24(20), 2085–2089; discussion 2090.
- Liu, H., Shen, Y., Xiong, Y., Zhou, H., Mao, Y., Shen, Q., Hong, W., Liu, M., Liu, Y., Qiu, L., Zhang, Z., & Jia, Y. (2022). Psychometric Properties of Four Common Clinical Tests for Assessing Hamstring Flexibility in Young Adults. *Frontiers in Physiology*, 13, 911240.
- March, L., Smith, E. U. R., Hoy, D. G., Cross, M. J., Sanchez-Riera, L., Blyth, F., Buchbinder, R., Vos, T., & Woolf, A. D. (2014). Burden of disability due to musculoskeletal (MSK) disorders. *Best Practice & Research. Clinical Rheumatology*, 28(3), 353–366.
- Marques, A., Rebelo, P., Paixão, C., Almeida, S., Jácome, C., Cruz, J., & Oliveira, A. (2020). Enhancing the assessment of cardiorespiratory fitness using field tests. *Physiotherapy*, 109, 54–64.

- Martínez-Romero, M. T., Ayala, F., De Ste Croix, M., Vera-García, F. J., Sainz de Baranda, P., Santonja-Medina, F., & Sánchez-Meca, J. (2020). A Meta-Analysis of the Reliability of Four Field-Based Trunk Extension Endurance Tests. *International Journal of Environmental Research and Public Health*, 17(9). <https://doi.org/10.3390/ijerph17093088>
- Mayorga-Vega, D., Merino-Marban, R., & Viciano, J. (2014). Criterion-Related Validity of Sit-and-Reach Tests for Estimating Hamstring and Lumbar Extensibility: a Meta-Analysis. *Journal of Sports Science & Medicine*, 13(1), 1–14.
- McDermott, A., Bolger, C., Keating, L., McEvoy, L., & Meldrum, D. (2010). Reliability of three-dimensional gait analysis in cervical spondylotic myelopathy. *Gait and Posture*, 32(4), 552–558. <https://doi.org/10.1016/j.gaitpost.2010.07.019>
- McGill, S. M., Childs, A., & Liebenson, C. (1999). Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database. *Archives of Physical Medicine and Rehabilitation*, 80(8), 941–944.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica: Casopis Hrvatskoga Društva Medicinskih Biokemicara / HDMB*, 22(3), 276–282.
- Merkle, S. L., Sluka, K. A., & Frey-Law, L. A. (2020). The interaction between pain and movement. *Journal of hand therapy : official journal of the American Society of Hand Therapists*, 33(1), 60–66. <https://doi.org/10.1016/j.jht.2018.05.001>
- Mokkink, L. B., Boers, M., van der Vleuten, C. P. M., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2020). COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Medical Research Methodology*, 20(1), 293.
- Moseley, G. L., & Vlaeyen, J. W. S. (2015). Beyond nociception: the imprecision hypothesis of chronic pain. *Pain*, 156(1), 35–38.
- Moseng, T., Tveter, A. T., Holm, I., & Dagfinrud, H. (2014). Patients with musculoskeletal conditions do less vigorous physical activity and have poorer physical fitness than population controls: a cross-sectional study. *Physiotherapy*, 100(4), 319–324.
- Nijs, J., Van Houdenhove, B., & Oostendorp, R. A. B. (2010). Recognition of central sensitization in patients with musculoskeletal pain: Application of pain neurophysiology in manual therapy practice. *Manual Therapy*, 15(2), 135–141.
- Oliveira, M. J., Marçôa, R., Moutinho, J., Oliveira, P., Ladeira, I., Lima, R., & Guimarães, M. (2019). Reference equations for the 6-minute walk distance in healthy Portuguese subjects 18-70 years old. *Pulmonology*, 25(2), 83–89.

- Ozcan Kahraman, B., Salik Sengul, Y., Kahraman, T., & Kalemci, O. (2016). Developing a Reliable Core Stability Assessment Battery for Patients with Nonspecific Low Back Pain. *Spine*, 41(14), E844–E850.
- Quitério, N. (2022). Valores de referência de testes funcionais para informar a prática clínica da fisioterapia. ESSUA, Dissertações de mestrado.
- Rabin, A., Shashua, A., Pizem, K., & Dar, G. (2013). The interrater reliability of physical examination tests that may predict the outcome or suggest the need for lumbar stabilization exercises. *The Journal of Orthopaedic and Sports Physical Therapy*, 43(2), 83–90.
- Radtke, T., Puhan, M. A., Hebestreit, H., & Kriemler, S. (2016). The 1-min sit-to-stand test—A simple functional capacity test in cystic fibrosis? *Journal of Cystic Fibrosis: Official Journal of the European Cystic Fibrosis Society*, 15(2), 223–226.
- Reiman MP, Krier AD, Nelson JA, Rogers MA, Stuke ZO, Smith BS. Comparison of different trunk endurance testing methods in college-aged individuals. *Int J Sports Phys Ther*. 2012 Oct;7(5):533-9. PMID: 23091786; PMCID: PMC3474305.
- Rezvani, A., Ergin, O., Karacan, I., & Oncu, M. (2012). Validity and reliability of the metric measurements in the assessment of lumbar spine motion in patients with ankylosing spondylitis. *Spine*, 37(19), E1189–E1196.
- Rice, D., Nijs, J., Kosek, E., Wideman, T., Hasenbring, M. I., Koltyn, K., Graven-Nielsen, T., & Polli, A. (2019). Exercise-Induced Hypoalgesia in Pain-Free and Chronic Pain Populations: State of the Art and Future Directions. *The Journal of Pain: Official Journal of the American Pain Society*, 20(11), 1249–1266.
- Ritchie, C., Trost, S. G., Brown, W., & Armit, C. (2005). Reliability and validity of physical fitness field tests for adults aged 55 to 70 years. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, 8(1), 61–70.
- Scholtes, V. A., Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury*, 42(3), 236–240.
- Sebbag, E., Felten, R., Sagez, F., Sibilila, J., Devilliers, H., & Arnaud, L. (2019). The world-wide burden of musculoskeletal diseases: a systematic analysis of the World Health Organization Burden of Diseases Database. *Annals of the Rheumatic Diseases*, 78(6), 844–848.
- Segura-Ortí, E., & Martínez-Olmos, F. J. (2011). Test-retest reliability and minimal detectable change scores for sit-to-stand-to-sit tests, the six-minute walk test, the one-leg heel-rise test, and handgrip strength in people undergoing hemodialysis. *Physical Therapy*, 91(8), 1244–1252.

- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Simmonds, M. J., Olson, S. L., Jones, S., Hussein, T., Lee, C. E., Novy, D., & Radwan, H. (1998). Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine*, 23(22), 2412–2421.
- Sterling, M., Jull, G., & Wright, A. (2001). The effect of musculoskeletal pain on motor activity and control. *The journal of pain*, 2(3), 135–145. <https://doi.org/10.1054/jpai.2001.19951>
- Strassmann, A., Steurer-Stey, C., Lana, K. D., Zoller, M., Turk, A. J., Suter, P., & Puhan, M. A. (2013). Population-based reference values for the 1-min sit-to-stand test. *International Journal of Public Health*, 58(6), 949–953.
- Troosters, T., Gosselink, R., & Decramer, M. (1999). Six minute walking distance in healthy elderly subjects. *The European Respiratory Journal: Official Journal of the European Society for Clinical Respiratory Physiology*, 14(2), 270–274.
- Tveter, A. T., Dagfinrud, H., Moseng, T., & Holm, I. (2014). Measuring health-related physical fitness in physiotherapy practice: reliability, validity, and feasibility of clinical field tests and a patient-reported measure. *The Journal of Orthopaedic and Sports Physical Therapy*, 44(3), 206–216.
- Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abolhassani, H., Aboyans, V., Abrams, E. M., Abreu, L. G., Abrigo, M. R. M., Abu-Raddad, L. J., Abushouk, A. I., ... Murray, C. J. L. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258), 1204–1222.
- Wakefield, C. B., Halls, A., Difilippo, N., & Cottrell, G. T. (2015). Reliability of goniometric and trigonometric techniques for measuring hip-extension range of motion using the modified Thomas test. *Journal of Athletic Training*, 50(5), 460–466.
- Williams, A., Kamper, S. J., Wiggers, J. H., O'Brien, K. M., Lee, H., Wolfenden, L., Yoong, S. L., Robson, E., McAuley, J. H., Hartvigsen, J., & Williams, C. M. (2018). Musculoskeletal conditions may increase the risk of chronic disease: a systematic review and meta-analysis of cohort studies. *BMC Medicine*, 16(1), 167.

APÊNDICES

Apêndice 1 – Panfleto Redes Sociais



É SAUDÁVEL?

NÃO APRESENTA RESTRIÇÕES PARA A PRÁTICA DE EXERCÍCIO FÍSICO?

TEM ENTRE 18 E 65 ANOS?

VIVE NA ZONA METROPOLITANA DE LISBOA?



VENHA PARTICIPAR NESTE ESTUDO!

Avaliação da fiabilidade de uma bateria de testes para avaliação da aptidão física em indivíduos saudáveis - The MyBack Physical Performance Battery

Deixe o seu contacto

<https://docs.google.com/forms/d/e/1FAIpQLSfqpFDabY2e-dVeOq-wYQ6yEEiUnOimuOQYkF29onLZNc4v11Q/viewform>

Entre em contacto connosco!



210512001@estudantes.ips.pt
210512004@estudantes.ips.pt

Apêndice 2 – Formulário Online

Avaliação da fiabilidade de uma bateria de testes para avaliação da aptidão física em indivíduos que recuperaram de lombalgia e em indivíduos saudáveis – The MyBack Physical Performance Battery

O estudo realiza-se no âmbito da Unidade Curricular de “Relatório de Investigação” integrado no 2º ano do Curso de Mestrado em Fisioterapia em Condições Músculo-Esqueléticas, lecionado pela Escola Superior de Saúde do Instituto Politécnico de Setúbal em parceria com a Nova Medical School/ Faculdade de Ciências Médicas e a Escola Nacional de Saúde Pública da Universidade Nova de Lisboa, a realizar pelas discentes Bruna Costa e Matilde Serra sob orientação científica da professora Rita Fernandes e coorientação do Professor Diogo Moço

Se tem interesse em obter mais informações sobre o estudo, preencha o questionário abaixo. A partilha do seu contacto não significa integração direta no estudo, e poderá mudar de opinião a qualquer momento.

O seu contacto e informações pessoais serão mantidas em confidencialidade e serão somente utilizadas no contexto deste estudo.

Contacto Telefónico *

A sua resposta

Email *

A sua resposta

*

Confirmo que li e entendi as informações acima

Enviar

Limpar formulário

Apêndice 3 – Carta Explicativa do Estudo



INSTITUTO POLITÉCNICO DE SETÚBAL - ESCOLA SUPERIOR DE SAÚDE

DEPARTAMENTO DE FISIOTERAPIA

Avaliação da fiabilidade de uma bateria de testes para avaliação da aptidão física em indivíduos que recuperaram de lombalgia e em indivíduos saudáveis – The MyBack

Physical Performance Battery

Costa, B., Serra, M., Moço, D., Fernandes, R.

CARTA EXPLICATIVA DO ESTUDO AOS PARTICIPANTES

Gostaríamos de convidá-lo(a) a participar neste estudo. Antes de tomar qualquer decisão, é importante que compreenda as razões pelas quais este estudo está a ser conduzido e o nível de envolvimento que lhe é pedido. Por favor, utilize o tempo que necessitar para ler a informação que se segue. Poderá falar com outras pessoas sobre este estudo, se o desejar.

Este documento inclui duas partes: a parte 1 apresenta-lhe a informação sobre o propósito deste estudo e o nível de envolvimento que lhe será pedido; a parte 2 oferece-lhe informação mais detalhada sobre a forma como o estudo será conduzido.

Se algum aspeto não for claro ou se desejar mais informações por favor não hesite em colocar-nos as suas questões. Utilize o tempo que precisar para decidir se deseja ou não participar neste estudo.

Parte 1 | O propósito do estudo e o nível de envolvimento que lhe é pedido

Qual é a finalidade do estudo?

Este estudo tem como objetivo avaliar a fiabilidade intra-avaliador de uma bateria de testes de aptidão física em dois grupos: indivíduos que recuperaram de lombalgia (há menos de 3 meses) e indivíduos saudáveis. Isto é, perceber se os testes aplicados (que são simples e de fácil aplicação na prática clínica) são fiáveis (produzindo resultados consistentes quando realizados em 2 momentos pelo mesmo avaliador). Este conhecimento poderá permitir uma avaliação fiável da aptidão física dos indivíduos após um episódio de lombalgia, permitindo caracterizar melhor as pessoas que recuperaram desta condição e, potencialmente, selecionar o tipo de exercício mais

indicado para cada pessoa, de acordo com os resultados da avaliação da(s) componente(s) física(s) mais afetada(s).

Porque fui convidado?

Foi convidado pelas razões enunciadas num dos seguintes pontos:

- ✓ Foi convidado por ter recuperado recentemente de um episódio de lombalgia e por se enquadrar nos critérios de elegibilidade definidos para o estudo, estando integrado nos seguintes contextos: serviços de fisioterapia no setor privado onde se encontram os fisioterapeutas colaboradores do projeto, ou é utente de uma das USF do ACES Arrábida onde decorre o projeto SPLIT; ou por se ter voluntariado através dos meios digitais para colaborar neste projeto.

- ✓ Foi convidado por ser uma pessoa saudável (sem qualquer condição músculo-esquelética, neurológica, pulmonar ou cardíaca diagnosticada), enquadrando-se nos critérios de elegibilidade do estudo, e por frequentar/pertencer ao campus do IPS ou por se ter voluntariado através dos meios digitais para colaborar neste projeto.

A sua participação irá ajudar-nos a determinar se será possível confiar num conjunto de testes de aptidão física para avaliar as competências em 4 domínios da performance física: capacidade cardiorrespiratória, força de resistência do tronco e membros inferiores, flexibilidade e controlo neuromuscular.

Tenho mesmo de participar?

A escolha de participar ou não no estudo é sua, e é voluntária. O presente estudo não acarreta qualquer risco, não trazendo também qualquer vantagem direta para os que nele participam. Se decidir participar no estudo, poderá abandonar o mesmo a qualquer momento, sem que tenha de o justificar, nem que saia prejudicado.

O que acontece se aceitar participar?

Se aceitar participar, ser-lhe-ão entregues dois exemplares do formulário de consentimento informado que deverá ler atentamente, preencher e assinar (um exemplar ficará na sua posse). O estudo e os respetivos procedimentos encontram-se descritos ao longo desta carta explicativa, terá o tempo que necessitar para a ler e colocar questões. De seguida será solicitado o seu consentimento informado.

O que terei de fazer?

Terá de responder a um questionário de caracterização sociodemográfica e a um questionário de atividade física. O tempo estimado de preenchimento de ambos é de 10 minutos, embora possa utilizar o tempo que necessitar para o efeito. Terá de realizar um conjunto de 11 testes de aptidão física (com uma duração aproximada de 20 minutos) em 2 momentos (dias) diferentes, que contemplam a avaliação da capacidade cardiorrespiratória, flexibilidade, força muscular e controlo neuromuscular.

Quais as possíveis vantagens em participar?

Não lhe podemos garantir que este estudo o(a) ajude de alguma forma. Contudo, podemos garantir-lhe que a informação que retirarmos dele poderá ajudar-nos a encontrar testes físicos fiáveis para caracterizar, do ponto de vista da aptidão física, indivíduos após recuperarem de um episódio de lombalgia e indivíduos saudáveis.

No futuro, este conhecimento, permitindo caracterizar melhor os indivíduos que recuperaram desta condição, poderá possibilitar selecionar o tipo de exercício mais indicado para cada utente de acordo com os resultados da avaliação da(s) componente(s) física(s) mais afetada(s). Estes resultados poderão permitir definir programas de prevenção de recorrência individualizados, melhorando os cuidados de saúde e reduzindo os custos associados a esta condição, e poderão ainda permitir caracterizar os indivíduos saudáveis relativamente aos parâmetros de aptidão física identificados.

Quais as possíveis desvantagens ou riscos se aceitar participar?

Não são esperadas quaisquer implicações negativas para os participantes neste estudo. Se, por alguma razão, se sentir prejudicado, poderá comunicar com os responsáveis e/ou abandoná-lo a qualquer momento sem necessidade de fornecer qualquer justificação.

E se houver algum problema?

Qualquer queixa que tenha sobre este estudo, sobre a forma como foi abordado(a) ou qualquer dano associado serão considerados. Na parte 2 deste documento, poderá encontrar mais informação sobre este aspeto.

A minha participação neste estudo será confidencial?

Sim. Serão adotados um conjunto de procedimentos de natureza ética de forma a assegurar que a sua participação será mantida em confidencialidade. Todo o material recolhido será codificado e tratado de forma anónima e confidencial, sendo conservado à responsabilidade de Bruna Alexandra Lopes Pratas Costa e Matilde Marques Serra.

Os resultados do estudo serão apresentados no âmbito da apresentação do “Relatório de Investigação” do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas, e em nenhum

momento os dados serão apresentados de forma individual. Os dados originais serão destruídos após a disseminação do estudo.

Se a informação anterior lhe despertou interesse em participar no estudo, por favor leia a informação adicional (parte 2) antes de tomar qualquer decisão.

Parte 2 | Condução do estudo

O que acontece se não aceitar participar no estudo?

Uma vez que a sua participação é totalmente voluntária, é livre de desistir do estudo a qualquer momento, sem que tenha de o justificar, não tendo qualquer impacto na sua vida atual ou no futuro. Se desistir do estudo não serão utilizados quaisquer dados que lhe digam respeito.

E se pretender algum esclarecimento ou se houver algum problema?

No caso de necessitar de algum esclarecimento sobre qualquer aspeto deste estudo, deverá contactar um membro da equipa de investigação. A sua dúvida será esclarecida e iremos responder às suas questões. Poderá contactar-nos através dos seguintes *e-mails*: 210512001@estudantes.ips.pt; 210512004@estudantes.ips.pt ou sensemotionlab@gmail.com.

Se pretender desejar fazer uma reclamação, poderá contactar o responsável pela unidade curricular “Relatório de Investigação” via *e-mail* (eduardo.cruz@ess.ips.pt) ou a CE-IPS (comissao.etica@ips.pt).

A minha participação no estudo será confidencial?

Sim, a sua participação será confidencial, assegurada através do uso de vários procedimentos de natureza ética. As suas respostas aos questionários, resultados dos testes e dados sociodemográficos e clínicos serão codificados e introduzidos pelo investigador responsável por esta tarefa numa base de dados, sem que sejam referenciados nomes ou outros dados que o identifiquem.

O que irá acontecer às informações que der sobre mim?

Os dados recolhidos serão agregados e nunca serão apresentados de forma individual, sendo a sua finalidade apenas caracterizar os participantes deste estudo e, através de cálculos estatísticos perceber o quão fiáveis são os testes realizados, isto é a sua capacidade para fornecer dados semelhantes quando aplicados ao mesmo indivíduo pelo mesmo avaliador (fiabilidade intra-avaliador). Ser-lhe-á atribuído um código no momento da avaliação, garantindo o anonimato no seu armazenamento. Os dados originais serão destruídos após a disseminação do estudo.

O que irá acontecer com os resultados deste estudo?

Os resultados serão apresentados no âmbito da Unidade Curricular “Relatório de Investigação” integrada no Mestrado em Fisioterapia em Condições Músculo-Esqueléticas, nunca sendo expostos de forma individual. Os resultados poderão vir a ser apresentados/ publicados em conferências/ revistas da especialidade, garantindo a impossibilidade de individualizar as respostas de cada participante. Uma vez apresentados os resultados, os dados originais serão destruídos após a respetiva disseminação pública através de dissertação, artigos científicos e apresentações em congressos ou conferências.

Obrigado pela leitura atenta deste documento.

Contactos:

Bruna Alexandra Lopes Pratas Costa

Email: 210512001@estudantes.ips.pt

Telemóvel: 910626477

Matilde Marques Serra

Email: 210512004@estudantes.ips.pt

Telemóvel: 917383727

Apêndice 4 – Consentimento Informado



Formulário de consentimento informado

É convidado(a) a participar num estudo no âmbito da Unidade Curricular de “Relatório de Investigação”, integrado no 2º ano do Curso de Mestrado em Fisioterapia em Condições Músculo-Esqueléticas, lecionado pela Escola Superior de Saúde do Instituto Politécnico de Setúbal em parceria com a Nova Medical School/ Faculdade de Ciências Médicas e a Escola Nacional de Saúde Pública da Universidade Nova de Lisboa, a realizar pelas discentes Bruna Costa e Matilde Serra, sob orientação científica da professora doutora Rita Fernandes e coorientação do professor Diogo Moço.

Os objetivos deste estudo são avaliar a fiabilidade intra e inter-avaliador, de um conjunto de testes de aptidão física em indivíduos que recuperaram de lombalgia e em indivíduos saudáveis.

Fui selecionado(a) porque:

Recuperei recentemente de um episódio de lombalgia não específica ou sou saudável;

Sou capaz de ler e falar em português;

Tenho um telemóvel capaz de receber e enviar mensagens de texto;

Não tenho nenhuma contraindicação médica para a realização de uma bateria de testes de avaliação da aptidão física (The MyBack Physical Performance Battery).

Tenho entre 18 e os 65 anos;

Confirmando que irei responder a um questionário relativo às minhas características sociodemográficas, a um questionário para classificar o meu nível de atividade física, e que irei realizar uma bateria de testes para avaliação da aptidão física em 3 momentos diferentes. Foram-me explicados todos os procedimentos a realizar e sei onde comparecer nos dias/horários agendados com os investigadores.

Compreendo igualmente que a participação no estudo não acarreta qualquer tipo de vantagens e/ou desvantagens potenciais. A minha identidade será preservada, através de um sistema de codificação o que permitirá que o estudo funcione em anonimato, ou seja, a equipa de investigadores que analisa os dados não terá acesso à minha identificação e todos os dados serão utilizados apenas para fins científicos. Apenas os investigadores terão acesso à sequência e código que me será atribuído. Todos os documentos em formato de papel serão guardados num arquivo seguro no IPS sob a responsabilidade do investigador principal, sendo destruídos após a disseminação do estudo.

Reconheço que não são esperados quaisquer riscos ou implicações negativas acerca da minha participação no estudo. Sei também que poderei contactar os investigadores do estudo a qualquer altura.

Por último, declaro saber que as minhas respostas serão apresentadas no âmbito da apresentação do "Relatório de Investigação" do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas, mas nunca de forma individual e que os resultados poderão vir a ser apresentados/publicados em conferências/ revistas da especialidade.

Mais uma vez, declaro que li a informação acima descrita, foi-me dada a oportunidade de colocar questões relacionadas com o estudo e que qualquer questão que coloquei foi esclarecida de forma clara.

Fui informado(a) que tenho direito a recusar participar e que a minha recusa em o fazer não terá consequências para mim. Compreendo que tenho o direito de colocar agora e durante o desenvolvimento do estudo, qualquer questão relacionada com o mesmo. Compreendo que sou livre de, a qualquer momento, abandonar o estudo sem ter de fornecer qualquer explicação.

Declaro que li e compreendi toda a informação facultada na ficha informativa e que pude esclarecer todas as minhas dúvidas com os investigadores

Declaro que aceito participar nesta investigação, com a salvaguarda da confidencialidade e anonimato e sem prejuízo pessoal de cariz ético e moral.

Assim, dou o meu Consentimento Informado relativamente à minha participação nesta investigação, com a salvaguarda da confidencialidade e anonimato e sem prejuízo pessoal de cariz ético ou moral.

Nome do Participante: _____

Assinatura do Participante: _____

Data: __/__/____

(dia/mês/ano)

Assinatura do Investigador

Apêndice 5 – Registo de Avaliações.



FORMULÁRIO DE REGISTO DA AVALIAÇÃO DA APTIDÃO FÍSICA (T0)

Código Completo

DATA: ___/___/_____

(Após a verificação dos Critérios de Inclusão e assinatura do Consentimento Informado)

Tempo de realização dos testes previsto: 60 minutos

Registar os resultados dos seguintes testes:

	Componentes de aptidão física	T0
Capacidade cardiorrespiratória	<i>6-Min Walk Test</i>	
Força de Resistência do tronco e dos membros inferiores	<i>Trunk Flexor Test</i>	
	<i>Side Bridge direita</i>	
	<i>Side Bridge esquerda</i>	
	<i>Biering-Sørensen Test</i>	
	<i>60s Sit to Stand</i>	
Flexibilidade	<i>Sit-and-Reach test</i>	
	Schober test (Flexão)	
	Schober test (Extensão)	
	Modified Thomas Test	

		Testes	Negativo	Positivo	Movimento(s) identificado(s)
T0	Controlo motor	<i>Aberrant Movement pattern</i>			Arco de dor em flexão
					Arco de dor em extensão
					"Gower sign"
					"Instability catch sign"
	Prone Instability test				Alteração do ritmo lombo-pélvico

FORMULÁRIO DE REGISTO DA AVALIAÇÃO DA APTIDÃO FÍSICA (T1)

Código Completo

DATA: ___/___/_____

(Após a verificação dos Critérios de Inclusão e assinatura do Consentimento Informado)

Tempo de realização dos testes previsto: 60 minutos

Registar os resultados dos seguintes testes:

	Componentes de aptidão física	T0
Capacidade cardiorrespiratória	<i>6-Min Walk Test</i>	
Força de Resistência do tronco e dos membros inferiores	<i>Trunk Flexor Test</i>	
	<i>Side Bridge</i> direita	
	<i>Side Bridge</i> esquerda	
	<i>Biering-Sørensen Test</i>	
	<i>60s Sit to Stand</i>	
Flexibilidade	<i>Sit-and-Reach test</i>	
	Schober test (Flexão)	
	Schober test (Extensão)	
	Modified Thomas Test	

		Testes	Negativo	Positivo	Movimento(s) identificado(s)
T0	Controlo motor	<i>Aberrant Movement pattern</i>			Arco de dor em flexão
					Arco de dor em extensão
					"Gower sign"
					"Instability catch sign"
					Alteração do ritmo lombo-pélvico
		Prone Instability test			

ANEXOS

Anexo 1 – Parecer da Comissão Especializada de Ética em Investigação



Comissão de Ética

Identificação do documento: CE-IPS n.º 47/2023

Título do projeto:

Avaliação da fiabilidade de uma bateria de testes para avaliação da aptidão física em indivíduos que recuperaram de lombalgia e em indivíduos saudáveis – The MyBack Physical Performance Battery

Investigadores principais:

Bruna Alexandra Lopes Pratas Costa (Estudante do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas)

Matilde Marques Serra (Estudante do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas)

Equipa de investigação:

Orientadores: Professora Doutora Rita Fernandes e Professor Diogo Moço.

Unidade Orgânica do IPS: Escola Superior de Saúde.

Outras Unidades/Participantes: Nova Medical School/Faculdade de Ciências Médicas

Parcerias: USF do ACES Arrábida e Unidades de Saúde Privadas (ainda a seleccionar).

ANÁLISE E JUSTIFICAÇÃO DO PARECER

Documentos recebidos:

- Requerimento dirigido à Presidente da Comissão de Ética;
- Dossier de suporte ao estudo incluindo:
 - Sinopse do estudo;
 - Manual de recrutamento de participantes (apêndice 1)
 - Termo de informação e consentimento para tratamento de dados pessoais (apêndice 2)
 - Panfleto de Divulgação do Estudo (apêndice 3)
 - Carta explicativa do Estudo aos participantes (apêndice 4)
 - Formulário de Consentimento informado (apêndice 5)
 - Caderno de instrumentos (questionários e testes) (apêndice 6)
 - Descrição dos Testes de Aptidão Física (apêndice 7)
 - Cronograma (apêndice 8)

- Declaração de ausência de conflito de interesses e incompatibilidades dos investigadores principais;
- Declaração/Termo de responsabilidade dos investigadores principais.
- CV dos investigadores principais.

- Parecer favorável relativo ao registo das atividades de recolha e tratamento de dados do IPS (RGPD - Art.º 30.º e 49.º)

Análise e Justificação do Parecer:

1. O Estudo é proposto no âmbito do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas da Escola Superior de Saúde do Instituto Politécnico de Setúbal (ESS-IPS), lecionado em parceria pela Nova Medical School/Faculdade de Ciências Médicas e Escola Nacional de Saúde Pública, da Universidade Nova de Lisboa.
2. O Estudo tem como objetivo “avaliar a fiabilidade (teste reteste, intra-avaliador e inter-avaliador) de uma bateria de testes para avaliação da aptidão física em indivíduos que recuperam recentemente de lombalgia e em indivíduos saudáveis”.
3. A equipa tem como investigadores responsáveis as estudantes do Mestrado em Fisioterapia em Condições Músculo-Esqueléticas Bruna Alexandra Lopes Pratas Costa e Matilde Marques Serra e tem a orientação científica da Professora Rita Fernandes e do Professor Diogo Moço.
4. Trata-se, a nível metodológico, de um estudo prospetivo com aplicação de uma bateria de testes (tipo teste-reteste intra e inter avaliador) que serão aplicados a “duas amostras não-probabilísticas, por conveniência: uma amostra de indivíduos que recuperam de lombalgia há menos de 3 meses e uma amostra de indivíduos saudáveis”.
5. Os critérios de inclusão e exclusão para cada grupo participante no estudo estão claramente definidos no estudo, assim como o número a recrutar (23 em cada grupo). Os indivíduos com lombalgia serão convidados a participar através das unidades de saúde públicas e privadas das áreas de Setúbal e metropolitana de Lisboa, assim como de Coimbra e Condeixa a Nova, sendo o convite feito pelos fisioterapeutas colaboradores dessas mesmas unidades. Os indivíduos saudáveis serão convidados a participar (das mesmas áreas territoriais) através de meios digitais, podendo responder positivamente ao convite através de um link colocado nas respetivas páginas de divulgação, através do qual também terão acesso aos objetivos e às características do estudo
6. Os potenciais participantes que preencham os critérios de inclusão no estudo serão convidados a participar através de contacto por telemóvel (preferencialmente) ou email, sendo-lhes “explicado o objetivo do estudo verbalmente e por escrito e as possíveis vantagens/riscos, bem como todos os procedimentos realizados”, assim como tomarão conhecimento da possibilidade de desistir do estudo a qualquer momento, sem que exista qualquer prejuízo. É solicitado a cada participante a assinatura de dois exemplares do formulário de consentimento informado (Apêndice 5 – Consentimento Informado), a fim de serem integrados no estudo.
7. São facultadas indicações claras aos participantes sobre o tempo que demorará a aplicação de cada instrumento de pesquisa (questionário e testes)
8. É referido o procedimento a adotar relativamente à confidencialidade e o anonimato dos dados dos participantes, haverá designadamente a atribuição de um código que permita a associação entre os o participante e os diferentes momentos de testagem em que será avaliado. É também mencionado que os dados recolhidos (questionários e registo dos testes) ficam ao encargo da equipa de investigadores e serão guardados em local seguro e de acesso restrito na ESS-IPS, assim como que destruição dos dados ocorrerá após a disseminação do estudo (num prazo máximo de 5 anos).
9. Conforme é descrito no estudo a “análise dos dados será realizada através do programa *IBM SPSS Statistics* ©, versão 28”. Será ainda utilizada “a estatística descritiva para as variáveis qualitativas, distribuição de frequência para as variáveis nominais e ordinais e cálculo de medidas de tendência central e de dispersão para as variáveis numéricas. Para avaliar a

fiabilidade (teste reteste, intra e inter-avaliador) dos testes com variáveis numéricas calcular-se-á o coeficiente de correlação intra-classe (de Vet et al., 2006a) e para os testes com variáveis categóricas utilizar-se-á o coeficiente de *Kappa* de *Cohen* ponderado (de Vet et al., 2006b). Serão também calculados os valores médios das diferenças entre as medições, os seus intervalos de confiança e o desvio padrão das diferenças (DPdif). Calcular-se-ão ainda os limites da concordância de Bland e Altman, o erro padrão de medida (EPM) (Vet et al., 2006a) e a diferença mínima detetável (DMD) (de Vet et al., 2006b)”

10. É facultada uma carta explicativa do estudo com o contacto dos investigadores (estudantes). É facultada uma declaração de consentimento informado, com espaço para assinatura do participante e dos investigadores, sendo esta disponibilizada em duplicado. Na declaração de consentimento informado não constam os contactos dos estudantes, do(s) orientador(es) da investigação ou da comissão de ética.

Parecer:

A CE-IPS considera que o estudo preenche requisitos éticos, com preocupações relativas à proteção dos direitos dos participantes, emitindo parecer favorável. Sugere-se que a capa do estudo inclua a designação dos nomes completos das estudantes e dos respetivos orientadores.

Relator/a: Carla Cibebe Figueiredo

Aprovação CE-IPS a 17 de abril de 2023

Assinado por: **SÓNIA ALEXANDRA PAIVA DOS SANTOS**
Num. de Identificação: 10091202
Data: 2023.04.17 18:56:06+01'00'

Vice-presidente da CE-IPS

Anexo 2 – Questionário Sociodemográfico e Ocupacional



Questionário de caracterização sociodemográfica e clínica



Código Completo

QUESTIONÁRIO DE CARACTERIZAÇÃO CLÍNICA E SOCIODEMOGRÁFICA

1. Idade _____
2. Sexo: Masculino Feminino
3. Peso: _____
4. Altura: _____
5. Quais são as suas **Habilitações Literárias** (escolha uma das seguintes opções)

Ensino Primário ou <input type="checkbox"/> inferior	Ensino Básico Completo <input type="checkbox"/> (9º ano de escolaridade)	Ensino Secundário ou <input type="checkbox"/> equivalente completo (12º ano de escolaridade)
Ensino Superior <input type="checkbox"/> completo		
13. Se não, qual a sua **situação profissional atual**?

Trabalhador(a) e/ou estudante <input type="checkbox"/>	Doméstico(a) <input type="checkbox"/>
Desempregado(a) <input type="checkbox"/>	Incapaz de trabalhar <input type="checkbox"/>
Reformado(a) <input type="checkbox"/>	
14. Selecione **todos** os atributos que melhor descrevem a sua **atividade profissional atual**:

Predominantemente em pé <input type="checkbox"/>	Trabalho manual com cargas inferiores a 11kg (levantar/carregar/mover) <input type="checkbox"/>
Predominantemente sentado <input type="checkbox"/>	Trabalho manual com cargas superiores a 11kg (levantar/carregar/mover) <input type="checkbox"/>
Predominantemente na mesma posição <input type="checkbox"/>	Trabalho manual com repetição dos mesmos movimentos <input type="checkbox"/>
Predominantemente agachado, ajoelhado e/ou subir/descer escadas <input type="checkbox"/>	Levantar/mover pessoas <input type="checkbox"/>

15. O seu trabalho implica uma velocidade elevada?

- Sempre Quase sempre Cerca de ¾ do tempo
Cerca de metade do tempo Cerca de ¼ do tempo Quase nunca
Nunca

16. Com que frequência realiza tarefas que envolvam posições estranhas (ex: costas dobradas ou torcidas)?

- Nunca ou raramente Ocasionalmente Frequentemente

17. Para cada uma das seguintes afirmações, selecione a resposta que melhor descreve a sua situação laboral, sendo 1 "sempre", 2 "a maior parte das vezes", 3 "algumas vezes", 4 "raramente" e 5 "nunca"

Os/As seus/suas colegas ajudam-no/a e apoiam-no/a	1	2	3	4	5
O/A seu/sua chefe apoia-o/a e ajuda-o/a	1	2	3	4	5
Pode fazer uma pausa quando desejar	1	2	3	4	5
Tem tempo suficiente para terminar o seu trabalho	1	2	3	4	5
Tem a sensação de realizar um trabalho útil	1	2	3	4	5
Sabe o que é esperado de si no trabalho	1	2	3	4	5
É tratado/a de forma justa no seu local de trabalho	1	2	3	4	5
Sente stress no seu trabalho	1	2	3	4	5
O seu trabalho requer que não mostre os seus sentimentos	1	2	3	4	5

Muito insatisfeito 0 1 2 3 4 5 6 7 8 9 10 Muito satisfeito

18. O seu trabalho afeta a sua saúde?

- Não Sim, de forma maioritariamente positiva Sim, de forma maioritariamente negativa
Ausência de stress 0 1 2 3 4 5 6 7 8 9 10 Stress extremo

19. Sente que tem todo o apoio necessário da sua família ou amigos quando necessita?

- Concordo Nem concordo nem discordo Discordo

20. Na sua opinião, quão grande é o risco de ter uma recorrência de lombalgia durante os próximos 12 meses?

Ausência de risco	0	1	2	3	4	5	6	7	8	9	10	Risco máximo
-------------------	---	---	---	---	---	---	---	---	---	---	----	--------------

21. Em média, quantas horas dorme por noite? (pense na última semana) _____

22. Nos últimos 12 meses, com que frequência teve algum destes problemas relacionados com o sono?
Sendo 1 "todos os dias", 2 "várias vezes por semana", 3 "várias vezes por mês", 4 "com menos frequência" e 5 "Nunca"

Dificuldade em adormecer	1	2	3	4	5
Acordar várias vezes durante o sono	1	2	3	4	5
Acordar com uma sensação de cansaço e fadiga	1	2	3	4	5

23. Em média, quantas horas está sentado por dia, incluindo enquanto viaja, trabalha ou em atividades de lazer? (pense na última semana) _____

24. Como descreve os seus hábitos tabágicos?

Nunca fui fumador Sou ex-fumador Sou fumador ocasional Sou fumador regular

25. É consumidor(a) de bebidas alcoólicas?

Não Sou consumidor(a) ocasional Sou consumidor(a) regular

26. Tem alguma doença ou problema de saúde que se tenha prolongado, ou que preveja prolongar-se por mais de 6 meses?

Não Sim

27. Na última semana, quantos dias fez pelo menos 30 minutos de atividade física? _____

28. No global, diria que a sua saúde é:

Má Razoável Boa Excelente

29. Na última semana, quão em baixo se sentiu?

Nada Ligeiramente Moderadamente Muito

30. Na última semana, quão ansioso se sentiu

Nada Ligeiramente Moderadamente Muito

Questionário cedido pelo Projeto Myback, financiado por fundos nacionais através da FCT – Fundação para a Ciência e Tecnologia. I.P., no âmbito do Projeto MyBack (REF. PTDC/SAL-SER/7406/2020)".

Anexo 3 – IPAQ Versão Curta



Código Completo

Questionário Internacional de Aptidão Física - IPAQ

Estamos interessados em conhecer os níveis de actividade física habitual dos Portugueses. As suas respostas vão ajudar-nos a compreender o quanto activos somos.

As questões referem-se ao tempo que dispense na actividade física numa semana. Este questionário inclui questões acerca de actividades que faz no trabalho, para se deslocar de um lado para outro, actividades referentes à casa ou ao jardim e actividades que efectua no seu tempo livre para entretenimento, exercício ou desporto.

As suas respostas são importantes. Por favor responda a todas as questões mesmo que não se considere uma pessoa activa.

Obrigado pela sua participação

Actividade física vigorosa refere-se a actividades que requerem muito esforço físico e tornam a respiração muito mais intensa que o normal.

Actividade física moderada refere-se a actividades que requerem esforço físico moderado e toma a respiração um pouco mais intensa que o normal.

Ao responder às questões considere apenas as actividades físicas que realize durante pelo menos 10 minutos seguidos.

Ao responder às seguintes questões considere o seguinte:

1a Habitualmente, por semana, quantos dias faz actividades físicas **vigorosas** como levantar e/ou transportar objectos pesados, cavar, ginástica aeróbica ou andar de bicicleta a uma velocidade acelerada?

____ dias por semana
____ Nenhum (passe para a questão **2a**)

1b Quanto tempo costuma fazer actividade física vigorosa por dia?

____ horas ____ minutos

2a Normalmente, por semana, quantos dias faz actividade física **moderada** como levantar e/ou transportar objectos leves, andar de bicicleta a uma velocidade moderada ou jogar ténis? Não inclua o andar/caminhar.

_____ dias por semana
_____ Nenhum (passe para a questão **3a**)

2b Quanto tempo costuma fazer actividade física moderada por dia?

_____ horas ___ minutos

3a Habitualmente, por semana, quantos dias **caminha** durante pelo menos 10 minutos seguidos? Inclua caminhadas para o trabalho e para casa, para se deslocar de um lado para outro e qualquer outra caminhada que possa fazer somente para recreação, desporto ou lazer.

_____ dias por semana
_____ Nenhum (passe para a questão **4a**)

3b Quanto tempo costuma caminhar por dia?

_____ horas ___ minutos

3c A que passo costuma caminhar?

_____ Passo **vigoroso**, que torna a sua respiração muito mais intensa que o normal;
_____ Passo **moderado**, que torna a sua respiração um pouco mais intensa que o normal;
_____ Passo **lento**, que não causa qualquer alteração na sua respiração;

As últimas questões referem-se ao tempo que está sentado diariamente no trabalho, em casa, no percurso para o trabalho e durante os tempos livres. Estas questões incluem o tempo em que está sentado numa secretária, a visitar amigos, a ler ou sentado/deitado a ver televisão.

4a Quanto tempo costuma estar sentado num **dia de semana**?

_____ horas ___ minutos

4b Quanto tempo costuma estar sentado num **dia de fim-de-semana**?

_____ horas ___ minutos

Campaniço, Helena (2016). Validade simultânea do questionário internacional de actividade física através da medição objectiva da actividade física por actigrafia proporcional. [Tese de Mestrado, Faculdade de Morticidade Humana, Universidade de Lisboa]. <https://www.repository.utl.pt/handle/10400.5/11866>