



UNIVERSIDADE NOVA DE LISBOA  
ESCOLA NACIONAL DE SAÚDE PÚBLICA



Faculdade de Ciências Médicas



Ana Cristina Vieira **Contributo para a validação da  
versão portuguesa da *Quebec  
Back Pain Disability Scale***

Fiabilidade e Poder de Resposta em utentes  
com Dor Lombar Crónica

Dissertação de Mestrado em Fisioterapia em  
Condições Músculo-Esqueléticas

Relatório de Projeto de Investigação

julho 2012

Relatório do Projeto 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, realizado sob a orientação científica do Professor Doutor Eduardo Cruz e coorientação da Professora Mestre Rita Fernandes.

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.

A candidata,

---

(Ana Cristina Pires Vieira)

Setúbal, 31 de julho de 2012

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

O orientador,

---

(Professor Doutor Eduardo Cruz)

Setúbal, 31 de julho de 2012

*“Procurai deixar o mundo um pouco melhor do que o que encontraste...”*

Baden – Powell

## **AGRADECIMENTOS**

Ao Professor Eduardo Cruz, o meu sincero agradecimento por tudo o que me ensinou, pelo exemplo, rigor científico, orientações objetivas e disponibilidade ao longo da realização deste trabalho;

À Professora Rita Fernandes, pela partilha dos seus conhecimentos e a sua cooperação fundamental na plataforma de registo dos dados;

Ao Conselho de Administração e Comissão de Ética, do Serviço Regional de Saúde da Região Autónoma da Madeira e à Santa Casa da Misericórdia de Machico por autorizarem a recolha de dados;

A todos os colegas que colaboraram no recrutamento dos participantes e na recolha dos dados, pela disponibilidade e espírito de entreatajuda;

Aos meus Pais, ao Nuno e à Joana, pelo apoio valioso ao longo desta etapa, pela compreensão e pelas “horas extra” que pude dedicar a esta dissertação;

Ao João Coelho, por me fazer feliz todos os dias e me dar confiança na concretização de mais um sonho.

À Eva Sousa por ter sempre um cantinho especial no “lote 3” e à Patrícia Ribeiro pela ajuda excecional nas viagens para Setúbal e partilha de emoções ao longo desta caminhada.

E a todos os outros amigos, que direta ou indiretamente contribuíram para a realização deste trabalho, o meu muito obrigada.

## RESUMO

### **Contributo para a validação da versão Portuguesa da *Quebec Back Pain Disability Scale*: Fiabilidade e Poder de Resposta em utentes com Dor Lombar Crónica.**

Ana Cristina Vieira

**PALAVRAS-CHAVE:** Dor Lombar Crónica, Incapacidade funcional, Fiabilidade, Poder de resposta.

**Introdução/Objetivo:** Um dos principais resultados das intervenções de Fisioterapia dirigidas a utentes com Dor Lombar Crónica (DLC) é reduzir a incapacidade funcional. A *Quebec Back Pain Disability Scale* (QBPDS) é um instrumento amplamente aceite a nível internacional na medição do nível de incapacidade funcional reportada pelos indivíduos com DLC. O objetivo deste estudo é dar um contributo para a adaptação cultural da versão portuguesa da QBPDS (QBPDS-VP) e investigar o poder de resposta e interpretabilidade da escala.

**Metodologia:** Realizou-se um estudo metodológico, multicentro, baseado num *coorte* prospetivo de 132 utentes com DLC. Os utentes foram recrutados a partir da lista de espera de 16 serviços de Medicina Física e de Reabilitação/Fisioterapia de várias áreas geográficas de Portugal. A QBPDS- VP foi administrada 3 vezes, em 3 momentos de recolha de dados distintos: T0 - momento inicial (utentes em lista de espera); T1 - 1 semana de intervalo (início dos tratamentos de Fisioterapia); e T2 - 6 semanas de intervalo (pós-intervenção de Fisioterapia). Os dados recolhidos em T0 foram utilizados para a análise fatorial e para o estudo da consistência interna (n=132); os dados da amostra emparelhada de T0 e T1 (n=132) para a fiabilidade teste-reteste; e os dados da amostra emparelhada de T0 e T2 (n=120) para a análise do poder de resposta e interpretabilidade. A âncora externa utilizada foi a perceção global de mudança, neste caso a PGIC- VP, que foi respondida em T1 e T2. O nível de significância para o qual os valores se consideraram satisfatórios foi de  $p \leq 0,05$ . O tratamento dos dados foi realizado no software *IBM SPSS Statistics* (versão 20).

**Resultados:** A QBPDS- VP é uma escala unidimensional, que revela uma excelente consistência interna ( $\alpha$  de Cronbach= 0,95) e uma fiabilidade teste-reteste satisfatória (CCI= 0,696; IC 95%: 0,581–0,783). Esta escala demonstrou um poder de resposta moderado, quando aplicada em utentes com DLC ( $\rho$ = 0,426 e AAC= 0,741; IC 95%: 0,645 – 0,837). A Diferença Mínima Detetável (DMD) estimada foi de 19 pontos e as estimativas da Diferença Mínima Clinicamente Importante (DMCI) variaram entre 7 (pelo método curva ROC) e 8 pontos (pelo método “diferença média de pontuação”). A estimativa pela curva ROC deriva do ponto ótimo de corte de 6,5 pontos, com Área Abaixo da Curva (AAC)= 0,741, sensibilidade de 72%, e especificidade de 71%. Uma análise complementar da curva ROC baseada nas diferenças de pontuações da QBPDS, expressa em percentagem, revelou um ponto ótimo de corte de - 24% (AAC= 0,737, sensibilidade de 71%, e especificidade de 71%). Para pontuações iniciais da QBPDS- VP mais altas ( $\geq 34$  pontos), foi encontrado um ponto ótimo de corte de 10,5 pontos (AAC= 0,738, sensibilidade de 73%, e especificidade de 67%).

**Conclusão:** A QBPDS-VP demonstrou bons níveis de fiabilidade e poder de resposta, recomendando-se o seu uso na medição e avaliação da incapacidade funcional de utentes com DLC. A DMD estimada, de 19 pontos, determinou uma amplitude válida da QBPDS-VP de 19 a 81 pontos. Este estudo propõe estimativas de DMCI da QBPDS- VP numa aplicação específica da escala (em utentes com DLC que são referidos para a intervenção de Fisioterapia). A pontuação inicial da QBPDS- VP deve ser considerada na interpretação de mudanças de pontuação, após a intervenção de Fisioterapia.

## ABSTRACT

### Contribute to the validation of the Portuguese version of the Quebec Back Pain Disability Scale: Reliability and Responsiveness for chronic low back pain patients

Ana Cristina Vieira

**Keywords:** Chronic low back pain, Functional disability, Reliability, Responsiveness.

**Background and Purpose:** One of the main results of physiotherapy interventions for patients with Chronic Low Back Pain (CLBP) is decrease the functional disability. The Quebec Back Pain Disability Scale (QBPDS) is an instrument widely accepted internationally, in measuring the level of disability reported by individuals with CLBP. The purpose of this study is to contribute to the cultural adaptation of the Portuguese version of QBPDS (QBPDS - PV) and investigate the Responsiveness and Interpretability of QBPDS-PV.

**Methodology:** This was a methodological and multicenter study, based on a sample of 132 subjects with CLBP. The patients were recruited from the waiting lists of 16 medicine rehabilitation service, in many Portugal districts. The Quebec Back Pain Disability Scale was administered in three different moments: T0 – baseline (patients in the waiting list); T1- one week after T0 (the beginning of treatment); and T2 – six weeks after T1 (the posttreatment). The data collected at T0 were used for factor analysis and to study the internal consistency (n = 132); paired sample data of T0 and T1 (n=132) were used for test-retest reliability, and sample data paired for T0 and T2 (n=120) used for responsiveness and interpretability analysis. The external anchor was the global perception of change, measured by the Portuguese version of *Patient's Global Impression of Change (PGIC) Scale*. The minimal level of significance established was  $p \leq 0,05$ . Data analysis was performed using the *IBM SPSS Statistics* software (version 20).

**Results:** The QBPDS-PV is a unidimensional scale, demonstrates an excellent internal consistency (Cronbach's  $\alpha=0.95$ ) and satisfactory test-retest reliability (ICC= 0.696, 95% CI: 0.581–0.783). The scale revealed moderate responsiveness when applied to patients with CLBP ( $\rho = 0.426$  and AUC= 0.741, 95% CI: 0.645 - 0.837). The Smallest Detectable Change (SDC) was 19 points, whereas the Minimal Clinically Important Change (MCIC) ranged between 7 (ROC curve method) and 8 points (by the "mean difference score"). The estimate was derived from the ROC curve by an optimal cutoff point of 6.5 points, with Area Under the Curve (AUC)= 0.741, sensitivity 72%, and specificity of 71%. A complementary analysis of the ROC curve based on differences in QBPDS scores from baseline, expressed in percentage, revealed an optimal cutoff point of -24% (AUC= 0.737, sensitivity of 71%, and specificity of 71%). For the highest initial scores of QBPDS-PV ( $\geq 34$  points) was found an optimal cutoff of 10.5 points (AUC= 0.738, sensitivity of 73%, and specificity 67%).

**Conclusion:** The QBPDS-PV demonstrated good levels of reliability and responsiveness, being recommended its use in the measurement and evaluation of disability of patients with CLBP. The SDC of 19 points determined the QBPDS' scale width of 19 to 81. This study proposes MCIC values for QBPDS –PV for this specific setting (in CLBP patients who are referred for physiotherapy intervention). The QBPDS –PV baseline score have to be taken into account while interpreting the score change after physiotherapy intervention.

## LISTA DE ACRÓNIMOS

AAC	- Área Abaixo da Curva
CCI	- Coeficiente de Correlação Intraclasse
COSMIN	- <i>Consensus-based Standards for the selection of health Measurement Instruments</i>
DL	- Dor Lombar
DLA	- Dor Lombar Aguda
DLC	- Dor Lombar Crónica
DMCI	- Diferença Mínima Clinicamente Importante
DMD	- Diferença Mínima Detetável
DP	- Desvio Padrão
EPM	- Erro Padrão de Medida
EPM <sub>repetição</sub>	- Erro Padrão de Medida associado à repetição das medições
EVA	- Escala Visual Análoga
GRI	- <i>Guyatt's Responsiveness Index</i>
IC	- Intervalo de Confiança
ICIDH	- <i>International Classification Impairment Disability Handicap</i>
ICF	- <i>Classification of Functioning, Disability and Health</i>
ODI	- <i>Oswestry Disability Index</i>
PGIC	- <i>Patient's Global Impression of Change</i>
QBPDS	- <i>Quebec Back Pain Disability Scale</i>
RMDQ	- <i>Roland Morris Disability Questionnaire</i>
RMP	- Resposta Média Padronizada
ROC	- <i>Receiver Operating Characteristic</i>
TE	- Tamanho do Efeito
VP	- Versão Portuguesa
WHO	- <i>World Health Organization</i>

# ÍNDICE

1. INTRODUÇÃO .....	1
2. REVISÃO DA LITERATURA .....	5
2.1. Dor lombar .....	5
2.2. Dor lombar Crónica .....	6
2.3. Epidemiologia da DLC .....	8
2.4. Instrumentos de avaliação da incapacidade .....	10
2.4.1. Quebec Back Pain Disability Scale .....	12
2.5. Propriedades psicométricas dos instrumentos .....	13
2.6. Propriedades psicométricas da Quebec Back Pain Disability Scale .....	15
2.6.1. Fiabilidade .....	15
2.6.1.1. Consistência interna .....	17
2.6.1.2. Fiabilidade (teste-reteste).....	19
2.6.1.3. Erro de Medição .....	21
2.6.2. Validade.....	23
2.7. Poder de resposta .....	23
2.7.1. Interpretabilidade (das mudanças).....	24
2.7.1.1. DMD .....	25
2.7.1.2. DMCI.....	27
2.7.1.3. Relação DMD/ DMCI.....	28
2.7.2. Métodos associados à estimativa do poder de resposta .....	29
2.8. Poder de resposta da QBPDS.....	32
2.8.1. Poder de resposta da QBPDS - Métodos baseados na distribuição .....	33
2.8.2. Poder de resposta da QBPDS - Métodos baseados na âncora .....	34
2.9. Interpretabilidade da QBPDS .....	36
1.9.1. Estimativa da DMD da QBPDS .....	36
2.9.2. Estimativa da DMCI da QBPDS .....	39
3. METODOLOGIA.....	44
3.1. Identificação dos Objetivos do estudo .....	44
3.2. Quebec Back Pain Disability Scale - VP .....	45

3.3. Avaliação das Propriedades Psicométricas .....	46
3.4. Tipo de estudo .....	47
3.5. Constituição da Amostra .....	48
3.6. Processo de recrutamento da amostra .....	49
3.7. Caracterização da amostra .....	50
3.8. Aspectos éticos .....	50
3.6. Instrumentos de medida .....	52
3.6.1. <i>Patient's Global Impression of Change (PGIC) Scale</i> - VP .....	52
3.7. Procedimentos de recolha de dados .....	53
3.7. Análise dos dados .....	55
3.7.1. Caracterização da amostra .....	55
3.7.2. Avaliação da Unidimensionalidade da QBPDS .....	56
3.7.2. Avaliação da consistência Interna .....	58
3.7.3. Avaliação da Fiabilidade Teste-reteste.....	58
3.7.3.1. Fiabilidade Teste-Reste pré intervenção da Fisioterapia (Intervalo de 4 a 7 dias); .....	59
3.7.4. Estimativa do EPM e da DMD.....	60
3.7.5. Determinação dos efeitos de “chão” e de “teto” .....	62
3.7.6. Estimar Poder de Resposta .....	62
3.7.6.1. Método “Coeficiente correlação âncora”.....	62
3.7.6.2. Método Curva ROC: área abaixo da curva .....	63
3.7.7. Determinação da Diferença Mínima Clinicamente Importante.....	63
3.7.7.1. Diferença Média de Pontuação .....	63
3.7.7.2. Método Curva ROC: ponto ótimo de corte.....	64
4. APRESENTAÇÃO DOS RESULTADOS .....	66
4.1. Caracterização da amostra T0 .....	67
4.2. Unidimensionalidade da QBPDS .....	69
3.2.1. Estrutura Fatorial da QPBDQ.....	69
4.3. Consistência Interna da QBPDS .....	72
4.4. Fiabilidade teste–reteste .....	75
4.5. Estimativa do EPM e DMD .....	76

4.5.2. Comparação dos resultados de incapacidade funcional em T0 e T2, para ambos os grupos .....	78
4.5.3. Equivalência dos grupos “sem alterações” e “melhoria clínica”, no ponto de partida .....	80
4.5.3.1. Equivalência da pontuação da QBPDS em T0 e da Intensidade da dor em T0 .....	81
4.5.3.2. Equivalência das variáveis sócio-demográficas e clínicas em T0 .....	82
4.5.4. Fiabilidade teste-reteste, pré-pós intervenção da Fisioterapia (intervalo de 7 semanas) .....	85
4.5.5. Cálculo do EPM e DMD .....	86
4.6. Determinação dos efeitos de “chão” e de “teto” .....	87
4.7. Poder de resposta da QBPDS .....	88
4.7.1. Método “Coeficiente correlação âncora” .....	88
4.7.2. Método Curva ROC: Área abaixo da curva .....	90
4.8. Determinação da Diferença Mínima Clinicamente Importante .....	91
4.8.1. Diferença média das pontuações nos indivíduos com PGIC=5 .....	91
4.8.2. Método Curva ROC: ponto ótimo de corte .....	92
4.8.3. DMCI - comparação de estimativas .....	96
4.9. Proporção de DMD e DMCI .....	97
5. DISCUSSÃO DOS RESULTADOS .....	99
5.1. Caracterização da amostra T0-T1 .....	99
5.2. Unidimensionalidade .....	101
5.3. Consistência Interna .....	102
5.4. Fiabilidade Teste Reteste .....	103
5.5. Fiabilidade Teste-Retest, EPM e DMD .....	105
5.6. Determinação dos efeitos de “chão” e de “teto” .....	112
5.7. Estimativa do Poder de resposta .....	113
5.8. Estimativa DMCI .....	115
5.9. Relação DMD e DMCI .....	120
6. CONCLUSÃO .....	122
6.1. Limitações do estudo .....	123
6.2. Implicações do estudo .....	125
6.3. Recomendações para estudos futuros .....	126

REFERÊNCIAS BIBLIOGRÁFICAS .....	128
APÊNDICE A: Estimativas de validade da QBPDS em diversos contextos culturais .....	i
APÊNDICE B: Protocolo para recrutamento dos participantes no estudo .....	ii
APÊNDICE C: Caderno de instrumentos .....	iii
APÊNDICE D: Pedido de autorização à administração do SESARAM, E.P.E. ....	iv
APÊNDICE E: Carta explicativa do estudo e consentimento informado .....	v
APÊNDICE F: Pedido de colaboração aos fisioterapeutas .....	vi
APÊNDICE G: <i>Output IBM SPSS Statistics</i> – Caracterização sócio-demográfica e clínica da amostra .....	vii
APÊNDICE H: <i>Output IBM SPSS Statistics</i> – Análise Fatorial.....	viii
APÊNDICE I: <i>Output IBM SPSS Statistics</i> – Consistência interna .....	ix
APÊNDICE J: <i>Output IBM SPSS Statistics</i> – Fiabilidade teste- resteste .....	x
APÊNDICE K: <i>Output IBM SPSS Statistics</i> – Verificação de condições prévias à estimativa do EPM e DMD .....	xi
APÊNDICE L: Recodificação das variáveis sócio-demográficas e clínicas .....	xii
APÊNDICE M: <i>Output IBM SPSS Statistics</i> – Verificação dos efeitos de “chão” e de “teto” .....	xiii
APÊNDICE N: <i>Output IBM SPSS Statistics</i> – Poder de Resposta e DMCI.....	xiv
ANEXO I: Autorização da comissão de ética e administração do SESARAM, E.P.E ...	xv
ANEXO II: Autorização da Santa Casa da Misericórdia de Machico.....	xvi

# 1. INTRODUÇÃO

A Dor Lombar (DL) continua a ser uma condição com elevada incidência e prevalência (Koes *et al.*, 2010). Apesar dos contínuos esforços de investigação e do desenvolvimento de novas intervenções, a DL mantém-se um desafio clínico (Mayer, Haldeman, Tricco & Dagenais, 2010) e uma condição que acarreta, às sociedades industrializadas, custos consideráveis em cuidados de saúde, bem como um impacto socioeconómico negativo (Apeldoorn *et al.*, 2010; van Middelkoop *et al.*, 2011).

Na maioria dos casos de DL (85% a 90%), a dor não pode ser atribuída a uma patologia específica conhecida (Burton *et al.*, 2004; Airaksinen *et al.*, 2006), a causa da dor não pode ser definitivamente identificada e não é possível atribuir um diagnóstico pato-anatómico preciso (Slade & Keating, 2009). Nestes casos, a DL recebe a designação de DL de origem não específica.

Este relatório de investigação incide particularmente na população com Dor Lombar Crónica (DLC), referindo-se à população que apresenta DL de origem não específica, que se manifesta de forma persistente, ou recorrente, por mais de 12 semanas, ou com múltiplos episódios de recorrência ao longo da sua condição crónica (Airaksinen *et al.*, 2006; Bekkering *et al.*, 2003; Krismer & van Tulder, 2007; Stanton, Latimer, Maher & Hancock, 2010).

A prevalência da DLC em Portugal não é conhecida porque até ao momento não foram desenvolvidos estudos epidemiológicos de investigação da DLC, contudo, um estudo recente, reconhece o impacto da dor crónica em Portugal e o “peso” que a DL representa nessas condições. Estima-se que a DL atinge mais de 40% dos indivíduos que reportam dor crónica (Castro-Lopes, Saramago, Romão & Paiva, 2010).

Os utentes com DLC recorrem frequentemente aos cuidados de fisioterapia. Porém, em Portugal, a maioria dos utentes com dor crónica, recorre em primeiro lugar ao médico de família ou clínico geral que decide acerca da necessidade de referenciar o utente para os cuidados secundários (Castro-Lopes *et al.*, 2010). É imperativo que a Fisioterapia demonstre a efetividade dos serviços prestados aos utentes com DLC, para que eles sejam de facto orientados para a Fisioterapia. Uma das estratégias na concretização deste objetivo passa pela medição e avaliação dos resultados de saúde nos utentes com essa condição.

A literatura sugere que os fisioterapeutas enfrentam constrangimentos de tempo na sua prática clínica e que essa é a maior barreira à utilização de instrumentos de medida formais (Abrams *et al.*, 2006; Delaune & Bemis-Dougherty, 2007). Esta lacuna nos hábitos de medição e registo dos resultados dificulta, por conseguinte, o conhecimento dos resultados da Fisioterapia dirigida à população com DLC.

Um resultado clínico fundamental e consensual da intervenção da Fisioterapia dirigida a utentes com DLC, é a redução da incapacidade funcional (Davidson & Keating, 2002; Ostelo *et al.*, 2008; Smeets *et al.*, 2011). Para avaliar a incapacidade funcional, quer no âmbito da prática clínica, quer para fins de investigação, têm sido utilizados, sobretudo, instrumentos de autorrelato (Smeets *et al.*, 2011), que vão de encontro às necessidades dos fisioterapeutas, uma vez que são de autopreenchimento e não consomem tempo aos profissionais.

A *Quebec Back Pain Disability Scale* (QBPDS) é uma das escalas mais recomendadas para avaliar a incapacidade funcional em utentes com DL (Rocchi *et al.*, 2005; Davidson & Keating, 2002). Está traduzida em muitas línguas e validada em vários países, o que atesta o seu interesse e aceitação como instrumento de avaliação do grau de incapacidade dos utentes com DL.

A preferência da QBPDS em relação a outras escalas deve ser fundamentada nas propriedades psicométricas das suas medições. A literatura recente sugere que a propriedade de poder de resposta, é tão ou mais importante do que a validade e fiabilidade das medições na escolha do instrumento de avaliação (Lauridsen *et al.*, 2006). O poder de resposta é a capacidade de um instrumento de medida detetar mudanças, e assume particular relevância quando um instrumento, tal como a QBPDS, é usado para avaliar a mudança da condição clínica dos utentes.

Recentemente tem sido dado ênfase à necessidade de estudar a interpretabilidade dos resultados dos instrumentos de medida, no sentido de reconhecer quando é que a alteração da pontuação representa uma mudança “real” e/ou quando é que representa uma mudança clinicamente importante. A partir desse conhecimento, é possível aplicar os resultados dos instrumentos no processo de tomada de decisão relativa à progressão clínica dos utentes e na avaliação dos efeitos da intervenção da Fisioterapia em utentes com DLC.

Desde 2005 encontra-se disponível a versão portuguesa (VP) da QBPDS, que se encontra em fase de adaptação cultural e carece de estudos que avaliem a sua precisão na medição da incapacidade funcional e na deteção de mudanças ao longo do tempo. Até à data a QBPDS- VP é aplicada sem conhecimento prévio do valor da Diferença Mínima Detetável (DMD) e Diferença Mínima Clinicamente Importante (DMCI).

Neste contexto parece ser particularmente relevante e necessário prosseguir a investigação dedicada à adaptação cultural da QBPDS-VP e desenvolver estudos de elevada qualidade metodológica, e com amostras suficientemente grandes, que permitam avaliar as propriedades psicométricas da QBPDS – VP.

Os principais objetivos deste trabalho de investigação são: contribuir para a adaptação cultural da QBPDS-VP e investigar a capacidade da QBPDS- VP detetar mudanças na incapacidade funcional dos utentes com DLC que frequentam tratamentos de Fisioterapia. Especificamente pretende-se:

- a. Averiguar a unidimensionalidade da QBPDS – VP e estimar a consistência interna e fiabilidade teste–reteste das suas medições;
- b. Avaliar o poder de resposta da QBPDS- VP e estimar medidas de interpretabilidade de mudanças de pontuação da QBPDS – VP: a DMD e a DMCI;

No final deste trabalho, espera-se que os resultados obtidos permitam julgar a QBPDS- VP quanto à fiabilidade das suas medições e ao poder de resposta da escala na avaliação e monitorização da incapacidade funcional dos utentes com DLC.

Ambiciona-se que os resultados deste trabalho deem a conhecer as estimativas de duas medidas essenciais de interpretabilidade e auxiliem os profissionais de saúde a interpretar a relevância clínica de mudanças na pontuação da QBPDS – VP, quando aplicada a utentes com DLC. Dessa forma espera-se que os profissionais de saúde possam ter duas referências de suporte que permitam dar resposta a questões frequentes como: A mudança observada na pontuação total da QBPDS-VP, após tratamento, corresponde a uma mudança real no nível de incapacidade funcional? O valor dessa mudança na pontuação total da QBPDS-VP é clinicamente importante e relevante para o utente com DLC?

Para os propósitos definidos, estabelece-se um estudo metodológico, que tem por base um *coorte* prospetivo de utentes com DLC, que é acompanhado por um período de 7 semanas (a primeira semana contempla o período entre a lista de espera e o

início dos tratamentos de Fisioterapia, e as 6 semanas seguintes correspondem ao período de intervenção de Fisioterapia). Todos os procedimentos realizados e dados obtidos com a realização desse estudo são apresentados de forma sistematizada e discutidos ao longo deste relatório de investigação.

A presente dissertação está organizada em 6 capítulos. O Capítulo 2 destina-se à Revisão da Literatura e pretende esclarecer os conceitos teóricos relevantes para a compreensão do trabalho de investigação realizado. Neste sentido, este capítulo engloba a caracterização da DLC e dos seus dados epidemiológicos; a temática da medição da incapacidade funcional em utentes com DLC e o aprofundamento das propriedades psicométricas. A grande maioria da revisão é dedicada à sistematização dos resultados de investigação relativos às propriedades psicométricas da QBPDS, nas suas variadas versões linguísticas.

O capítulo 3 apresenta a descrição e argumentação da metodologia do estudo de investigação, incluindo objetivos, tipo/desenho de estudo, a amostra e respetivo processo de recrutamento, os aspetos éticos, os instrumentos de avaliação, os procedimentos e a estrutura para a análise dos dados.

Os capítulos 4 e 5 são dedicados à apresentação e discussão dos resultados, que se encontram organizados pela mesma sequência dos objetivos propostos. Os resultados são discutidos em função da evidência anteriormente publicada sobre a QBPDS e do conhecimento mais atual.

No capítulo 6, enuncia-se as principais conclusões, focando os resultados obtidos, o contributo dos mesmos para o conhecimento na área de investigação e as implicações clínicas do trabalho. Por último, apontam-se algumas direções em que pode ser desenvolvido o trabalho de investigação com interesse para a solução do problema da efetividade das intervenções de Fisioterapia dirigidas a utentes com DLC.

## 2. REVISÃO DA LITERATURA

### 2.1. Dor lombar

A dor lombar (DL) é definida como dor e desconforto, localizada abaixo da margem costal e acima das pregas glúteas inferiores, com ou sem dor na perna (Burton *et al.*, 2004; Krismer & van Tulder, 2007). A DL é uma condição complexa que se manifesta sob uma variedade de situações clínicas que devem ser corretamente distinguidas.

Para a classificação da DL tem sido internacionalmente aceite a categorização genérica, designada por triagem de diagnóstico (Airaksinen *et al.*, 2006), que divide a DL em 3 categorias, de acordo com as características clínicas apresentadas pelos utentes. Trata-se de uma classificação simples, prática e crescentemente utilizada no âmbito da investigação. Segundo Waddell (1987), a DL pode ser classificada da seguinte forma:

- Dor com origem em patologia vertebral específica;
- Dor com origem radicular ou no nervo;
- Dor com origem não específica;

Na DL com origem em patologia vertebral específica, os sintomas são atribuídos à suspeita, ou confirmação, de patologias da coluna vertebral, como infeção, tumor, osteoporose, fratura, deformidade estrutural, doença inflamatória (ex. espondilite anquilosante), entre outros. Na DL com origem radicular ou no nervo, os sintomas são associados a uma dessas duas estruturas e a DL manifesta-se por síndrome radicular ou síndrome da cauda equina ou sinais de compressão nervosa (por exemplo, do nervo ciático) (Koes *et al.*, 2010). Por último, a DL de origem não específica, que será objeto de estudo neste trabalho, diz respeito à DL cuja origem não é atribuída a uma patologia vertebral específica e/ou causa conhecida (Burton *et al.*, 2004).

Apenas, em aproximadamente 10% dos utentes com DL, podem ser identificados mecanismos pato-fisiológicos específicos subjacentes (Krismer & van Tulder, 2007). Assim, estima-se que a DL de origem não específica está presente entre 85% (Burton *et al.*, 2004) e 90% (Krismer & van Tulder, 2007) dos utentes que

apresentam DL.

A DL de origem não específica pode ser referida para o membro inferior, desde que não seja acompanhada de sinais de défice neurológico (alterações nos reflexos, sensibilidade ou força); e pode ainda compreender alterações de espondilose e de espondilolistesis. Estas alterações (espondilose e/ou espondilolistesis) associadas a sintomas de DL, são classificadas como DL de origem não específica, dado que uma proporção considerável de utentes que as apresenta, revela-se assintomática (Soler & Calderón, 2000).

## **2.2. Dor lombar Crónica**

Para além da classificação anterior, a DL é geralmente classificada de acordo com a duração dos sintomas como: aguda (menos de 6 semanas), subaguda (entre 6 semanas e 12 semanas) ou crónica (mais de 12 semanas) (Bekkering *et al.*, 2003; Krismer & van Tulder 2007). Por definição, a dor crónica “é a dor que persiste para lá do tempo normal de reparação dos tecidos. Se a dor não for decorrente de patologia maligna, considera-se que o período 3 meses é o mais conveniente para a divisão entre dor aguda e dor crónica” (Merskey & Bogduk, 1994).

Em geral, a DL tem uma história natural e um prognóstico favorável. A história natural da DL de origem não específica prevê que a DL é autolimitada a um intervalo de poucas semanas (Melloh *et al.*, 2009). Os estudos revelam que, 30 a 60% dos indivíduos com DL aguda recupera no espaço de uma semana, e que 60 a 90% dos indivíduos recupera num período de 6 semanas (Du Bois & Donceel, 2008; Manchikanti, Singh, Datta, Cohen & Hirsch, 2009).

Segundo Koes *et al.* (2010), após um novo episódio de DL, é expectável que a dor melhor substancialmente durante as primeiras 4 a 6 semanas, mas não se resolva completamente. Existe uma pequena minoria de indivíduos com DL, 3 a 10% (Manchikanti *et al.*, 2009; Melloh *et al.*, 2009) que mantêm a dor, de forma persistente, por mais de 12 semanas e desenvolvem uma condição crónica (DLC), com um prognóstico menos favorável.

O conceito de DLC supramencionado, tem sido largamente questionado. Os estudos mostram que a classificação da DLC, baseada numa duração para além de 12 semanas, não é suficiente. Existem manifestações de DL, com múltiplos episódios ou recorrências, que apesar de não persistirem por 12 semanas, também podem ser consideradas condições crónicas. A forma recorrente da DLC pode assumir uma grande variedade de definições, no que se refere à frequência dos episódios e/ou duração da condição, que pode dificultar a interpretação da investigação, bem como, a prática clínica (Stanton *et al.*, 2010).

Dada a falta de consenso relativa à classificação temporal da DLC, atualmente, a condição caracteriza-se sobretudo pela sua origem não específica, associada a uma das várias manifestações (número de episódios, frequência, etc.) reportadas na literatura.

A população com DLC e/ou recorrente é a população alvo deste estudo. Ao longo do trabalho, a DLC de origem não específica e/ou recorrente, é reportada, simplesmente por DLC. Para efeitos deste trabalho, caracteriza-se a DLC, por:

- a) presença de dor persistente na região lombar, com duração de pelo menos 12 semanas ou presença de recorrências de dor lombar por um período de seis meses, sem causa física específica, isto é, de origem desconhecida (Airaksinen *et al.*, 2006; Bekkering *et al.*, 2003; Krismer & van Tulder, 2007), **ou**
- b) “presença de dor em menos de metade dos dias num período de 12 meses, ocorrendo múltiplos episódios no decorrer desse período” (Von Korff, 1994). De acordo com Von Korff (1994), estes episódios de agudização são definidos como “períodos (normalmente uma semana ou menos) em que os sintomas da condição de dor lombar se encontram mais exacerbados do que o normal para o utente”, **ou**
- c) múltiplas situações de agudização dos sintomas, correspondendo assim a vários episódios de recorrência ao longo da sua condição crónica (Stanton *et al.*, 2010), **ou**
- d) utentes com flutuações e recorrências na dor, embora tenham sempre presente a condição. Esta presença constante é interpretada como uma condição de DLC que envolve as limitações funcionais “normais” (isto é, com presença, ou não, de dor, mas com a existência de incapacidade mínima na realização das suas atividades diárias) alternando com a existência de flutuações e episódios de

agudização dos sintomas, com elevados níveis de dor e incapacidade, comprometendo a sua participação na maioria das atividades diárias e atividade profissional (Young, Wasiak, Phillips & Gross, 2011).

### 2.3. Epidemiologia da DLC

Existe evidência consistente de que a prevalência e os custos associados à DLC estão a aumentar (Freburger *et al.*, 2009). A DLC é um problema que ganha cada vez mais expressão nas sociedades industrializadas, face à elevada utilização dos cuidados de saúde, aos custos crescentes desses cuidados e à eficácia limitada de muitos tratamentos atuais (Slade & Keating, 2009). É uma fonte significativa de incapacidade a longo-prazo e de absentismo laboral, que representa 70% a 90% de todos os custos associados à condição de DL, inclusive em cuidados de saúde e compensações de incapacidade (Manchikanti *et al.*, 2009).

Em Portugal, não existem estudos à escala nacional que tenham descrito a prevalência específica da DLC na população, nem o seu impacto a nível de indicadores de saúde e económicos. No ano de 2002, foi realizado um estudo pelo Observatório Nacional de Saúde (ONSA) do Instituto Nacional de Saúde Dr. Ricardo Jorge, sobre a dor na população portuguesa. Os dados foram recolhidos por entrevista telefónica, tendo sido questionada a presença de dor nos 7 dias anteriores à entrevista. Das 1414 respostas válidas, 73,7% dos inquiridos manifestaram ocorrência de dor, onde as dores lombares foram as mais referidas (51,3%) (Rabiais, Nogueira & Falcão, 2003).

Recentemente foi desenvolvido um levantamento sobre a dor crónica em Portugal (Castro-Lopes *et al.*, 2010), enquadrado numa iniciativa europeia, a *Pain Proposal*, que foi lançada para obter um consenso sobre o impacto da dor crónica em toda a Europa. Observou-se que em Portugal, 36% (Intervalo de Confiança<sup>1</sup> 95%: 34,4 - 37,4) da população adulta sofre de dor crónica, sendo que em quase metade dos casos (16%, para IC 95%: 14,3 - 16,6) a dor é moderada ou forte ( $\geq 5$  numa escala numérica 0-10); e que a prevalência da dor é significativamente maior no género feminino e aumenta com a idade em ambos os géneros. Das principais causas de dor crónica,

---

<sup>1</sup> Intervalo de Confiança (IC)

surgem em primeiro lugar, as patologias osteoarticulares, em particular a DL, que atinge mais de 40% dos indivíduos. Estes dados, permitem inferir que num grupo de 10 indivíduos portugueses, aproximadamente 4 apresentam dor crónica, dos quais 2 queixam-se de DLC.

Considerando a elevada prevalência da dor crónica, pode concluir-se que estamos perante um grave problema de saúde pública que urge combater com todos os meios ao nosso alcance. Segundo Castro - Lopes et. al, (2010), “o impacto da dor crónica na qualidade de vida dos doentes é devastador e superior ao provocado por muitas outras patologias, principalmente pelo sofrimento que acarreta e pelas limitações que impõe em múltiplas atividades do dia a dia das pessoas. Frequentemente, a dor do doente afeta também, de forma muito significativa, o bem-estar e as atividades dos familiares e outros cuidadores”.

Os indivíduos com DLC reportam frequentemente comprometimento na capacidade de realizar as suas atividades diárias (Lin *et al.*, 2011). Vestir-se, sentar-se, estar em pé, andar e limpar a casa são exemplos dessas atividades. Segundo Bicer *et. al.* (2005), o impacto da dor lombar está intimamente relacionado com a incapacidade funcional.

Vários estudos têm investigado a prevalência da dor e da incapacidade funcional associada à DLC. Cassidy, Carroll & Cote (1998), determinaram essas prevalências entre os adultos de uma cidade do Canada (Saskatchewan), através de um estudo populacional transversal, cuja amostra aleatória foi contactada via e-mail (n=2184). Obtiveram 55% de respostas. Concluíram que nos 6 meses prévios, 8,9%, (IC 95%: 45,9-52,0) da população tinha experienciado DL de baixa intensidade e baixa incapacidade, 12,3% (IC 95%:10,3-14,4) tiveram DL de elevada intensidade e baixa incapacidade; e 10,7% (IC 95%: 8,8-12,5) relataram que nos últimos 6 meses tinham experienciado DL associada a elevada incapacidade.

Foram obtidos resultados semelhantes no estudo de Walker, Muller & Grant (2004), desenvolvido com uma amostra aleatória estratificada da população adulta da Austrália (n= 3000). Obtiveram uma taxa de resposta de 69% e uma taxa de prevalência de DLC no período de 6 meses de 64.6% (IC 95%: 62.4 a 66.7); No período de 6 meses prévios, 42,6% (IC 95%: 40,4 a 44,8) da população adulta tinha experimentado dor e incapacidade de baixa intensidade com origem na DL. Outros 10,9% (IC 95%: 9,6 a 12,3) tiveram dor de elevada intensidade mas baixos níveis de incapacidade decorrente

da DL. No entanto, 10,5% (IC 95%: 9,2 a 11,9) tiveram elevada incapacidade relacionada com a DLC.

Os dois estudos anteriores, utilizaram a mesma escala de classificação, o Questionário da Dor Crónica (Von Korff *et al.*, 1992), e ambos concluíram que mais de 10% da amostra esteve nos 6 meses anteriores, significativamente incapacitada devido à DL.

#### **2.4. Instrumentos de avaliação da incapacidade**

O tratamento dos utentes com DLC e respetiva incapacidade funcional, visa, em primeiro lugar, melhorar os níveis de atividade e participação dos mesmos (Smeets, *et al.*, 2011). A restauração de uma funcionalidade normal é considerada um resultado chave da Fisioterapia dirigida a utentes com DL (Davidson & Keating, 2002).

A incapacidade física e funcional são preocupações na área da saúde, e a importância da avaliação do estado funcional é amplamente aceite (Kopeck *et al.*, 1995). A avaliação da incapacidade tem progredido e acompanhado as alterações da sua definição. Em 1980, o termo “incapacidade” em vigor, definido pela *World Health Organization* (WHO) na “*International Classification Impairment Disability Handicap*” (ICIDH) era: "qualquer restrição ou falta de capacidade para realizar uma atividade de uma maneira ou dentro do intervalo considerado normal para um ser humano" (WHO, 1980). Essa definição foi revista pela WHO, em 2007, e publicada através da “*Classification of Functioning, Disability and Health*” (ICF), como um termo abrangente que cobre 3 aspetos amplos da saúde: funções do corpo e estruturas; limitações de atividade e restrições de participação (WHO, 2007).

Ao nível da clínica, constatou-se que as medidas de deficiência, como mobilidade da coluna vertebral estão muitas vezes pouco relacionadas com a condição ou com o nível de funcionalidade de um utente (Schoppink, van Tulder, Koes, Beurskens & de Bie, 1996). Em consequência desse fato, Schoppink, *et al.* (1996) sugeriram que o estado funcional devia ser medido diretamente em vez de inferido a partir medições de deficiências, de força muscular ou de amplitude de movimento.

Por outro lado, tem existido um crescente reconhecimento da necessidade de considerar que a avaliação dos utentes do seu próprio estado funcional ou incapacidade, desempenha um papel significativo na avaliação dos resultados da intervenção. Para avaliar a capacidade funcional diária, quer na prática clínica, quer para fins de investigação, têm sido utilizados sobretudo questionários de autorrelato (Smeets, *et al.*, 2011).

Na última década têm surgido vários questionários específicos para a DL, que aparentemente avaliam o mesmo constructo e que têm sido adequadamente validados, tornando desafiante a escolha de um instrumento adequado para uma determinada situação (Lauridsen *et al.*, 2006). A dificuldade da escolha deve-se à variabilidade no conteúdo analisado, à diferença de perguntas e tipo de escalas utilizadas, bem como à variabilidade de tempo necessário para a administração dos instrumentos (Schoppink *et al.*, 1996). Segundo Davidson & Keating (2002), um questionário é considerado com potencial utilidade clínica se for autopreenchido, curto e fácil de completar, simples de calcular a pontuação final e que não tenha demonstrado ter sérios efeitos de “chão” e de “teto”, na população clínica geral.

Para avaliar a incapacidade funcional associada à DLC existem diversas escalas e questionários suportadas em diferentes conceitos de incapacidade funcional. Esses instrumentos, para além do enfoque na diminuição da capacidade de desempenho e na alteração do desempenho de atividades diárias, abrangem outras limitações da saúde subentendidas na definição da ICF (Lin *et al.*, 2011).

Dos vários instrumentos de medição estabelecidos na literatura, que podem ser aplicados para avaliar a incapacidade em utentes com DL, os três mais utilizados pelos clínicos, para seguir o curso da condição em utentes com DL e avaliar o impacto da intervenção terapêutica, são: *Roland Morris Disability Questionnaire* (RMDQ), *Oswestry Disability Index* (ODI) e QBPDS (Davidson & Keating, 2002; Muller, Roder, & Greenough, 2006). Esses três instrumentos foram totalmente validados do ponto de vista psicométrico (Rocchi *et al.*, 2005), sendo a QBPDS, um dos instrumentos mais recomendados para avaliar a capacidade funcional em utentes com DL (Rocchi *et al.*, 2005; Davidson & Keating, 2002).

### 2.4.1. Quebec Back Pain Disability Scale<sup>2</sup>

A QBPDS foi originalmente desenvolvida em língua inglesa, por Kopec e seus colaboradores (Kopec *et al.*, 1996). Comparativamente a outras escalas semelhantes, existem pelo menos três fatores que suportam a QBPDS como o instrumento de medida mais indicado para avaliar a incapacidade funcional de utentes com DL. Esses fatores são: um construto bem definido, a explicação da sua origem e a ausência de efeitos de “chão” e de “teto”.

No que diz respeito ao construto, a QBPDS que foi desenvolvida a partir do modelo conceptual de incapacidade “*International Classification Impairment Disability Handicap*” (ICIDH) com o objetivo de avaliar o nível de incapacidade funcional de indivíduos com DL, resulta de um conceito de incapacidade bem definido.

Ao nível da conceção, a QBPDS é o único instrumento que explica os métodos que estão na origem da seleção e geração dos itens da escala (Muller *et al.*, 2006). Inicialmente constituída por 48 itens, a QBPDS sofreu uma redução para 20 itens de atividades da vida diária. Kopec *et al.*, 1996, utilizaram uma análise fatorial e a avaliação dos utentes e peritos para selecionar os vinte itens incluídos na versão final da escala. Esses itens são representativos das atividades comuns da vida diária em que os utentes podem enfrentar dificuldade na sua execução, como um resultado direto da sua DL (Kopec *et al.*, 1996; Muller *et al.*, 2006).

Do ponto de vista das propriedades psicométricas, a QBPDS não possui efeitos de “chão”, quando comparada com a escala de ODI, e também não apresenta efeitos de “teto” em relação à RMDQ (Davidson & Keating, 2002; Muller *et al.*, 2006)<sup>3</sup>.

Kopec *et al.* (1995) avaliaram as propriedades psicométricas da versão final da QBPDS (20 itens), que desde então tem sido a única versão utilizada e traduzida para muitas línguas: Alemão (Schoppink *et al.*, 1996); Francês (Yvanes-Thomas *et al.*, 2002); Turco (Bicer *et al.*, 2005); Iraniano (Mousavi *et al.*, 2006); Português do Brasil (Rodrigues *et al.*, 2009) e Árabe (Beneddouché *et al.*, 2011).

---

<sup>2</sup> Também é referida na literatura como “*Quebec Back Pain Disability Questionnaire*”

<sup>3</sup> Se a escala tivesse um efeito teto, o que normalmente acontece quando os itens de teste não são suficientemente exigentes para um grupo de utentes gravemente afetados pela DL, a escala só seria aplicável aos utentes com casos leves a moderados (Muller *et al.*, 2006).

A adaptação cultural da QBPDS para a população portuguesa foi da autoria de Nunes, Ribeiro & Cruz (2005), que realizaram a tradução para a língua portuguesa; versão de consenso/síntese; retroversão; validação por comité de peritos; validação facial com uma amostra de utentes; e uma auditoria final. Até ao momento, ainda não foi desenvolvido nenhum estudo que teste a validade psicométrica da QBPDS – VP. Neste sentido, este estudo propõe-se a realizar a análise fatorial da QBPDS - VP, avaliar a fiabilidade e poder de resposta das suas medições, quando aplicada em utentes com DLC.

## **2.5. Propriedades psicométricas dos instrumentos**

Antes de um instrumento de medição do estado de saúde poder ser utilizado em investigação ou prática clínica, as suas propriedades psicométricas, isto é, fiabilidade, validade e poder de resposta, devem ser avaliadas e consideradas adequadas (Mokkink *et al.*, 2010a). Só assim, os resultados obtidos a partir de um instrumento de medição podem ser utilizados, com confiança, na tomada de decisões de intervenção e de progressão da intervenção. Nesse sentido, devem ser desenvolvidos estudos de avaliação das propriedades psicométricas de elevada qualidade metodológica, que garantam conclusões efetivas sobre essas propriedades do instrumento (Mokkink *et al.*, 2010a).

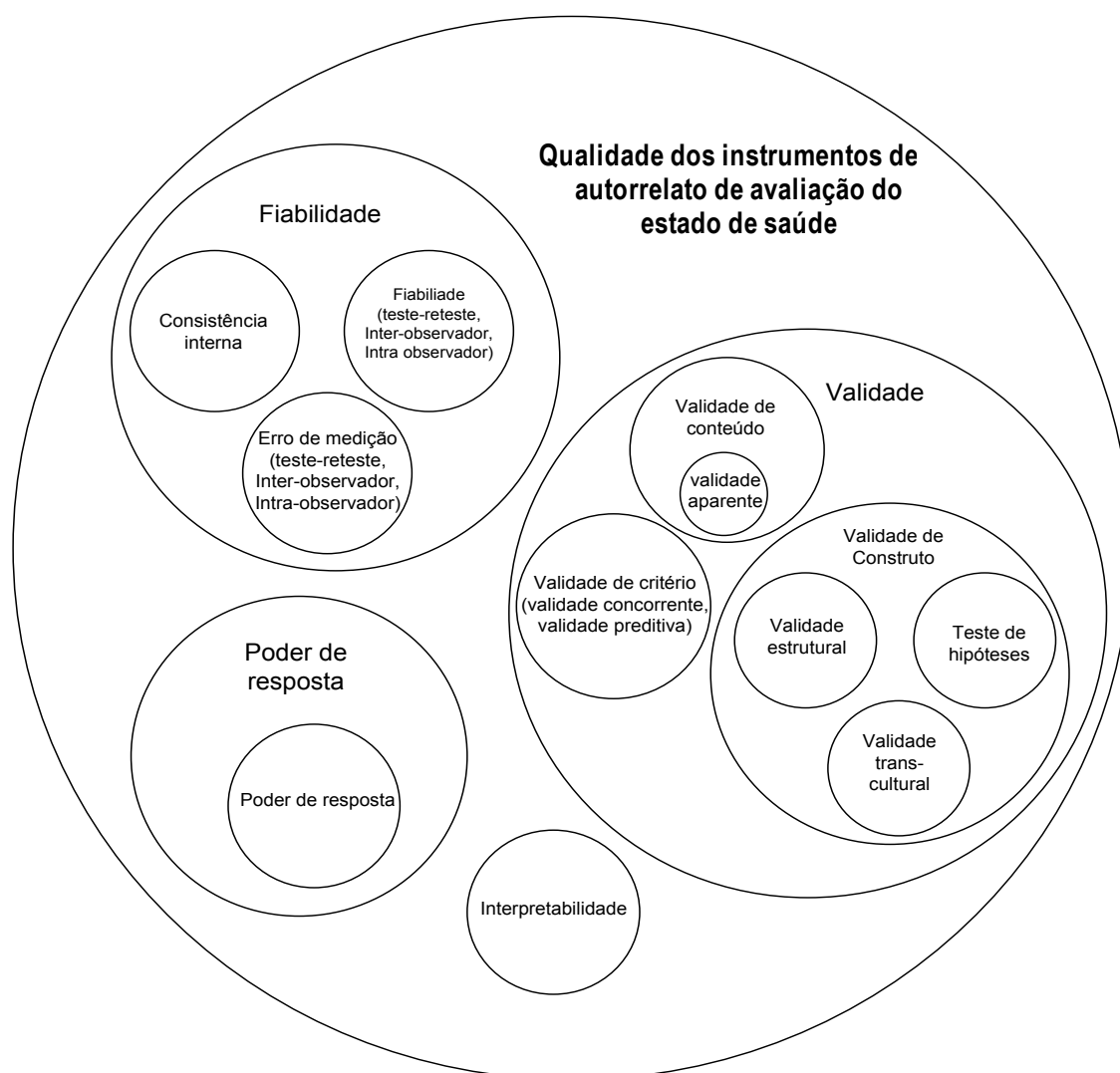
A investigação das propriedades psicométricas é particularmente importante para medidas de saúde que são diretamente relatadas por utentes (Mokkink *et al.*, 2010a). A condição funcional é um exemplo de construto, que não é diretamente mensurável e que pela sua natureza subjetiva, torna-se muito importante avaliar quando é que o instrumento de medida permite medir este construto de uma forma válida e fiável. Caso contrário, há um sério risco de obter resultados imprecisos e enviesados que podem conduzir a conclusões erradas (Mokkink *et al.*, 2010a).

Para facilitar a organização e exposição das propriedades psicométricas da QBPDS, ao longo deste trabalho será adotada a taxonomia, representada na figura 1, recentemente desenvolvida por um painel de peritos da iniciativa COSMIN (*Consensus-based Standards for the selection of health Measurement Instruments*). Esta taxonomia esclarece as relações entre as propriedades psicométricas que são relevantes para

”instrumentos de autorrelato de avaliação do estado de saúde”, aplicados com a intenção de avaliar, como é o caso específico da QBPDS.

Segundo Mokkink *et al.* (2010a), os requisitos das propriedades psicométricas variam com a aplicação do instrumento. Por exemplo, os instrumentos supramencionados, com uma aplicação de avaliação, isto é, com uma aplicação longitudinal que visa avaliar os efeitos do tratamento ou mudanças de saúde ao longo do tempo, necessitam de ser “sensíveis à mudança”, ao contrário dos instrumentos discriminativos.

A taxonomia compreende 3 domínios: a fiabilidade, a validade e o poder de resposta, em que cada qual contém propriedades psicométricas, como se pode visualizar na Figura 1. A interpretabilidade também foi incluída na taxonomia, apesar de não ser considerada uma propriedade psicométrica (Mokkink *et al.*, 2010a).



**Figura 1** – Taxonomia da COSMIN - relações entre as propriedades psicométricas. Tradução do diagrama de Mokkink *et al.* (2010).

## 2.6. Propriedades psicométricas da Quebec Back Pain Disability Scale

Foi realizada uma pesquisa e revisão da literatura publicada até janeiro de 2012 sobre as propriedades psicométricas de QBPDS. Constatou-se que têm sido desenvolvidos vários estudos sobre as propriedades psicométricas de fiabilidade e validade, que apresentam bons resultados (Tabelas 1 e 2). Contudo, o poder de resposta e interpretabilidade apenas foram investigados em poucos estudos (Tabelas 4, 5, 6 e 7). Este subcapítulo é alusivo às propriedades: fiabilidade e validade.

### 2.6.1. Fiabilidade

A fiabilidade refere-se à consistência da medição de uma variável latente (Hill & Hill, 2005). A existência de fiabilidade adequada é necessária, mas não suficiente, para garantir validade adequada (Hill & Hill, 2005). Ou seja, o fato da medida ser fiável e medir alguma variável de forma consistente, não garante que a mesma meça a variável pretendida. Assim, é fundamental que a medida de uma variável latente verifique coeficientes de fiabilidade e de validade adequados.

Segundo a teoria clássica da Fiabilidade, a variância total entre uma série de pontuações observadas é resultado de uma variância verdadeira entre as pontuações conjuntamente com a variância dos erros (Portney & Watkins, 2009).

A essência da fiabilidade é baseada na quantidade de erro que está presente numa série de pontuações. Uma medida é considerada mais fiável quanto maior for a proporção de variância observada, que é representada, por variância verdadeira nas pontuações. Assim, a fiabilidade é definida pela razão (adaptada de Portney & Watkins, 2009):

$$\frac{\text{Variância verdadeira}}{\text{Variância verdadeira} + \text{Variância do erro}} = \frac{\text{Variância verdadeira}}{\text{Variância Total}} = \frac{\text{Variância total} - \text{variância do erro}}{\text{Variância total}}$$

O coeficiente de fiabilidade pode tomar valores entre 0 e 1. Uma fiabilidade igual a 0, indica que toda a variação das medições é atribuída ao erro. Pelo contrário, uma fiabilidade de 1, significa que as medidas não têm componente de erro. Assim, à medida que o coeficiente se aproxima de 1, aumenta o grau de confiança de que as

pontuações observadas são representativas de pontuações verdadeiras (Portney & Watkins, 2009).

Para um determinado valor de variância do erro, e tendo em conta a equação da fiabilidade, quanto maior for a variância total, mais adequada será a fiabilidade. Ou seja, é necessária uma grande variação observada entre os indivíduos numa amostra, para que seja verificada fiabilidade adequada. Assim, o coeficiente de fiabilidade depende da heterogeneidade da amostra.

A estimativa da fiabilidade envolve a utilização de coeficientes de correlação. O Coeficiente de Correlação Intraclassa (CCI) é um desses coeficientes, que varia igualmente entre 0 e 1, refletindo, respetivamente, a ausência e a total correlação e concordância entre as medições.

Da mesma forma que é necessária uma grande variabilidade entre os sujeitos para aumentar o coeficiente de fiabilidade, é necessária uma grande variabilidade entre os sujeitos para aumentar o CCI. Deste modo, uma das principais razões para encontrar valores de CCI baixos é a falta de variabilidade entre os sujeitos. Quando as pontuações dos sujeitos são homogêneas, a fonte de variância entre eles não é significativa, impedindo que a CCI atinja os limites teóricos de 0 e 1 (Portney & Watkins, 2009).

Segundo a taxonomia da COSMIN, a fiabilidade pode ser estimada mediante três abordagens (Mokkink *et al.*, 2010a): consistência interna, fiabilidade (teste-reteste, interobservador e intraobservador) e erro de medição (teste-reteste, interobservador ou intraobservador).

Portney & Watkins (2009), propõem a particularidade do erro de medição ser interpretado no âmbito da estabilidade da resposta. Ou seja, o erro de medição constitui uma forma de avaliar a consistência e a estabilidade das respostas ao longo de tempo, que complementa as medidas de fiabilidade dos instrumentos e dos observadores. Os autores defendem que a estabilidade da resposta é a base para o estudo de todos os outros tipos de fiabilidade.

Ao nível deste trabalho e para fins do estudo da Fiabilidade da QBPDS – VP, apenas será estimada a consistência interna e fiabilidade teste-reteste. O erro de medição teste-reteste também será estimado, embora enquadrado no estudo da interpretabilidade da QBPDS.

A fiabilidade e o erro de medição, interobservador e intraobservador, não são

objeto de estudo, porque a QBPDS é um instrumento de autorrelato, que não requer um observador externo, logo, esse não constitui uma fonte de erro de precisão.

#### **2.6.1.1. Consistência interna**

A consistência interna foi definida como a inter-relação dos itens de um instrumento de medida (Mokkink *et al.*, 2010b). Refere-se à homogeneidade da escala (Cook & Beckman, 2006), ou seja, avalia se os itens de uma escala, planejados para medir o mesmo construto se correlacionam fortemente uns com os outros. Os itens também devem-se correlacionar com a pontuação total (Cook & Beckman, 2006).

O painel de peritos do estudo COSMIN chegou ao consenso de que a estatística da consistência interna só recebe um significado interpretável, quando (1) o inter-relacionamento entre os itens é determinado a partir de um conjunto de itens que no seu conjunto formam um modelo reflexivo e (2) todos os itens avaliam o mesmo construto, isto é, eles formam uma (sub) escala unidimensional (Mokkink *et al.*, 2010b)

Um modelo reflexivo é um modelo, no qual todos os itens são manifestações do mesmo construto subjacente. Nesse caso, os itens, são designados indicadores de efeito e é esperado que sejam altamente correlacionados e intermutáveis (Mokkink *et al.*, 2010b). Geralmente, os autores não descrevem explicitamente se o seu instrumento de autorrelato é baseado num modelo reflexivo ou formativo. Para decidir em qual dos modelos é baseado um determinado instrumento, pode-se questionar se todas as pontuações dos itens mudam aquando de mudanças do construto. Se se considerar que sim, está em questão um modelo reflexivo. Caso contrário, o instrumento é provavelmente baseado num modelo formativo (Mokkink *et al.*, 2010b).

A unidimensionalidade de uma escala pode ser investigada, por exemplo, com uma análise fatorial, mas não com a avaliação da consistência interna. Em vez disso, a unidimensionalidade de uma escala é um pré-requisito para uma interpretação clara das estatísticas de consistência interna.

Na Tabela 1, apresentam-se os estudos de QBPDS que estimaram a consistência interna. No estudo original da escala, de Kopec *et al.* (1995), foi obtido um grau relativamente elevado de correlação item-item (apesar das correlações variarem entre

0,24 a 0,87), assim como uma alta correlação item-total (intervalo de correlações de 0,59 a 0,86). Em estudos posteriores, são reportadas correlações item-item inferiores a 0,80 (Yvanes-Thomas *et al.*, 2002), e correlações item-total variando de 0,28 a 0,83 (Schoppink *et al.*, 1996; Yvanes-Thomas *et al.*, 2002, Bicer *et al.*, 2005), que podem revelar alguma “fragilidade” da escala em relação à sua homogeneidade.

**Tabela 1** – Estimativas de fiabilidade consistência interna da QBPDS em diversos contextos culturais

	Versão da QBPDS	Amostra/ Recolha de dados	Valores de fiabilidade consistência interna
Kopec <i>et al.</i> (1995)	Inglesa	n = 242 DL/ T0= ponto de partida; T1 = 2 a 4 dias depois; T2= 2 a 6 meses depois.	$\alpha$ de Cronbach em T0 (n=225) = 0,96
Schoppink <i>et al.</i> (1996)	Inglesa	n= 120 DLC/ T0 = ponto de partida; T1 = 1 semana; T2= 4 meses.	Correlação Item-total em T0 = 0,58 – 0,83; Correlação Item-total em T2 = 0,58 – 0,81; $\alpha$ de Cronbach em T0 = 0,95; $\alpha$ de Cronbach em T2 = 0,96
Yvanes-Thomas <i>et al.</i> (2002)	Francesa	n= 32 DLC/ Um único momento de avaliação	Correlação Item-item < 0,80 (não mostra matriz de correlação); Correlação Item-total: 0,44 - 0,76
Bicer <i>et al.</i> (2005)	Turca	n = 83 DLC/ Um único momento de avaliação	Correlação Item-total: 0,28- 0,76 (p <0,0001); $\alpha$ de Cronbach = 0,93 (p < 0,0001)
Mousavi <i>et al.</i> (2006)	Iraniana	n = 100 DLC/ T0 = ponto de partida; T1 = 24horas depois (n=31 DLC, aleatoriamente);	$\alpha$ de Cronbach = 0,92
Melikoglu <i>et al.</i> (2009)	Árabe	n = 100 DLC/ T0 = ponto de partida; T1 = 24horas depois;	$\alpha$ de Cronbach (T0) = 0,9405; $\alpha$ de Cronbach (T1) = 0,9537
Rodrigues <i>et al.</i> (2009)	Português do Brasil	n = 54 DL / (sem informação)	$\alpha$ de Cronbach = 0,97

A elevada consistência interna da QBPDS é confirmada nas mais variadas versões linguísticas, com valores de  $\alpha$  de Cronbach superiores a 0,90. Foram determinados valores de: 0,96 (Kopec *et al.*, 1995); 0,95 (Schoppink *et al.*, 1996); 0,93 (Bicer *et al.*, 2005); 0,92 (Mousavi *et al.*, 2006); 0,94 (Melikoglu *et al.*, 2009) e 0,97 (Rodrigues, Michel-Crosato, Cardoso & Traebert, 2009).

No âmbito deste trabalho e com a finalidade de estimar a consistência interna, primeiro será testada a unidimensionalidade da escala (análise fatorial), para poder prosseguir com o cálculo do  $\alpha$  de Cronbach.

### **2.6.1.2. Fiabilidade (teste-reteste)**

A fiabilidade teste-reteste requer duas medições numa população estável e representa a estabilidade do tipo temporal.

A reprodutibilidade ao longo do tempo (fiabilidade do tipo estabilidade temporal) da versão inglesa é considerada boa, com valores de CCI: 0,92 (Kopec *et al.*, 1995), 0,84 (IC 95%: 0,73 - 0,91) (Davidson & Keating, 2002) e 0,94 (IC 95%: 0,90-0,97) (Hicks & Manal, 2009). Os estudos realizados em outras versões linguísticas também revelaram uma boa fiabilidade, com um CCI em geral  $\geq 0,9$  (Schoppink *et al.*, 1996; Rodrigues *et al.*, 2009, Melikoglu *et al.* 2009) (Tabela 2).

No estudo de Davidson & Keating (2002), os valores de CCI foram ligeiramente inferiores aos reportados anteriormente. Essa discrepância pode ter origem nas particularidades da amostra, que incluía participantes com Dor Lombar Aguda (DLA) e DLC; ou no desenho de estudo. Do ponto de vista metodológico, um intervalo teste-reteste de 6 semanas para a estimativa da fiabilidade do tipo estabilidade temporal, de um instrumento que avalia a incapacidade funcional associada a DL, pode ser considerado uma má opção uma vez que a história natural de um novo episódio de DL dita que 60% dos sujeitos experienciam uma melhoria substancial após um período de 6 semanas (Morris, 1983); e efetivamente, vários estudos de intervenção em utentes com DLA reportam resultados entre as 4 e 6 semanas (Stratford., Binkley, Solomon, Finch, Gill & Moreland, 1996). Desta forma, um intervalo tão longo aumenta a probabilidade dos indivíduos experienciarem uma alteração da condição clínica.

**Tabela 2** – Estimativas de fiabilidade estabilidade temporal da QBPDS em diversos contextos culturais

	Versão da QBPDS	Amostra/ Recolha de dados	Valores de fiabilidade estabilidade Temporal
Kopec <i>et al.</i> (1995)	Inglesa	n = 242 DL/ T0= ponto de partida; T1 = 2 a 4 dias depois; T2= 2 a 6 meses depois.	Teste-reteste com intervalo de 1-14 dias, em média 3,8 dias: CCI (n=98) = 0,92
Schoppink <i>et al.</i> (1996)	Holandesa	n= 120 DLC/ T0 = ponto de partida; T1 = 1 semana; T2= 4 meses.	Teste-reteste com intervalo de 1 semana: Coeficiente de correlação de pearson= 0,90 CCI = 0,90
Fritz & Irrgang (2001)	Inglesa	n =67 DL aguda/ T0= ponto de partida; T1 = após 4 semanas de Fisioterapia;	Teste reteste com intervalo de 4 semanas: CCI = 0,55 (IC 95%: 0,20- 0,78)
Davidson & Keating (2002)	Inglesa	n= 106 DL/ T0= ponto de partida; T1 = após 6 semanas	Teste reteste com intervalo de 6 semanas: CCI no grupo classificado “sem alterações” (n=47) = 0,84 (IC 95%: 0,73-0,91) CCI do subgrupo que se identificou “aproximadamente igual” = 0,89 (IC 95%: 0,72 - 0,96)
Mousavi <i>et al.</i> (2006)	Iraniana	n = 100 DLC/ T0 = ponto de partida; T1 (n= 31 aleatoriamente) = 24horas depois	Teste-reteste com 1 dia de intervalo: CCI = 0,86 (p< 0,01)
Melikoglu <i>et al.</i> (2009)	Árabe	n = 100 DLC/ T0 = ponto de partida; T1 = 24horas depois;	Teste-reteste com 1 dia de intervalo: CCI = 0,9221 (p< 0,000)
Hicks & Manal (2009)	Inglesa	n =55 DL (> 62 Anos)/ T0 = ponto de partida; T1 = 1 semana depois.	Teste reteste com intervalo médio de 11 dias: CCI= 0,94 (IC 95%: 0,90-0,97).

De forma a controlar essas mudanças na condição clínica, Davidson & Keating (2002), recorreram a uma escala de percepção global de mudança (aplicada às 6

semanas). No estudo da fiabilidade apenas foram incluídos os participantes que reportaram uma condição estável (grupo “sem alterações”). Pelo motivo referido anteriormente, relativo à melhoria clínica reportada às 6 semanas, este grupo (“sem alterações”) correspondeu a menos de metade da amostra (47 em 106). Além dessa grande redução do número de indivíduos incluídos na estimativa da fiabilidade do tipo estabilidade temporal, a utilização desta estratégia de estabilidade clínica, condicionou, naturalmente, uma menor variância total nas pontuações. Essa menor variância total pode estar na origem da diminuição do valor de fiabilidade teste-reteste.

O estudo de Fritz e Irrgang (2001), cuja fiabilidade teste-reteste foi estimada com 4 semanas de intervalo (CCI = 0,55; IC 95%: 0,20- 0,78), não replicou os elevados valores de fiabilidade reportados nos estudos anteriores. Este estudo (Fritz & Irrgang, 2001), ao contrário dos estudos anteriores que se focaram em utentes com condições crónicas ou mistas, foi aplicado, exclusivamente, a um grupo de utentes com DLA. Deste modo, a diminuição da fiabilidade pode refletir instabilidade da escala quando aplicada a utentes com DLA.

No presente estudo, e para avaliação da fiabilidade teste-reteste da QBPDS – VP, optou-se por um intervalo teste-reteste curto, de 4 a 7 dias e pela aplicação de uma questão de perceção global de mudança no momento do reteste, no sentido de garantir a estabilidade da condição clínica.

### **2.6.1.3. Erro de Medição**

O erro de medição pode ser estimado através de 3 métodos: Erro padrão de medida (EPM), Diferença Mínima Detetável (DMD) e Limites de Acordo (Mokkink et al., 2010a). O painel de peritos COSMIN elege o EPM como o método preferido para medir o erro de medição, apesar de considerarem que os outros dois métodos também são apropriados para medir o erro de medição (Mokkink *et al.*, 2010a). Este painel de peritos argumenta que a DMD é uma transformação linear do EPM e que os Limites de acordo também são diretamente relacionados com o EPM.

Similarmente, Terwee *et al.* (2009) e muitos outros autores nomeiam o EPM para a análise de interpretação de mudanças de pontuação. Segundo os mesmos, o EPM tem muitas propriedades que o tornam numa operação estatística atrativa para

determinar mudança estatística significativa, em questionários sobre o estado de saúde. Em primeiro lugar, o EPM considera a possibilidade de que alguma mudança observada numa medição em particular, seja atribuída a um erro aleatório. Em segundo lugar, o EPM é considerado uma característica fixa de uma medição, independente da amostra sob investigação. Ou seja, é expectável que o EPM se mantenha relativamente constante para todas as amostras extraídas de uma dada população. Adicionalmente, o resultado do EPM é expresso na métrica original da medição (nas mesmas unidades), ajudando à sua interpretação.

A abordagem mais comum para estimar o EPM é a utilização da fórmula seguinte (Fritz & Irrgang, 2001; Davidson & Keating, 2002; de Vet, Terwee, Ostelo, Beckerman, Knol & Bouter, 2006):

$$EPM = DP \times \sqrt{1 - r}$$

Para o parâmetro de fiabilidade ( $r$ ), pode ser utilizado o valor de fiabilidade teste-reteste, ou o valor de  $\alpha$  de Cronbach. No contexto de mudanças de saúde, a fiabilidade teste-reteste é mais apropriada do que o  $\alpha$  de Cronbach, uma vez que as mudanças são baseadas em medições obtidas em dois momentos diferentes (Haley & Fragala-Pinkham, 2006).

Segundo Davidson & Keating (2002) e facultativamente, pode ainda ser calculado, o EPM associado à repetição das medições ( $EPM_{\text{repetição}}$ ), pela seguinte fórmula:  $EPM_{\text{repetição}} = \sqrt{2} \times EPM$ , que considera o erro adicional introduzido pela repetição das medições, em 2 momentos diferentes (Haley & Fragala-Pinkham, 2006).

Os estudos que estimaram o EPM na QBPDS são apresentados no ponto 2.9., no âmbito da interpretabilidade da QBPDS. O EPM surge sistematicamente como um cálculo precedente da estimativa da DMD. Assim, neste estudo que também tem a pretensão de estimar a DMD da QBPDS, estabelece-se a necessidade do cálculo do EPM e  $EPM_{\text{repetição}}$  da QBPDS.

### **2.6.2. Validade**

A validade refere-se ao grau de legitimidade com que se pode confiar nos resultados de um instrumento, quando interpretados com uma finalidade específica (Cook & Beckman, 2006). Segundo Cook & Beckman (2006) a validade não é uma propriedade do instrumento, mas sim das pontuações do instrumento e da inferência ou interpretação dos resultados.

A taxonomia proposta por COSMIN, apresenta 3 propriedades psicométricas de validade: validade de conteúdo; validade de critério e validade de construto.

Um instrumento só pode ser válido se o conceito que usa é claramente definido antes do instrumento ser desenvolvido. O conceito utilizado na QBPDS é consistente com a definição de incapacidade da *International Classification Impairment Disability Handicap* (ICIDH), publicada pela WHO (1980). A QBPDS explora a área limitada de “atividades físicas simples” (Yvanes-Thomas *et al.*, 2002).

Uma vez que a validade da QBPDS – VP, não é objeto de estudo para este trabalho, toda a informação relativa às estimativas de validade da QBPDS, noutras versões linguísticas, estão disponíveis no Apêndice A. Os resultados ao nível da validade de construto, evidenciam que, a QBPDS tem correlações moderadas a fortes, com vários questionários de incapacidade; e uma correlação fraca a moderada, com a intensidade da dor.

### **2.7. Poder de resposta**

Nas últimas décadas tem sido dado maior ênfase à capacidade de um instrumento de medida detetar mudanças, porque o objetivo final de uma escala de avaliação é distinguir entre os utentes ou grupos de utentes cujo estado de saúde, melhorou, agravou ou manteve-se estável.

Para a maioria dos autores, o poder de resposta refere-se à capacidade de um instrumento de medida detetar mudanças; mas há uma grande variedade de opiniões quanto à natureza da mudança que está a ser detetada. Vários autores (Wyrwich *et al.*, 2005; Lauridsen *et al.*, 2006; Terwee *et al.* 2007 citado por Demoulin, Ostelo,

Knottnerus & Smeets, 2010), referem-se à capacidade de detetar **mudanças clínicas relevantes** ao longo do tempo (mesmo que as mudanças sejam pequenas); e de diferenciá-las dos erros de medição (Terwee *et al.* 2007 citado por Demoulin *et al.*, 2010). Outros, definem o poder de resposta como a capacidade de um **instrumento detetar mudanças (significativas) com precisão**, quando essas ocorrem (de Bruin *et al.* 1997 citado por Beaton; Davidson & Keating, 2002; e Mokkink *et al.*, 2010).

Para efeitos deste trabalho, opta-se pela definição original de poder de resposta, de de Bruin *et al.* (1997 citado por Beaton *et al.*, 2001), que foi adotada em dois estudos de revisão de taxonomia (Beaton *et al.*, 2001; Mokkink *et al.*, 2010). Essa definição não especifica a natureza da mudança, mas exige que haja alguma determinação quanto à ocorrência de mudança, de modo que o poder de resposta de um instrumento inclui a capacidade de detetar mudanças reais e/ou mudanças clinicamente significativas, desde que seja com precisão. A diferenciação dessas mudanças é apresentada de seguida.

### **2.7.1. Interpretabilidade (das mudanças)**

A interpretabilidade é o grau com que se pode atribuir significado qualitativo às pontuações quantitativas (Terwee *et al.* 2007 citado por Demoulin *et al.*, 2010). Segundo a taxonomia da COSMIN, a interpretabilidade não representa uma propriedade psicométrica (Mokkink, *et al.*, 2010a), no entanto, é amplamente referida pela literatura pelo seu valor na interpretação de mudanças de pontuação. O facto das medidas serem fiáveis e consistentes, não permite inferir que uma alteração de uma determinada medida num segundo momento de avaliação de um utente, é uma alteração “real”. Para avaliar os efeitos de uma intervenção, o profissional de saúde precisa de conhecer quando é que a mudança da pontuação representa uma mudança “real”. Apenas uma mudança superior ao erro de medida, pode ser considerada uma diferença “real”.

Considera-se que um indivíduo experimentou uma mudança real quando a diferença entre a pontuação anterior e a pontuação atual excede a combinação da variabilidade intraindividual associada a ambas as medições. A magnitude deste erro designa-se de “**Diferença Mínima Detetável**” (**DMD**) (Finch, Brooks, Stratford & Mayo, 2002).

No entanto, não se pode depreender que essa diferença real representa uma mudança “importante” para os utentes ou profissionais de saúde, ou do ponto de vista social. Existe portanto a necessidade de determinar a quantidade mínima de mudança na pontuação, que os utentes ou profissionais de saúde consideram “importante” (Terwee *et al.*, 2009) e distingui-la da DMD.

Considera-se que um utente obteve uma mudança clinicamente importante, quando a diferença entre a pontuação anterior e a pontuação atual é relevante para o indivíduo. A magnitude dessa diferença designa-se “**Diferença mínima clinicamente importante**” (DMCI) (Finch, Brooks, Stratford & Mayo, 2002).

A compreensão destes dois conceitos anteriores, na literatura disponível, pode ser um esforço desconcertante, particularmente devido ao grande número de termos e anacronismos que são usados em diferentes áreas e tradições (Haley & Fragala-Pinkham, 2006).<sup>4</sup> Por outro lado, a maioria dos autores (Stratford *et al.*, 1996; Haley & Fragala-Pinkham, 2006; Terwee *et al.*, 2009; Hicks & Manal, 2009) assume a terminologia que distingue duas designações para dois valores distintos, calculados por diferentes métodos. Para efeitos deste trabalho, será assumida a mesma distinção: **DMD e DMCI**.

#### 2.7.1.1. DMD

A **diferença mínima detetável (DMD)** é a diferença mínima que pode ser detetável entre duas medições e que não se deve aos erros de medição (Haley & Fragala-Pinkham, 2006). Segundo Stratford *et al.* (1996), a DMD pode ser definida como a quantidade de mudança requerida para que haja 90% (ou 95%) de confiança que a alteração observada nas pontuações reflete uma mudança real na variável subjacente.

A fórmula adotada para a estimativa da DMD, através do EPM, é unânime para os vários autores (Fritz & Irrgang, 2001; Davidson & Keating, 2002; Haley & Fragala-Pinkham, 2006; de Vet *et al.*, 2006; van der Roer, Ostelo, Bekkering, van Tulder & de Vet, 2006); Kovacs *et al.*, 2007; Demoulin *et al.*, 2010; Mokkink *et al.*, 2010a):

---

<sup>4</sup> Na literatura esses valores de mudança da pontuação são muitas vezes sobrepostos e englobados na mesma classificação, designados por valores de “Diferença Clínica Significativa” (Wyrwich *et al.*, 2005; Haley & Fragala-Pinkham, 2006) ou de “Diferença mínima importante” (Demoulin *et al.*, 2010).

$$DMD = Valor z_{\text{Nível de confiança}} \times \sqrt{2} \times EPM$$

Esta fórmula propõe que o EPM seja multiplicado por  $\sqrt{2}$ , porque reconhece que há um erro associado às duas medições; e ainda, por um valor z, de um determinado percentil, equivalente ao nível de confiança, com que se pretende estimar a DMD.

A escolha do intervalo de confiança na literatura é dúbia. Alguns investigadores recomendam  $1,65 \times \sqrt{2} \times EPM$ , para IC de 90% (Fritz & Irrgang, 2001; Davidson & Keating, 2002; Hicks & Manal, 2009) e outros,  $1,95 \times \sqrt{2} \times EPM$ , para corresponder a um intervalo de confiança de 95% (de Vet *et al.*, 2006; Kovacs *et al.*, 2007; van der Roer *et al.*, 2006; Demoulin *et al.*, 2010; Mokkink *et al.*, 2010a).

O intervalo de confiança de 95%, aumenta a precisão da estimativa do valor de DMD (Steffen & Seney, 2008). Neste estudo, será adotado um intervalo de confiança de 90%, na estimativa da DMD da QBPDS – VP, porque considerou-se a precisão dessa estimativa suficiente.

Na interpretação da DMD, uma mudança observada num utente que seja inferior a esse valor, é indistinguível do erro de medição. Por consequência, é considerado que o utente não sofreu mudança na sua condição (Stratford *et al.*, 1996). Por outro lado, as alterações na pontuação superiores ao valor de DMD devem-se à melhoria do utente no teste, mais do que aos erros de medição. Os erros de medição incluem a expectável ou típica variabilidade no desempenho dos utentes (Steffen & Seney, 2008).

Como a DMD é expressa nas mesmas unidades do questionário, melhora a interpretação da mudança das pontuações.

Para avaliar a DMD, primeiro é necessário medir a fiabilidade teste-reteste. As fontes de erro, que perturbam essa fiabilidade, podem incluir inconsistências causadas pela condição física ou mental dos participantes, variações nos procedimentos do teste, ou erros por parte do avaliador. Manter a consistência e usar protocolos padronizados para o teste-reteste, pode melhorar a sua fiabilidade. Por exemplo, deve-se recorrer sempre ao mesmo avaliador, às mesmas instalações, às mesmas ordens de avaliação e à mesma hora do dia.

### 2.7.1.2. DMCI

A DMCI<sup>5</sup> representa a menor alteração da pontuação, que é considerada importante e/ou significativa para o utente (Haley & Fragala-Pinkham, 2006); ou, uma alteração que faz uma diferença real na vida dos utentes e familiares, e que é suficientemente grande para ter uma implicação no tratamento ou cuidados do utente (Wyrwich *et al.*, 2005).

A DMCI é definida por Jaeschke, Singer e Guyatt (1989), como “a menor diferença na pontuação no domínio de interesse, que os utentes percebem como benéfica e que vai determinar, na ausência de importunos efeitos colaterais e excessivos custos, uma alteração na intervenção do utente”

Esta medida é julgada pelo utente ou peritos na matéria e determina-se questionando ou observando o utente (Steffen & Seney, 2008). Ou seja, a sua determinação requer que se ligue a mudança da pontuação a uma âncora externa significativa (Lauridsen *et al.*, 2006). De acordo com o painel de peritos do estudo COSMIN, quando se estuda a DMCI de instrumentos de autorrelato, devem ser os utentes a decidir “quanto é importante”, e portanto devem ser utilizadas âncoras centradas na perceção de mudança do utente (Mokkink *et al.*, 2010a). Os instrumentos com essas características são genericamente designados de escalas de avaliação global de mudança (estas escalas são abordadas nos métodos baseados na âncora, na página 30).

Tem sido demonstrado, que DMCI varia entre os utentes e grupos de utentes e que tem limitada generalização (Haley & Fragala-Pinkham, 2006). Podem ser obtidos diferentes valores de DMCI, utilizando diferentes âncoras e/ou diferentes métodos. Desta forma, é importante que nas investigações se corroborem os resultados entre métodos e múltiplas âncoras. Além disso, uma estimativa DMCI será associada a um grau de incerteza e de variação (Haley & Fragala-Pinkham, 2006).

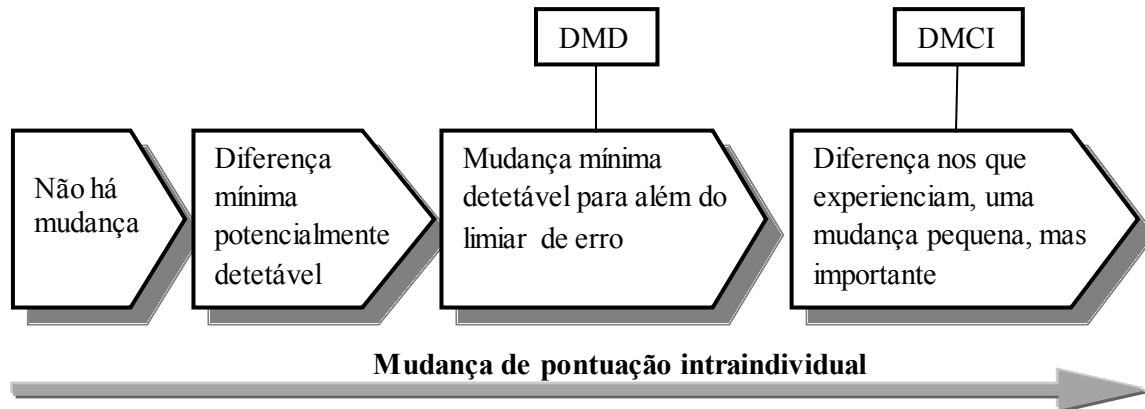
O painel de peritos da COSMIN convencionou que a DMCI, centra-se na interpretação de mudanças de pontuação, mas não julga a validade da mudança de pontuação (Mokkink *et al.*, 2010b).

---

<sup>5</sup> A DMCI também é designada na literatura por “diferença mínima importante” (Haley & Fragala-Pinkham, 2006; Terwee *et al.*, 2009).

### 2.7.1.3. Relação DMD/ DMCI

Na interpretação de mudanças de pontuações de um momento para outro, a quantidade de mudança pode ser representada ao longo de um referencial contínuo, conforme esquematizado na Figura 2.



**Figura 2** - Mudanças intraindividuais ao longo de um referencial contínuo (adaptado de Portney & Watkins, 2009)

O referencial contínuo começa com a diferença mínima potencialmente detetável que depende da precisão do instrumento de medida utilizado (Portney & Watkins, 2009). No caso da QBPDS, a diferença mínima potencialmente detetável é de um ponto, uma vez que não podem ser contadas frações de um ponto.

Observada uma determinada diferença nas pontuações, é necessário confirmar que essa diferença reflete uma mudança verdadeira e que não se deve a um mero erro de medição aleatório. Ou seja, visam-se mudanças que ultrapassem a DMD.

A DMD é apenas o ponto de partida para definir mudanças de pontuações, porque normalmente é muito pequena para representar uma diferença importante na condição do utente. Progredindo no referencial contínuo, pretende-se também identificar o valor que identifica uma DMCI (Portney & Watkins, 2009).

Embora seja dada uma crescente atenção à interpretação das pontuações do questionário e determinação da DMCI, a ligação com o erro de medição raramente é realizada (Terwee *et al.*, 2009). É importante considerar a DMCI em relação à DMD de um questionário, porque o fato da relação da DMCI ser superior à DMD, nem sempre se

verifica. Se a DMD é menor que a DMCI, é possível distinguir uma mudança clinicamente importante de um erro de medição com uma grande quantidade de certeza. No entanto, se a DMD é maior do que a DMCI, é muito mais difícil distinguir mudanças observadas iguais ou superiores à DMCI do erro de medição, porque há uma hipótese considerável de que a mudança observada se deva ao erro de medição.

Este estudo propõe-se a estimar a DMD através do cálculo da EPM e estimar a DMCI utilizando uma âncora centrada na percepção global de mudança do utente. Este estudo ambiciona ainda estabelecer a relação entre essas duas medidas e aplicá-la na interpretabilidade de mudanças da pontuação da QBPDS – VP.

### **2.7.2. Métodos associados à estimativa do poder de resposta**

Com a definição de poder de resposta adotada neste estudo, aceita-se que os métodos de exploração do poder de resposta possam ser classificados como métodos que avaliam a mudança por si só (métodos baseados na distribuição), ou como métodos que medem mudanças clínicas significativas (métodos baseados na âncora) (Davidson & Keating; 2002). Originalmente introduzidos por Eva Lydick e Robert Epstein em 1993 (citados em Wyrwich *et al.*, 2005), estes métodos representam uma perspetiva estatística e clínica respetivamente. É reconhecido o valor das duas abordagens complementares e distintas. As duas apresentam vantagens e limitações, pelo que os autores recomendam a utilização de ambas (Tabela 3).

Os **métodos baseados na distribuição** pretendem expressar as alterações na pontuação em termos da distribuição da amostra subjacente, quer em unidades de desvio padrão entre indivíduos, quer em unidades de desvio padrão intraindividuais, ou através de alguma variação do “Erro Padrão de Medida”. Este método baseia-se na significância estatística, na variabilidade da amostra e na precisão da medição (Haley & Fragala-Pinkham, 2006).

Na literatura são descritos alguns métodos baseados na distribuição: Tamanho do efeito (TE) (Stratford, Binkley & Riddle, 1996; Wyrwich *et al.*, 2005); Resposta média padronizada (RMP) (Stratford, Binkley & Riddle, 1996; Davidson & Keating, 2002; Lauridsen *et al.*, 2006, Wilhelm *et al.*, 2010) e o EPM (Davidson & Keating, 2002; Wyrwich *et al.*, 2005; Haley & Fragala-Pinkham, 2006). Os dois primeiros são

representações relativas de mudança (isto é, sem unidades), que permitem estimar o poder de resposta.

Com base no EPM é calculado um dos índices de mudança mais comuns dos métodos baseados na distribuição, a DMD.

A maior desvantagem dos métodos baseados na distribuição é que eles são puramente medidas estatísticas da magnitude da mudança e não expressam nada sobre a sua importância clínica (Beaton *et al.*, 2001).

**Tabela 3** – Métodos associados à estimativa do poder de resposta (Stratford, Binkley & Riddle, 1996; Mokkink *et al.*, 2010a).

Métodos baseados na distribuição	Métodos baseados em âncora
- Tamanho do efeito ( <i>Effect Size</i> );	- Coeficiente de correlação com a âncora
- Resposta média padronizada ( <i>Standardized Response Mean</i> );	- Curva ROC ( <i>Receiver Operating Characteristic</i> )**;
- Erro Padrão de Medida ( <i>Standard Error of Measurement</i> )*;	- Índice de poder de resposta de Guyatt ( <i>Guyatt's Responsiveness Index</i> );
- Eficiência Relativa ( <i>Relative Efficiency</i> ).	- Coeficiente de poder de resposta de Norman ( <i>Norman's Responsiveness Coefficient</i> );

\* Também é utilizado na estimativa da DMD; \*\*É um método que permite calcular a DMCI;

Os métodos baseados na âncora (ou Métodos baseados em critério) ligam ou relacionam a mudança da pontuação a um critério padrão externo e independente, para “ancorar” o significado de importância clínica. Essa âncora pode ser um evento, taxa ou condição (Wyrwich *et al.*, 2005), que deve ser, por si só, interpretável e pelo menos moderadamente correlacionada com o instrumento (Haley & Fragala-Pinkham, 2006).

Estes métodos incluem estudos com um desenho, quer transversal quer longitudinal, que ligam a mudança da pontuação a uma âncora externa significativa (Lauridsen *et al.*, 2006). Têm sido utilizadas muitas âncoras diferentes, que podem variar desde a percepção do próprio utente, que aparentemente é a mais óbvia, à percepção do familiar e do próprio profissional de saúde. Por exemplo, pode-se comparar a

mudança da pontuação com os objetivos do tratamento pré-definidos (se foram atingidos ou não), com a percepção global da mudança dos indivíduos, ou com indicadores mais objetivos, como valores de laboratório ou marcadores de saúde, dependendo do tipo de instrumento em causa (Lauridsen *et al.*, 2006; Haley & Fragala-Pinkham, 2006).

Segundo o painel de peritos da COSMIN, não há uma “medida de padrão de ouro” para “instrumentos de autorrelato de avaliação do estado de saúde”, que permita comparar as mudanças na pontuação desses instrumentos com as mudanças verificadas numa “medida de padrão de ouro”. Em alternativa à “medida de padrão de ouro”, usam-se as escalas de avaliação global de mudança como âncora externa (Mokkink *et al.*, 2010b).

As escalas de avaliação global de mudança são frequentemente utilizadas na investigação clínica, particularmente na área músculo-esquelética e são recomendadas como medida de resultado central dos estudos de dor crónica (Kamper, Maher & Mackay, 2005). Essas escalas são desenhadas para quantificar a melhoria, ou agravamento, da condição clínica do utente ao longo do tempo (desde o início até ao momento do reteste), através de uma única pergunta<sup>6</sup>. A magnitude dessa mudança é assinalada numa escala de *likert* com um número de pontos variável (por exemplo, de 7 ou de 15 pontos), variando desde “muito pior” a “muito melhor”.

Previamente à aplicação deste tipo de critério externo é importante que se defina quais as opções da escala que representam uma mudança clínica, quer positiva ou negativa, como a ausência de mudança. Um problema verificado na literatura é a definição do que é uma mudança importante e da organização dessas categorias. Por exemplo, alguns autores classificam que as opções "mínima melhoria" ou "ligeiramente melhor" devem ser interpretadas como uma “mudança clinicamente importante”, enquanto outros autores incluem as mesmas opções na categoria de "nenhuma mudança importante" ou "sem alterações".

No âmbito destes métodos, baseados na âncora, também há possibilidade de calcular a DMCI (Davidson & Keating, 2002; Lauridsen *et al.*, 2006; Haley & Fragala-Pinkham, 2006).

---

<sup>6</sup> Por esse motivo, a escala de avaliação global de mudança também é reportada na literatura por “Questão global de transição”

As abordagens de âncora não tomam em consideração a precisão da medição. Por conseguinte, estes métodos não implicam necessariamente uma significância estatística. Para ultrapassar esse problema, é recomendado que ambos os métodos sejam calculados separadamente.

Neste estudo, considera-se o método baseado na distribuição, em termos da interpretabilidade da escala, através do EPM para o cálculo da DMD. No âmbito da estimativa do Poder de resposta da escala QBPDS, opta-se apenas pelos métodos baseados na âncora. Esta decisão deriva das conclusões recentes da COSMIN, que considera que as medidas derivadas de métodos baseados na distribuição, são parâmetros “inapropriados” para estimar o poder de resposta - “Essas medidas, são consideradas medidas da magnitude da intervenção ou do evento, mais do que medidas da qualidade de medição do instrumento” (Mokkink *et al.*, 2010b); onde demasiada atenção é colocada na magnitude da sua estimativa numérica e pouca atenção é prestada à natureza e significado da mudança que está a ser quantificada (Beaton *et al.*, 2001).

No próximo tópico, reúne-se a evidência disponível sobre o poder de resposta de QBPDS e desenvolve-se os métodos que serão utilizados neste estudo.

## **2.8. Poder de resposta da QBPDS**

Os poucos estudos que avaliam o poder de resposta da QBPDS (Fritz & Irrgang, 2001; Davidson & Keating, 2002; van der Roer *et al.*, 2006, Demoulin *et al.*, 2010) são muito heterogêneos em relação a vários parâmetros, por exemplo, em relação à população em estudo (DL aguda, subaguda e crónica), à escolha e interpretação do critério externo usado para determinar se um dado utente melhorou ou não, e aos métodos utilizados (baseados na distribuição e baseados na âncora), resultando em uma ampla gama de valores.

Alguns autores não diferenciam poder de resposta e interpretabilidade e apresentam resultados de forma pouco clara em relação ao objetivo. Neste trabalho, opta-se por separar os dois conceitos para facilitar a comparação dos resultados.

### **2.8.1. Poder de resposta da QBPDS - Métodos baseados na distribuição**

Os estudos disponíveis sobre a estimativa do poder de resposta da QBPDS através de Métodos baseados na distribuição, utilizaram o TE, a RMP e o EPM.

Tal como já foi referido, apesar do EPM ser considerado um método baseado na distribuição, os seus resultados são mais utilizados na interpretabilidade da mudança de pontuações. No âmbito do poder de resposta da QBPDS, consideram-se apenas os resultados dos métodos TE e RMP. Ambos são representações relativas de magnitude de mudança (isto é, sem unidades) que permitem estimar poder de resposta. A Tabela 4 reúne os resultados do TE e do RMP da QBPDS.

O TE obtém-se pela divisão da diferença média das pontuações (entre os dois momentos de medição), pelo desvio padrão das medições iniciais (Stratford, Binkley & Riddle, 1996). Desta forma, a sua estimativa pode variar entre amostras com diferente nível de variabilidade no ponto de partida (Portney & Watkins, 2009).

A RMP é calculada através da divisão da diferença média das pontuações (entre os dois momentos de medição), pelo desvio padrão das diferenças de pontuações (Stratford, Binkley & Riddle, 1996). Segundo Portney & Watkins (2009), podem decorrer diferentes resultados da RMP, na consequência da diferente variabilidade associada à mudança das pontuações. Os mesmos autores entendem que a RMP varia em função da efetividade do tratamento.

Os valores absolutos do TE e da RMP são interpretados da mesma maneira. Entre 0,20 e 0,50 é considerado um poder de resposta pequeno, entre 0,51 e 0,80 é moderado e acima de 0,80 é elevado (Cohen *et al.* 1977 citado por Wilhelm *et al.*, 2010).

As estimativas de TE e RMP da QBPDS, efetuadas com os resultados de utentes com DLC (Chaory *et al.*, 2004; Wilhelm *et al.*, 2010), determinaram valores compreendidos entre 0,62 e 0,80, que se traduzem num poder de resposta moderado.

Os únicos estudos, que estimaram a RMP da QBPDS, aplicada em utentes com DL, sem critérios de duração dos sintomas da DL, obtiveram valores de RMP de 0,47 (Kopec *et al.*, 1995) e de 0,49 (Davidson & Keating, 2002), para um intervalo de 2 a 6 meses e de 6 semanas, respetivamente. Esses valores correspondem a um poder de

resposta pequeno, que se aproxima do moderado. A estimativa da RMP para um intervalo de 2 a 4 dias (Kopec *et al.*, 1995), resultou num valor muito baixo (0,29), que decorre, provavelmente, da menor variabilidade na mudança das pontuações verificada nesse curto espaço de tempo.

Tabela 4 – Estimativas do poder de resposta da QBPDS através de Métodos baseados na distribuição: TE e RMP

	Amostra/ Intervalo	Métodos baseados na distribuição	
		TE	RMP
Kopec <i>et al.</i> (1995)	n = 242 DL/ T0= ponto de partida; T1 = 2 a 4 dias depois; T2= 2 a 6 meses depois.	-----	T0-T1 = 0,29 T0-T2 = 0,47
Davidson & Keating (2002)	n= 106 DL; 6 semanas	-----	0,49 (IC 95%: 0,47-1,44)
Chaory <i>et al.</i> (2004)	n= 70 DLC; 5 ou 6 semanas (treino funcional)	0,80	0,78
Wilhelm <i>et al.</i> (2009)	n= 30 DLC; 4 semanas (treino funcional)	0,62	0,80

### 2.8.2. Poder de resposta da QBPDS - Métodos baseados na âncora

Até ao momento, os métodos baseados na âncora utilizados para estimar o poder de resposta da QBPDS, foram o Coeficiente correlação âncora, a Curva ROC e o Índice de poder de resposta de Guyatt (GRI) (Tabela 5).

O método “Coeficiente correlação âncora”, correlaciona as diferenças individuais nas pontuações da QBPDS, com o nível de perceção global de mudança (ou outro critério de padrão de mudança). Utiliza o coeficiente de correlação,  $r$  de Pearson ou  $\rho$  de Spearman, conforme o critério de padrão de mudança é quantificado numa escala intervalar ou ordinal, respetivamente. A intensidade da associação representa o poder de resposta do instrumento. O estudo de Fritz e Irrgang (2001), estimou o coeficiente de correlação de Pearson, no valor de 0,67. Esse resultado indica uma

correlação positiva e moderadamente forte, que dessa forma expressa um poder de resposta da QBPDS também moderadamente forte.

A curva ROC (“*Receiver Operating Characteristic*”) é um dos métodos baseados em âncora mais utilizados para reportar mudanças ao longo do tempo, e representa uma perspectiva alternativa de interpretar o poder de resposta. Este método considera que o poder de resposta é uma forma de discriminar entre os utentes que melhoraram e não melhoraram. Assim, utiliza a informação da âncora externa, para dicotomizar a amostra em grupo “sucesso” e grupo “insucesso” e depois estabelecer um gráfico da amostra de “sensibilidade” vs “1-especificidade” (taxa de falsos positivos). Esse gráfico avalia a relação entre o sucesso dos verdadeiros positivos e os erros dos falsos positivos, para os vários pontos de corte, que representam vários valores de diferença de pontuação (Lauridsen *et al.*, 2006).

Uma sensibilidade no valor de 1, indica que todos os sujeitos com uma mudança na pontuação superior ao limite de pontuação selecionada (ponto de corte), relataram melhorias clínicas (níveis de perceção de melhoria considerados como sucesso). Uma especificidade de 1 indica que todos os indivíduos com uma mudança de pontuação inferior ao ponto de corte selecionado, relataram “sem alterações” (ou sem grandes alterações) (níveis de perceção de melhoria considerados como insucesso) (Demoulin *et al.*, 2010)

A Área Abaixo da Curva (AAC) reflete a capacidade do instrumento para discriminar os sujeitos que melhoraram e os indivíduos que não sofreram alterações na sua condição. Se a AAC for igual a 0,5 o modelo não consegue discriminar os sujeitos que melhoraram vs os que não melhoraram. Por outro lado, quanto mais próximo a área for de 1, maior é a capacidade do modelo para fazer essa discriminação (Marôco, 2010).

Os estudos que reportam resultados do método curva ROC, aplicado à escala de QBPDS, referem valores de AAC de: 0,87 ( $\pm 0,048$ ) (Fritz & Irrgang, 2001); 0,74 (IC 95%: 0,64-0,84) (Davidson & Keating, 2002); 0,850 (IC 95%: 0,786 – 0,914) e 0,856 (IC 95%: 0,791 - 0,922) (Demoulin *et al.*, 2010). Todos esses valores expressam uma precisão relativamente elevada da QBPDS em detetar mudanças clinicamente significativas.

O Índice de GRI é outro método baseado na âncora, mas apenas foi referido num estudo da QBPDS (Fritz & Irrgang, 2001). É definido como a razão da mudança média

em utentes identificados como melhores, a dividir pelo desvio padrão da mudança dos utentes identificados como estáveis. Um valor de GRI maior, indica maior poder de resposta (Fritz & Irrgang, 2001). No estudo de Fritz e Irrgang (2001), com uma amostra de utentes com DLA, foi atribuída à QBPDS um índice de GRI de 1,82, que serviu de comparação com o índice de GRI de outro questionário Oswestry, no valor de 3,49. A diferença nos índices é sugestiva de que o questionário Oswestry, tem maior poder de resposta, do que a QBPDS.

**Tabela 5** – Estimativa do poder de resposta da QBPDS através de Métodos baseados na âncora.

	Amostra/ Intervalo	Poder de Resposta: Métodos baseados na âncora		
		Coefficiente Correlação Âncora	Curva ROC	Índice GRI
Fritz & Irrgang (2001)	n =67 DLA; 4 semanas de Fisioterapia	Coefficient e de correlação de Pearson = 0,67	AAC = 0,87 ( $\pm 0,048$ );	1,82 (IC 95%: 1,10 -2,55)
Davidson & Keating (2002)	n= 106 DL; 6 semanas		AAC (n=99) = 0,74 (IC 95%: 0,64-0,84)	—
Demoulin <i>et al.</i> (2010)	n = 212 DLC; 10 semanas de tratamento multidisciplin ar		AAC= 0,850, (IC 95%: 0,786 – 0,914); Curva ROC (mudanças de pontuação expressa em %): AAC = 0,856 (IC 95%: 0.791 - 0.922)	

## 2.9. Interpretabilidade da QBPDS

### 1.9.1. Estimativa da DMD da QBPDS

A literatura é coerente na estimativa da DMD. Nos vários estudos de estimativa da DMD da QBPDS, essa é precedida pelo método baseado na distribuição: EPM. Para todos os estudos foi calculado o EPM decorrente da repetição de medições da QBPDS,

de um grupo que é considerado *a posteriori* estável do ponto de vista clínico, com base na resposta simultânea de uma questão de percepção global de mudança (Tabela 6).

**Tabela 6** –Estimativas do EPM e DMD da QBPDS em diversas aplicações da escala.

	<b>Amostra/ Intervalo</b>	<b>Questão global de transição (percepção global de mudança)</b>	<b>EPM</b>	<b>DMD</b>
<b>De 15 pontos/ 3 categorias:</b>				
Fritz & Irrgang (2001)	n= 67 DLA; 4 semanas de Fisioterapia;	<ul style="list-style-type: none"> <li>• Agravamento clínico: -7, -6, -5, -4;</li> <li>• Sem alterações: -3, -2, -1, 0 (aproximadamente igual), +1, +2, +3;</li> <li>• Melhoria clínica: +4, +5, +6, +7.</li> </ul>	EPM (categoria “sem alterações”) = 13, 08 (IC 95%: 10, 54 – 17,47)	DMD <sub>90</sub> = 30, 52
<b>De 7 pontos/ 3 categorias:</b>				
Davidson & Keating (2002)	n= 106 DL; 6 semanas	<ul style="list-style-type: none"> <li>• Melhoria clínica: 1) completamente resolvido; 2) muito melhor; 3) melhor;</li> <li>• Sem alterações: 4) ligeiramente melhor; 5) aproximadamente igual; 6) ligeiramente pior;</li> <li>• Agravamento clínico: 7) Muito Pior.</li> </ul>	EPM (categoria “sem alterações”) = 8 (IC 95%: 6-10)	DMD <sub>90</sub> (categoria “sem alterações”) = 19 (IC 95%: 14-24)
			EPM (“aproximadamente igual”) =7 (IC 95%: 4-11)	DMD <sub>90</sub> (“aproximadamente igual”) = 15 (IC 90%: 9-24)
<b>De 6 pontos/ 3 categorias:</b>				
van der Roer <i>et al.</i> (2006)	n = 442 (DLA e DLC); 12 semanas	<ul style="list-style-type: none"> <li>• Melhoria clínica: 1) “completamente resolvido”; 2) muito melhor”;</li> <li>• Sem alterações: 3) ligeiramente melhor”; 4) “aproximadamente igual”; 5) “ligeiramente pior”;</li> <li>• Agravamento clínico: 6) “Muito pior”</li> </ul>	-----	DMD <sub>95</sub> para <b>DLA</b> (categoria “sem alterações”; n=52) = 32,9 (IC 95%: 24,6 – 49,8)
			-----	DMD <sub>95</sub> para <b>DLC</b> (categoria “sem alterações”; n=62) = 24,6 (IC 95%: 19,9 – 32,4)

<b>De 15 pontos/3 categorias:</b>				
Hicks & Manal, 2009	n=55 DL (> 62 A); 1 semana	<ul style="list-style-type: none"> <li>•Agravamento clínico: -7, -6, -5, -4;</li> <li>•Sem alterações: -3, -2, -1, 0 (“aproximadamente igual”), +1, +2, +3.</li> <li>•Melhoria clínica: +4, +5, +6, +7</li> </ul>	EPM (categoria “sem alterações”) = 4,73;	DMD <sub>90</sub> = 11,04
<b>De 7 pontos/ 3 categorias:</b>				
Demoulin <i>et al.</i> (2010)	N= 212 DLC; 10 semanas de tratamento;	<ul style="list-style-type: none"> <li>•Agravamento clínico: 1) Pior do que nunca; 2) Muito pior;</li> <li>•“Sem alterações” (ou sem alterações importantes): 3) Ligeiramente pior; 4) sem alterações; 5) Ligeiramente melhor;</li> <li>•Melhoria clínica: 6) muito melhor; e 7) completamente recuperado;</li> </ul>	EPM (categoria “sem alterações”) = 5,7 (Calculado de acordo com Terwee <i>et al.</i> , 2009)	DMD <sub>95</sub> (categoria “sem alterações”; n=143) = 15,8

As estimativas da DMD da QBPDS variam em 2 aspetos. Em relação ao nível de confiança que é fixado, em 90 ou 95% (DMD<sub>90</sub> ou DMD<sub>95</sub>, respetivamente) e em relação ao grupo considerando como referência da estabilidade clínica (de acordo com os itens da escala de perceção global de mudança que são pré-definidos, como tal). Apesar das variações, segundo a literatura a DMD é predominantemente estimada com base no EPM obtido a partir dos participantes que se avaliam na categoria “sem alterações” (incluindo os utentes que assinalaram qualquer umas das 3 opções da pergunta de escala de avaliação global de mudança).

No cálculo primordial da EPM, obtiveram os seguintes valores: 13, 08 (IC 95%: 10,54–17,47) (Fritz & Irrgang, 2001); 8 (IC 95%: 6 –10) e 7 (IC 95%: 4-11), consoante foi considerado para o seu cálculo, a categoria “sem alterações” ou a opção “aproximadamente igual” respetivamente (Davidson & Keating, 2002); 4,73 (Hicks & Manal, 2009) e 5,7 (Demoulin *et al.*, 2010).

Os valores de DMD reportados na literatura (Tabela 5), tomam valores entre 11,04 (Hicks & Manal, 2009) e 32,9 (van der Roer *et al.*, 2006), independentemente do nível de confiança estabelecido. Apesar de haver uma variedade considerável, os

valores dos extremos foram obtidos de estudos com particularidades. Os valores superiores: DMD<sub>95</sub> de 32,9 pontos (van der Roer *et al.*, 2006) e DMD<sub>95</sub> de 30,52 pontos (Fritz & Irrgang, 2001) foram obtidos a partir de amostras de utentes com DLA; e o valor inferior, DMD<sub>90</sub> de 11,04 pontos, foi aferido de uma amostra geriátrica (Hicks & Manal, 2009).

No estudo de Davidson e Keating (2002), cuja amostra não particulariza a duração dos sintomas dos participantes com DL, comparam os resultados decorrentes da definição de estabilidade clínica segundo a categoria "sem alterações" e da definição de estabilidade clínica segundo o nível de percepção de melhoria "aproximadamente igual". Os resultados obtidos revelam que o EPM derivado da primeira definição, no valor de 8 pontos (IC 95%: 6-10), determinou uma DMD<sub>90</sub> de 19 pontos (IC 90%: 14-24); ao passo que o EPM do grupo "aproximadamente igual", de 7 pontos (IC 95%: 4-11), determinou uma DMD<sub>90</sub> de 15 pontos (IC 90%: 9-24). Estes resultados sugerem que a restrição do grupo estável a uma opção mais específica de percepção de melhoria, reduz o erro de medida associado à repetição das medições (EPM) e conseqüentemente o valor de DMD. No entanto, pela concordância de todos os outros estudos, ao considerar o grupo "sem alterações", considera-se que a estimativa da DMD da QBPDS de 19 pontos é mais fidedigna.

Os dois estudos que calcularam os valores de DMD apenas com utentes com DLC (van der Roer *et al.*, 2006; Demoulin *et al.*, 2010), também consideraram os resultados das medições dos participantes da categoria "sem alterações" e obtiveram valores de DMD<sub>95</sub> de 24,6 e 15,8 pontos. Os resultados do estudo mais recente (Demoulin *et al.*, 2010), obtido a partir de uma amostra maior (n=143) sugerem que, para essas características específicas (utentes com DLC referidos para o tratamento multidisciplinar), parece ser mais apropriado um valor de diferença mínima detetável menor do que os anteriormente publicados (DMD<sub>95</sub>=15,8).

### **2.9.2. Estimativa da DMCI da QBPDS**

A DMCI é estimada pela aplicação de métodos baseados na âncora, através de resultados particulares desses métodos.

Nos estudos realizados ao nível da QBPDS, a estimativas de DMCI foram obtidas através do método “Diferença média de pontuação” (van der Roer *et al.*, 2006) e do método curva ROC (Fritz & Irrgang, 2001; van der Roer *et al.*, 2006; Demoulin *et al.*, 2010). As estimativas de DMCI da QBPDS, quando aplicada a utentes com DL (Tabela 7) variam entre 5 (Demoulin *et al.*, 2010) e 20,4 (Fritz & Irrgang, 2001).

**Tabela 7** – Estimativa da DMCI da QBPDS através de Métodos baseados na âncora.

	Amostra/ Intervalo	Estimativa da DMCI	
		Diferença média de pontuação	Curva ROC: ponto ótimo de corte
Fritz & Irrgang (2001)	n =67 DLA; 4 semanas de Fisioterapia		Ponto ótimo de corte= 15; Sensibilidade = 82% (IC 95%: 70% -93%); Especificidade = 83% (IC 95%: 67% -98%);
van der Roer <i>et al.</i> (2006)	n = 442 (n <sub>DLA</sub> =304; n <sub>DLC</sub> =138);	Pela diferença média dos participantes que se consideraram “2) muito melhor”: <b>Amostra DLA</b> = 20,4 (±16,1);	<b>Amostra DLA (n=304)</b> Ponto ótimo de corte = 17.5; Sensibilidade = 0,64; Especificidade = 0.79.
	12 semanas	<b>Amostra DLC</b> = 12,7 (±14,8);	<b>Amostra DLC (n=138)</b> Ponto ótimo de corte = 8.5; sensibilidade = 0.65 e especificidade = 0.76
Demoulin <i>et al.</i> (2010)	n = 212 DLC; 10 semanas de tratamento		Ponto ótimo de corte = 5; Sensibilidade = 0.783; Especificidade = 0.769; ou Ponto ótimo de corte = 18,1%; Sensibilidade = 0.717; Especificidade = 0,853

O método, “Diferença média de pontuação”, originalmente proposto por Jaeschke *et al.* (1989), estima a DMCI pela diferença média da pontuação (entre dois momentos de avaliação - T0 e T1) do grupo dos utentes que reportam “ligeiramente

melhores” na escala de percepção global de mudança, menos a diferença média na pontuação dos questionários no grupo dos utentes que reportam “sem alterações”.

No caso de estudos com amostras de utentes com DLC, o único estudo que calculou a DMCI segundo o método “Diferença média de pontuação” (van der Roer *et al.*, 2006) utilizou uma variante do método. Os autores estimaram a DMCI da QBPDS apenas pela diferença média das pontuações dos participantes que se consideraram “2 - muito melhor”. Ou seja, como a esse valor, não subtraíram a diferença média na pontuação do grupo dos utentes que reportaram “sem alterações”, depreende-se que o valor de DMCI obtido seja superior ao esperado pelo método proposto por Jaeschke *et al.* (1989). A DMCI estimada para a amostra de utentes com DLA foi de 20,4 ( $\pm 16,1$ ) e para a amostra de DLC, estimaram uma DMCI de 12,7 ( $\pm 14,8$ ).

No método curva ROC, a área abaixo da curva (AAC) incide no poder de resposta, mas o ponto ótimo de corte, complementado com os níveis de sensibilidade e especificidade associados, relaciona-se com a interpretabilidade. O ponto ótimo de corte é o ponto mais próximo do canto superior esquerdo do gráfico, (Demoulin *et al.*, 2010). Este ponto é estimado pela sensibilidade e especificidade, que igualmente equilibradas e em conjunto, minimizam o erro geral de classificação (Demoulin *et al.*, 2010). O ponto ótimo de corte é considerado uma expressão da DMCI, de modo que constitui uma estratégia para calcular a DMCI (Lauridsen *et al.*, 2006).

Os estudos realizados com utentes com DLA, estimaram os valores de DMCI de: 15 (Fritz & Irrgang, 2001) e 17,5 (van der Roer *et al.*, 2006), associados a níveis de sensibilidade e especificidade, moderados a elevados.

As estimativas de DMCI, resultantes de medições de utentes com DLC, revelaram valores de DMCI inferiores, em relação aos valores encontrados para utentes com condição aguda. Nas estimativas efetuadas, determinaram-se valores de DMCI de: 8,5 (van der Roer *et al.*, 2006) e 5 pontos (Demoulin *et al.*, 2010), acompanhados dos respetivos valores de sensibilidade e especificidade, também moderados a elevados.

Pela análise dos dados anteriores, pode-se aferir que há uma tendência transversal aos dois métodos, de que as estimativas de DMCI para amostras de utentes com DLA são superiores às estimativas inferidas de amostra de utentes com DLC. Constata-se também, que os valores de DMCI, calculados pelo método “Diferença

média de pontuação”, segundo as condições descritas por van der Roer *et al.* (2006), são superiores comparativamente aos estimados pela curva ROC.

A literatura sugere que a pontuação inicial dos utentes parece influenciar as propriedades de medição dos questionários de incapacidade, tais como, a interpretabilidade e o poder de resposta (Kovacs *et al.*, 2007). Nos estudos de van der Roer *et al.* (2006) e de Demoulin *et al.* (2010), foi investigada e confirmada a influência das pontuações iniciais na estimativa da DMCI.

Embora não se apresente essa particularidade na Tabela 7, van der Roer *et al.* (2006), procederam a uma análise da sensibilidade da DMCI, ou seja, testaram a sua vulnerabilidade às pontuações iniciais. Estes autores procederam à dicotomização dos utentes num grupo das pontuações baixas e outro das pontuações altas, com base no primeiro e terceiro tercil das pontuações iniciais da QBPDS. Contudo, nos resultados não indicam as pontuações de “corte” (da QBPDS) que determinam cada um desses tercis, limitando assim a interpretação desses resultados. Os autores estimaram que a DMCI, para utentes com DLC e com baixas pontuações iniciais pode variar, em valores absolutos, entre 4,8 e 14,5; E que, no caso de utentes com DLC e com altas pontuações iniciais, a DMCI é identificada mediante diferenças na pontuação total da QBPDS, que podem variar, em valores absolutos, entre 11,5 e 33,1.

No estudo de Demoulin *et al.* (2010), para além da categorização dicotómica da amostra ( $\leq 45$  e  $\geq 46$  ou seja, inferior e superior à mediana), testaram a categorização de três sub-grupos (pela distribuição dos tercis:  $< 40$  e  $> 50$ ). Estes autores observaram que o procedimento de agrupar os utentes (em 2 ou 3 categorias), de acordo a sua pontuação inicial, teve impacto sobre a magnitude da pontuação de corte ideal.

Ambos os estudos encontraram maiores valores de DMCI, para os subgrupos de utentes com altas pontuações iniciais. Concluíram igualmente, que o oposto também se verifica, ou seja, os utentes com baixa pontuação inicial identificam um valor de DMCI inferior ao dos utentes com altas pontuações iniciais, quer em utentes com DLA ou DLC (van der Roer *et al.*, 2006).

Estes resultados sublinham o fato de ser ilusório considerar a DMCI como um único valor e que, entre outros fatores, parece ser importante, considerar a pontuação inicial. Desta forma, Demoulin *et al.* (2010) sugerem que é mais apropriado usar a percentagem de melhoria, uma vez que a percentagem tem em consideração a

pontuação inicial. Assim, em termos percentuais, Demoulin *et al.* (2010) concluíram que uma mudança da pontuação inicial na ordem dos 18,1% era considerada uma DMCI.

Para além dos resultados de DMCI apresentados, houve um estudo desenvolvido por Ostelo *et al.* (2008) sobre a interpretação de mudanças de pontuação de instrumentos frequentemente utilizados em utentes com DL, quer para medir intensidade da dor (escala visual analógica e escala numérica da dor), quer para medir o estado funcional (RMDQ, ODI e QBPDS). Com esse estudo pretendiam um consenso internacional e uma recomendação prática quanto à DMCI desses instrumentos de medida. Através de uma revisão da literatura (até 2008) e do debate de um painel de peritos, propuseram que uma DMCI na QBPDS, correspondia a uma diminuição de 20 pontos ou de 30%, relativamente à pontuação inicial (sem ter em consideração o método estatístico utilizado).

O estudo de Ostelo *et al.* (2008), foi a primeira tentativa de reunir consenso e desenvolver recomendações para a DMCI, de medidas frequentemente utilizadas em utentes com DL, no entanto, as suas conclusões podem ser criticadas por erros conceptuais (por exemplo, incluir na pesquisa, estudos de DMD e DMCI) e por não fazer qualquer ressalva à duração dos sintomas dos utentes com DL. Entretanto, Demoulin *et al.* (2010) testou estes valores recomendados através dos pontos de corte da curva ROC. Obteve valores de sensibilidade e especificidade baixos e concluiu que essas referências revelavam um fraco valor em contexto clínico.

Neste trabalho de investigação, de acordo com o conhecimento científico mais atual, opta-se por complementar a estimativa da DMCI com a análise da sua sensibilidade às pontuações iniciais da QBPDS e ainda com a sua estimativa em valor percentual.

### 3. METODOLOGIA

#### 3.1. Identificação dos Objetivos do estudo

A literatura evidencia que a DLC é um problema atual das sociedades. É uma condição com uma enorme prevalência nos países ocidentais da qual resultam elevados custos, quer a nível da sociedade em geral, quer a nível individual. Os utentes com DCL são frequentemente referidos para os cuidados de Fisioterapia (Airaksinen *et al.*, 2006; Koes *et al.*, 2010), sendo que a diminuição da intensidade da dor e da incapacidade funcional para a realização das atividades da vida diária, constituem os principais objetivos e resultados da intervenção (Davidson & Keating, 2002; Smeets, *et al.*, 2011).

Existem atualmente diversos instrumentos desenvolvidos para medir o nível e a mudança na incapacidade funcional destes utentes, sendo que a QBPDS é um dos mais recomendados na literatura (Rocchi *et al.*, 2005; Davidson & Keating, 2002). Os estudos acerca das propriedades psicométricas, quer da sua versão original, quer das adaptações a outras línguas, têm revelado elevados valores de validade e fiabilidade das suas medições. Contudo, até ao momento, poucos estudos analisaram a propriedade de poder de resposta e interpretabilidade da QBPDS. Esses estudos, além de escassos, são heterogéneos em vários aspetos metodológicos, e dessa forma, pouco conclusivos.

A validação da QBPDS – VP, para a população portuguesa com DLC está em desenvolvimento. Foi concluída a primeira fase de adaptação cultural do instrumento e avaliação da validade facial e de conteúdo. Para prosseguir com a utilização do instrumento é fundamental que se teste as propriedades psicométricas do questionário adaptado, na população alvo, neste caso na população portuguesa com DLC. Este estudo vem de encontro a essa necessidade, com a finalidade de estimar a fiabilidade e o poder de resposta da QBPDS – VP e investigar a interpretabilidade da QBPDS – VP, quando aplicada a utentes com DLC.

A avaliação da fiabilidade da VP da QBPDS proposta, inclui a avaliação da consistência interna das respostas ( $\alpha$  de Cronbach) e da estabilidade temporal (teste-reteste). Espera-se que essas estimativas, à semelhança das realizadas em outras versões

linguísticas, concluaam altos níveis de fiabilidade, que possam assegurar a precisão e reprodutibilidade das medições da QBPDS – VP.

O estudo do poder de resposta e da interpretabilidade de QBPDS- VP parece também particularmente relevante e necessário, porque no contexto atual, não se conhece essa capacidade da QBPDS-VP, nem a DMD e/ou DMCI das medições da QBPDS – VP, de modo que é impossível fazer um julgamento clínico das mudanças de pontuações observadas.

Pretende-se ainda, comparar os valores de DMCI observados com os valores recentemente propostos (em estudos realizados noutros países) e avaliar, de forma crítica, potenciais diferenças entre os valores de diferença mínima importante para QBPDS.

Em última instância, este estudo é um contributo para a validação da escala de QBPDS – VP, que visa também, aumentar o potencial de aplicação desse instrumento, através da estimativa da DMD e DMCI em utentes com DLC. A possibilidade de conhecer valores de DMD e DMCI, permite uma leitura clínica dos resultados e melhores inferências quanto à efetividade das intervenções de Fisioterapia.

### **3.2. Quebec Back Pain Disability Scale - VP**

A QBPDS é uma escala de autopreenchimento que pode ser preenchida em menos de 5 minutos (Kopec *et al.*, 1995). É constituída por 20 itens que representam atividades diárias elementares em que os utentes com DL podem percecionar dificuldade na sua execução. Esses itens podem ser agrupados em seis domínios de atividade considerados relevantes para os utentes com DL: Dormir/descansar; sentar/levantar; andar; mover-se; inclinar/agachar e levantar objetos grandes e/ou pesados (Kopec *et al.*, 1996).

Cada um dos 20 itens é graduado através de uma escala de *Likert* de 6 pontos, que varia entre 0 (sem dificuldade nenhuma) e 5 (incapacidade total na realização da atividade).

Os utentes são convidados a responder à QBPDS de acordo com a dificuldade que sentem no presente dia ("hoje") para realizar cada uma das 20 atividades. A

pontuação final é calculada pela soma da pontuação de cada um dos itens, podendo variar entre 0 e 100, onde 0 significa ausência de incapacidade e 100 corresponde à incapacidade total para realizar determinadas atividades da vida diária (Kopeck *et al.*, 1995).

### 3.3. Avaliação das Propriedades Psicométricas

Neste estudo procedeu-se à avaliação das seguintes propriedades psicométricas da escala QBPDS – VP:

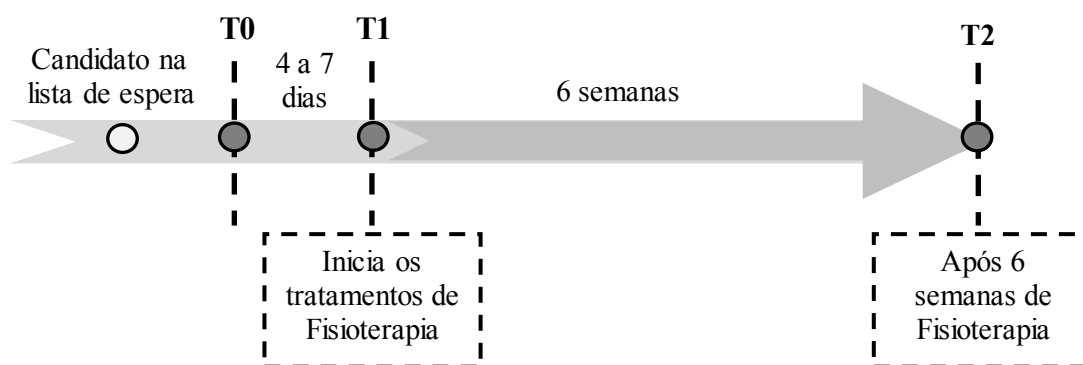
- a. Avaliou-se a unidimensionalidade através de uma Análise Fatorial Exploratória;
- b. Estimou-se a consistência interna, segundo o  $\alpha$  de Cronbach. Adicionalmente fez-se a análise da correlação item a item e item –total da QBPDS;
- c. Procedeu-se à análise da fiabilidade teste-reteste utilizando o Coeficiente de Correlação Intraclasse do tipo 2,1 [CCI (2,1)];
- d. Estimou-se a DMD com recurso ao EPM;
- e. Determinou-se a presença (ou não) de efeitos de “chão” e/ou de “teto” da escala;
- f. Avaliou-se o poder de resposta da escala por método baseado na âncora: a AAC da curva ROC;
- g. Estimou-se a DMCI através de métodos baseados na âncora. Como critério de avaliação externo, utilizou-se a avaliação do efeito efetuada pelos participantes, através da escala de percepção global de mudança – PGIC. Recorreu-se ao método “Diferença Média de Pontuação” (método de Jaeschke *et al.*, 1989) e método Curva ROC.

### 3.4. Tipo de estudo

Considerando os objetivos de investigação, realizou-se um estudo metodológico, multicentro, baseado num *coorte* prospetivo de indivíduos com DLC (ver definição na página 7).

O estudo é metodológico e observacional, porque o estudo das propriedades psicométricas, limita-se a observar as variáveis, não tendo nenhum controlo ou intervenção (propositada) sobre elas. O estudo é ainda multicentro, porque advém de observações obtidas em vários serviços de Medicina Física e de Reabilitação/Fisioterapia de Portugal.

Optou-se por um estudo de *coorte*, pela necessidade de ser um grupo com a característica de interesse, ou seja, com DLC. Para fins do estudo proposto, inclusive desenho teste-reteste e observação de diferenças das pontuações após um intervalo de tempo, decidiu-se por um estudo prospetivo, em que a amostra é seguida desde o presente (T0) e acompanhada por um período de tempo predefinido (7 semanas), onde são registados os eventos específicos de interesse (em T1 e T2). A Figura 3 esquematiza o desenho de estudo.



**Figura 3** - Esquema ilustrativo do desenho de estudo

Este estudo, apesar de apresentar momentos de avaliação, antes e após a exposição de intervenção em Fisioterapia, não teve por objetivo avaliar a efetividade das intervenções de Fisioterapia. No decurso deste estudo não existiu qualquer interferência no plano de intervenção de Fisioterapia estabelecido nos vários contextos

clínicos. Deste modo, as intervenções foram diversificadas e o seu efeito foi “diluído”.

### 3.5. Constituição da Amostra

Para este estudo selecionou-se uma amostra não probabilística, por conveniência, de utentes com DLC que integravam as listas de espera para iniciar intervenção em Fisioterapia, face à acessibilidade dos fisioterapeutas colaboradores na recolha de dados.

Eram elegíveis os indivíduos que cumprissem os seguintes critérios de inclusão:

- Dor localizada na região lombar e/ou associada a sintomatologia dos membros inferiores (Kendall *et al.*, 1997), com duração de pelo menos 12 semanas, sem causa específica, ou presença de recorrências de dor lombar por um período de seis meses, sem causa física específica, (Airaksinen *et al.*, 2006; Bekkering *et al.*, 2003; Krismer & van Tulder, 2007), nomeadamente doença neoplásica, infecciosa e/ou inflamatória, osteoporose, fratura, osteoporose, deformidade estrutural, síndrome da cauda equina e radicular (Waddell, 1987); Bekkering *et al.*, 2003; Krismer & van Tulder, 2007) **ou**
- Presença de dor em menos de metade dos dias num período de 12 meses, ocorrendo múltiplos episódios no decorrer desse período” (VonKorff, 1994). De acordo com VonKorff (1994), estes episódios de agudização são definidos como “períodos (normalmente uma semana ou menos) em que os sintomas da condição de dor lombar se encontram mais exacerbados do que o normal para o utente” **ou**
- Múltiplas situações de agudização dos sintomas, correspondendo assim a vários episódios de recorrência ao longo da sua condição crónica (Stanton, Latimer, Maher & Hancock, 2010);
- Idade compreendida entre os 18 e os 65 anos. A população idosa não é incluída pelo facto das alterações funcionais e biológicas associadas ao processo de envelhecimento poderem estar de alguma forma associadas à dor lombar dos participantes;

- Ausência de sintomas de compressão radicular<sup>7</sup>, cauda equina, doença sistémica, inflamatória e/ou infecciosa, dor de origem visceral/ maligna ou fratura/risco de fratura associado a osteoporose (Smeets *et al.*, 2006);
- Ausência de realização de cirurgia lombar nos 6 meses prévios e de realização de Fisioterapia ou outro tratamento conservador por sintomas lombares nos 3 meses prévios, com exceção da medicação para a dor;
- Mulheres que não se encontrem numa situação de gravidez. A dor lombar durante a gravidez é comum, tendo sido demonstrado que as mulheres grávidas experienciam algum tipo de DL durante o período de gestação, no entanto, os fatores etiológicos apontados são divergentes dos fatores associados à DCL de origem não específica (Garshasbi & Faghîh Zadeh, 2005).
- Saber ler e escrever. A condição de analfabetismo impossibilita o preenchimento dos instrumentos de avaliação da condição dos participantes.

### 3.6. Processo de recrutamento da amostra

Os participantes foram recrutados a partir das listas de espera de 16 serviços de Medicina Física e de Reabilitação/Fisioterapia de Portugal, no período entre novembro de 2011 e maio de 2012. O recrutamento foi efetuado nos respetivos serviços de Fisioterapia, pelos fisioterapeutas colaboradores na investigação, segundo um protocolo para recrutamento dos participantes no estudo (Apêndice B). Esse protocolo estava dividido em três passos consecutivos. Primeiro foram identificados os potenciais participantes, através da verificação do processo clínico dos utentes enviados para a Fisioterapia (que ainda estavam na lista de espera) e do diagnóstico médico de dor lombar - Ex: Dor Lombar, Tensão Lombar, Lombalgia, Lumbago ou Raquialgia Lombar (de acordo com a *International Classification of Diseases*, adotado pela WHO, 2010).

---

<sup>7</sup> Note-se que existe uma diferenciação importante entre dor de origem específica (ex síndrome de compressão radicular) e dor de origem não específica. Em ambos os casos pode existir dor referida para o membro inferior mas na dor lombar de origem não específica não existe *deficit* neurológico (alterações nos reflexos, sensibilidade ou força). É considerada a presença de sintomas de compressão radicular (origem específica) quando se verifica a presença de pelo menos dois testes neurológicos positivos: reflexos, dermatomas (sensibilidade) ou miotomas (força) (Ferreira *et al.*, 2011).

Uma vez confirmada a natureza não específica da DLC, verificaram-se os critérios de inclusão e exclusão para o presente estudo, através de uma *check-list* com 6 itens de inclusão. Nalguns contextos, essa pôde ser aferida no primeiro contacto com o utente (pela via telefónica) e noutras situações, aguardou-se a primeira sessão de Fisioterapia.

Após a verificação dos critérios, todos os utentes elegíveis foram convidados a participar no estudo.

### **3.7. Caracterização da amostra**

Os participantes foram caracterizados do ponto de vista sócio-demográfico e clínico, através do preenchimento do Questionário de Caracterização Sócio-demográfica e Clínica, que foi incorporado no caderno de instrumentos (Apêndice C). Foram recolhidos dados relativos ao género, idade, estado civil, grau de escolaridade e situação profissional, uma vez que são aspetos sócio-demográficos que parecem apresentar particularidades no que diz respeito aos utentes com DLC. Ainda no âmbito sócio-demográfico, os participantes foram caracterizados quanto ao peso e à altura, com o intuito de ser calculado o Índice de Massa Corporal, que é um aspeto sócio-demográfico associado ao desuso, que parece estar presente nos utentes com DLC.

Para caracterizar a condição clínica desses participantes foram recolhidos dados clínicos como a duração da dor, a localização da dor, a utilização de medicação para a dor, faltas ao trabalho por dor e baixas remuneradas, no sentido de se compreender melhor a história de dor dos participantes e o respetivo impacto na vida profissional, dado que estes são aspetos também muito associados aos utentes com DLC. Foi ainda recolhida a perceção de intensidade da dor, nos últimos 7 dias, que foi identificada segundo uma Escala Visual Análoga (EVA).

### **3.8. Aspetos éticos**

Este estudo de investigação, sendo de natureza clínica, considerou e assegurou aspetos éticos abordados na Declaração de Helsínquia (Aguiar, 2007). Neste sentido, o

protocolo desde estudo foi submetido à avaliação da comissão de ética de uma das instituições onde foi desenvolvido – SESARAM, E.P.E.<sup>8</sup>

Após ter sido apresentado o pedido de autorização às instituições envolvidas e facultadas informações relativas ao tipo de estudo, anonimato e confidencialidade dos dados (Apêndice D), o estudo recebeu parecer positivo de todas as instituições, nomeadamente da comissão de ética e administração do SESARAM, E.P.E. (Anexo I) e da provedoria da SCMM<sup>9</sup> (Anexo II).

No decorrer do estudo, foram garantidos aspetos éticos mediante a elaboração e aplicação de uma carta explicativa do estudo, que incluía o consentimento informado (Apêndice E). Assim, todos os utentes que foram convidados a participar no estudo, tiveram obrigatoriamente que ler ou ser informados do conteúdo da carta explicativa. Desta forma, tomaram conhecimento dos objetivos deste estudo e do contexto em que foi efetuado; foram clarificados da voluntariedade da sua participação e assegurados da garantia de que poderiam interromper a sua participação em qualquer altura, sem que esse facto implicasse qualquer constrangimento ou dano relativamente ao seu tratamento; foram igualmente assegurados do anonimato da sua participação e de todos os dados recolhidos durante a execução deste estudo; e foram esclarecidos dos procedimentos de recolha de dados. Os indivíduos que aceitaram participar assinaram o termo de consentimento informado.

A confidencialidade foi outro dos aspetos considerados no decorrer do estudo. Para salvaguardá-la foi criado um sistema de codificação dos dados. Esse código era constituído por 9 algarismos. Os 3 primeiros identificavam o investigador responsável pela recolha dos dados daquele participante; os 3 algarismos do meio representavam o local de recolha dos dados e os 3 últimos indicavam o número do participante de cada local de recolha (Por exemplo: 001002001). O número de codificação atribuído a cada participante foi registado no envelope e nos respetivos questionários do participante.

---

<sup>8</sup> SESARAM, E.P.E. (Serviço de Saúde da Região Autónoma da Madeira, E.P.E.)

<sup>9</sup> SCMM (Santa Casa de Misericórdia de Machico)

### **3.6. Instrumentos de medida**

Decorrente dos objetivos deste estudo e da literatura consultada, decidiu-se que o critério externo (âncora) utilizado seria a percepção global de mudança reportada pelo utente. Assim, além da QBPDS – VP, utilizou-se o seguinte instrumento de medida:

#### **3.6.1. *Patient's Global Impression of Change (PGIC) Scale - VP***

A PGIC é uma medida unidimensional na qual os indivíduos classificam a sua melhoria associada à intervenção, numa escala de 7 itens que varia entre “1= sem alterações” e “7= Muito melhor” (Hurst & Bolton, 2004). Esta escala tem sido utilizada em diversos estudos relativos à dor crónica (Dworkin *et al.* 2003 e Wernicke *et al.* 2006, citados por Domingues & Cruz, 2011), disponibilizando informação facilmente interpretável acerca da importância clínica das mudanças do estado de saúde percebidas pelos indivíduos quando submetidos a determinados tipos de intervenção. A PGIC, tem sido igualmente utilizada para determinar diferenças mínimas clinicamente importantes de instrumentos de avaliação da dor, função física e qualidade de vida (Cella, Hahn, & Dineen, 2002; Farrar *et al.*, 2001).

A sua adaptação cultural para português de Portugal e avaliação da validade de constructo foi realizada por Domingues & Cruz (2011). Concluíram que a PGIC – VP é um instrumento compreensível, adequado, de fácil e rápida utilização, capaz de medir a percepção de mudança do estado de saúde e satisfação com o tratamento de indivíduos com dor crónica de natureza músculo-esquelética, correlacionando-se com instrumentos que evidenciam alterações noutros domínios, nomeadamente na percepção de intensidade da dor.

Para a interpretação dos resultados da escala PGIC, que traduzem a percepção global de mudança experienciada pelos utentes, Hurst & Bolton (2004) referem uma possível divisão das pontuações da PGIC, em duas categorias: “sem alterações ou alterações não importantes” (pontuações 1,2,3,4 e 5) e “muito melhores” (pontuações 6 e 7). Contudo, os mesmos autores, referem que esta divisão é arbitrária e pode variar de amostra para amostra. Neste sentido, foram analisados os estudos anteriores nesta área

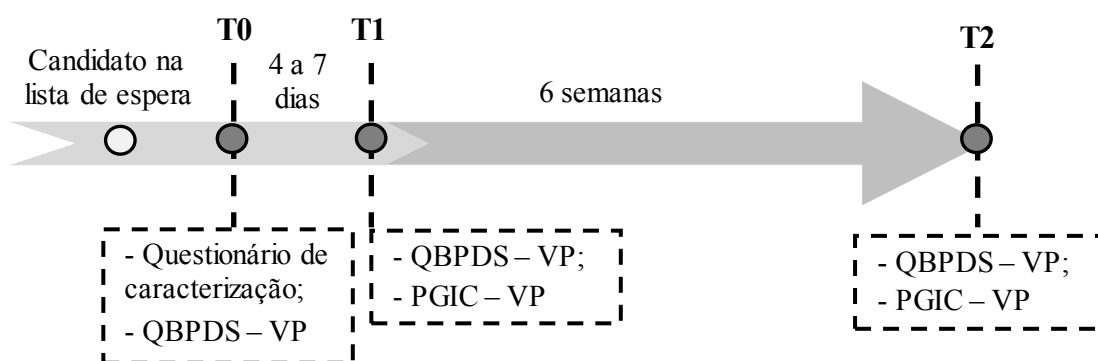
de investigação, que utilizaram uma questão de transição global semelhante (de 7 e 15 itens) com o mesmo propósito deste estudo (Tabela 6) e foi pesquisado o significado dos itens considerados em cada categoria. Verificou-se que os itens que expressam “ligeiramente melhor/com algumas melhorias”, foram incluídos na categoria “sem alterações”, e que os itens que se referem a “melhor” e “muito melhor”, correspondem a uma categoria de “melhoria clínica”. Nesta escala, o item 5 “moderadamente melhor” completa-se com a expressão “...mudança ligeira **mas** significativa”, de modo que a sua semântica traduz uma mudança clínica significativa para o utente. Desta forma, o método de dicotomização da amostra optado no presente estudo, define-se em duas categorias: “sem alterações” (pontuações 1, 2, 3 e 4 na PGIC) e “melhoria clínica” (pontuações 5, 6, 7 na PGIC).

### **3.7. Procedimentos de recolha de dados**

Uma vez recebidas as autorizações dos locais de recolha, pediu-se a colaboração aos fisioterapeutas dessas instituições (Apêndice F) e agendaram-se reuniões com os fisioterapeutas que aceitaram participar na recolha de dados. Nessas reuniões, contextualizou-se o trabalho de investigação e apresentou-se o objetivo do estudo; Descreveram-se os procedimentos envolvidos na identificação dos utentes e na entrega/recolha dos questionários; E disponibilizou-se ajuda para qualquer esclarecimento e treino específico na utilização dos instrumentos de medida.

A recolha de dados sucedeu-se de acordo com o desenho de estudo, representado na Figura 4, que identifica 3 momentos de avaliação dos participantes (T0, T1 e T2). O primeiro momento (T0), ocorreu preferencialmente com uma antecipação de 4 a 7 dias da primeira sessão de Fisioterapia, ou, quando assim não foi possível, ocorreu na primeira sessão de Fisioterapia. O segundo momento (T1), foi recolhido precisamente 4 a 7 dias após o primeiro momento de avaliação. Por último, o terceiro momento (T2) aconteceu após 6 semanas de intervenção de Fisioterapia.

Os questionários foram organizados num caderno de instrumentos (Apêndice C). No primeiro momento (T0), os participantes no estudo preencheram o questionário de caracterização sócio-demográfica (inclui a EVA) a QBPDS - VP. No segundo e terceiro momento (T2 e T3) preencheram a QBPDS - VP e a PGIC - VP.



**Figura 4** - Esquema ilustrativo do protocolo de avaliação

A recolha de dados da QBPDS em T0 e T1, com um intervalo entre os dois momentos, no máximo de 7 dias, foi propositada para avaliar a fiabilidade teste-reteste (do tipo estabilidade temporal). O intervalo entre o teste e o reteste foi de 7 dias, porque considera-se que este é um intervalo suficientemente curto, que permite assumir a variável que está a ser medida não sofreu alterações e, simultaneamente, porque é o tempo mínimo que garante que o utente não se recorda das opções que assinalou no teste (Schoppink *et al.*, 1996). O resultado simultâneo da PGIC em T1, permitiu confirmar, *a posteriori*, os participantes cuja condição clínica não se alterou e que por essa razão, foram incluídos no estudo da fiabilidade teste-reteste. Foi assumida a dicotomização da amostra, segundo os critérios supramencionados, que constituíram dois grupos em T1: grupo “sem alterações na PGIC” (pontuações na PGIC de 1,2,3 e 4) e grupo “com alterações na PGIC” (pontuações na PGIC de 5,6 e7). Desta forma, para a estimativa fiabilidade teste-reteste apenas foram incluídos os participantes do grupo “sem alterações na PGIC”.

O terceiro momento de recolha da QBPDS, ocorreu após 6 semanas de intervenção, por conformidade com a história natural da DLC. Segundo a história natural da condição, 60% dos utentes, revelam diferenças significativas num espaço de 4 a 6 semanas (Roland & Moris, 1983). Efetivamente, 6 semanas, é um intervalo vulgarmente usado na prática para a reavaliação global dos utentes com dor lombar, principalmente se eles não retomaram suas atividades normais. (Fordyce, 1995, Waddell, Feder & McIntosh 1996 citados por Davidson & Keating, 2002). A opção de um intervalo T0 – T2, que oferecesse maior garantia de melhorias clínicas foi

intencional, para se obter a DMCI da QBPDS- VP. Neste sentido, a PGIC preenchida em T2, foi a “âncora” utilizada para a estimativa da DMCI.

Para organizar a recolha de dados, os instrumentos necessários para cada participante foram reunidos num envelope e foram submetidos ao sistema de codificação dos dados, conforme está descrito nos aspetos éticos.

Este sistema de codificação foi muito eficaz para a recolha de dados de vários momentos de avaliação, relativos ao mesmo participante, e também constituiu uma fonte de anonimato. Depois de concluída a recolha de dados, os questionários codificados separaram-se do consentimento informado e do nome do participante. A análise de dados procedeu-se com dados independentes e anónimos.

Para o registo de dados, foram criadas plataformas de registo online, específicas de cada momento de avaliação. Nesses registos, acrescentou-se ao código dos participantes, o código do momento de avaliação (T0, T1 ou T2) a que os dados correspondiam.

### **3.7. Análise dos dados**

A análise de dados efetuada neste estudo, foi decidida em função dos objetivos específicos pretendidos. O nível de significância estabelecido para todos os testes e para o qual os valores se consideraram satisfatórios foi de  $p \leq 0,05$ . O tratamento dos dados foi realizado no software *IBM SPSS Statistics* (versão 20).

#### **3.7.1. Caracterização da amostra**

Para caracterizar a amostra quanto aos aspetos sócio-demográficos e clínicos, e inclusive ao nível da perceção de intensidade da dor, foi utilizada a análise estatística descritiva. Na análise das variáveis nominais e ordinais, recorreu-se à distribuição das frequências; E, na análise das variáveis numéricas, utilizaram-se as medidas de tendência central (médias) e medidas de dispersão (desvio padrão e intervalos máximo e mínimo).

Na caracterização da amostra no momento inicial (T0), analisaram-se apenas as variáveis sócio-demográficas e clínicas mais referidas na literatura em estudos anteriores da fiabilidade da QBPDS. Ao nível do estudo do poder de resposta da QBPDS, a caracterização da amostra que constitui o terceiro momento de avaliação (T2), faz-se mediante a comparação das características no ponto de partida dos grupos “sem alterações” e “melhoria clínica”.

No questionário de caracterização sócio-demográfica, a questão 7 relativa à atividade profissional, era uma questão de resposta aberta, que foi interpretada segundo a Classificação Portuguesa das Profissões de 2010 (Instituto Nacional de Estatística, 2011a), que permitiu a organização das respostas em 10 categorias possíveis<sup>10</sup>. Além destas, foi considerada a categoria “estudante”.

A variável numérica dos resultados da IMC foi ainda transformada numa variável qualitativa ordinal com 3 categorias, de acordo com o sistema de classificação internacional do IMC ( $\text{Kg/m}^2$ ) proposto pela WHO (2000): Peso corporal abaixo do normal ( $\text{IMC} < 18$ ); Peso normal ( $18,5 < \text{IMC} \leq 24,99$ ); Excesso de peso ( $\text{IMC} \geq 25$ ).

### **3.7.2. Avaliação da Unidimensionalidade da QBPDS**

O contributo para a validação do questionário QBPDS – VP visa, em primeiro lugar, verificar que o questionário mede a presumível e única variável latente, que é a incapacidade funcional associada à DLC. Se o questionário assim o fizer, diz-se que é unidimensional. Neste sentido, decidiu-se por realizar uma Análise Fatorial Exploratória (AFE). A AFE analisa as correlações observáveis entre os itens e verifica se essa associação entre itens resulta da partilha de uma característica comum não diretamente observável (isto é, um fator comum latente). Assim, a AFE verifica se esses itens medem uma ou mais variáveis latentes. O objetivo primordial da AFE é atribuir uma

---

<sup>10</sup> As 10 categorias das profissões, foram: 1) “Profissões das Forças Armadas”; 2) “Representantes do poder legislativo e de órgãos executivos, dirigentes, diretores e gestores executivos”; 3) “Especialistas das actividades intelectuais e científicas”; 4) “Técnicos e profissões de nível intermédio”; 5) “Pessoal administrativo”; 6) “Serviços pessoais de protecção e segurança e vendedores”; 7) “Qualificados da agricultura, pesca e floresta”; 8) “Qualificados da indústria, construção e artífices”; 9) “Operadores de instalações e máquinas e trabalhadores de montagem”; 10) “Trabalhadores não qualificados”. Além destas, foi considerada a categoria “estudante”.

pontuação (quantificação) a “construtos” ou fatores que não são diretamente observáveis (Marôco, 2011).

Um dos pressupostos da AFE é que as correlações entre as variáveis originais são elevadas o suficiente para que a AFE tenha utilidade na estimativa de fatores comuns (Marôco, 2011). Para avaliar a qualidade dos dados recorreu-se ao teste de Esfericidade de Barlett e à “medida de adequação da amostragem de Kaiser-Meyer-Olkin” (KMO).

O Teste de Esfericidade de Bartlett, testa a hipótese nula de que a matriz das correlações é uma matriz diagonal (isto é, uma matriz onde as correlações entre as variáveis são iguais a zero). Quando o valor da estatística do Qui-quadrado é significativo, as correlações entre as variáveis são adequadas para fazer uma análise fatorial (Hill & Hill, 2005).

O KMO é uma medida da homogeneidade das variáveis, que compara as correlações simples com as correlações parciais observadas entre as variáveis (Marôco, 2011). A partir do valor de KMO obtemos um nível de recomendação para a AF. Se o valor encontrado for inferior a 0,5, é inaceitável realizar uma análise fatorial, se estiver entre 0,8 e 0,9, tem um nível de recomendação bom para avançar com AFE, e se o valor obtido estiver entre 0,9 e 1 é considerado excelente (Marôco, 2011).

Com recurso à programação informática, existem vários métodos iterativos para obter uma solução para o problema da estimativa das comunalidades. Para efeitos deste estudo, optou-se pelo “Método das componentes principais”, porque é o método mais comumente utilizado (Marôco, 2011). Depois de obtida a solução fatorial inicial do questionário, isto é, depois de estimados os fatores que explicam o comportamento correlacional das variáveis, o passo seguinte é decidir qual o número mínimo de fatores que se deve reter de forma a representar e resumir, apropriadamente, a informação presente nas variáveis originais. Para a decisão do número apropriado de fatores a reter, estabelecem-se os seguintes critérios:

1. Critério de Kaiser (ou regra do “eigenvalue<sup>11</sup> superior a 1”)
2. Critério do *Scree plot*;

---

<sup>11</sup> *Eigenvalue* ou “valor próprio” (Marôco, 2011)

3. Os itens devem ter um peso fatorial igual ou superior a 0.5 (*factor loadings*  $\geq$  0.50);

4. A percentagem da variância explicada pelos fatores retidos deve ser no mínimo de pelo menos 40% (Streiner & Norman, 2003), ou de pelo menos 50% (Marôco, 2011). Segundo Streiner & Norman (2003), a unidimensionalidade de um instrumento está refletida num fator predominante, quando se obtém um fator predominante com um *eigenvalue* mais elevado que os outros fatores e que explique uma proporção elevada da variância em comparação com os outros fatores. Nesses casos, se os restantes fatores forem relativamente fracos, devem ser eliminados.

### **3.7.2. Avaliação da consistência Interna**

Para a avaliação da consistência interna do questionário QBPDS – VP (em T0) recorreu-se ao coeficiente  $\alpha$  de Cronbach. Complementou-se o estudo com a análise dos itens (correlações item-total e correlações entre itens), que permitiu averiguar se essas correlações eram relativamente fortes e verificadas por todos os itens.

### **3.7.3. Avaliação da Fiabilidade Teste-reteste**

Para a estimativa da fiabilidade teste-reteste (do tipo estabilidade temporal) foi necessário decidir o intervalo de tempo mais apropriado entre as duas medidas.

De acordo com revisão da literatura (Capítulo 2), os intervalos de tempo dos estudos anteriores, que estimaram a fiabilidade teste – reteste da QBPDS, foram muito diversos. Apesar da diversidade, distinguem-se intervalos mais curtos e intervalos maiores, associados a duas finalidades distintas. Por um lado, para períodos curtos, de 1 a 7 dias, durante os quais os utentes não recebiam qualquer intervenção, assumiu-se que a condição se mantinha estável. Estes intervalos foram utilizados para avaliar a fiabilidade do tipo estabilidade temporal. Por outro lado, estudos em que os intervalos teste-reteste foram maiores, de 4 a 12 semanas, e durante os quais os utentes sofreram alguma intervenção. Nestes, foi necessário garantir que os utentes se mantinham

estáveis relativamente à condição a ser medida. Foi constituído um grupo clinicamente estável, para o qual foi estimado o Coeficiente de Correlação Intraclasse (CCI). As estimativas destes últimos, serviram de coeficiente de correlação das pontuações da QBPDS, entre a pré intervenção e a pós intervenção, para o grupo estável. Esse coeficiente possibilitou o cálculo do EPM e subsequente DMD.

Da mesma forma, no presente trabalho, opta-se por avaliar a fiabilidade teste-reteste em dois momentos, com intervalos diferentes e duas finalidades distintas.

A fiabilidade do tipo estabilidade temporal da QBPDS, pretendida neste subcapítulo, foi estimada entre o momento inicial (T0) e 4 a 7 dias depois (T1). A outra estimativa da fiabilidade (T0-T2) é necessária para o cálculo do EPM e DMD.

### **3.7.3.1. Fiabilidade Teste-Reste pré intervenção da Fisioterapia (Intervalo de 4 a 7 dias);**

A fiabilidade do tipo estabilidade temporal da QBPDS, foi estimada entre o momento inicial (T0) e 4 a 7 dias depois (T1), na medida que é o intervalo teste-reteste que apresenta argumentos mais válidos (Schoppink *et al.*, 1996).

Para aumentar a garantia de que os utentes mantinham a sua situação funcional estável, que é um aspeto relevante para a avaliação da estabilidade, foi utilizada a resposta simultânea da PGIC, em T1. Apenas foram incluídos os utentes que perceberam que sua condição clínica estava estável (grupo “sem alterações na PGIC” em T1), isto é, os que se auto-classificaram na PGIC como “sem alterações”, “quase na mesma”, “ligeiramente melhores” ou “moderadamente melhores, mas cuja mudança não representou qualquer diferença real”, respetivamente as pontuações 1, 2 3 ou 4 na escala de PGIC.

Assim, a fiabilidade teste-reteste foi investigada, através dos resultados das pontuações totais da QBPDS das amostras emparelhadas, obtidos em T0-T1, no grupo de participantes identificados a *posteriori* (*post-hoc*) como não tendo alterado a sua condição no espaço de 7 dias (grupo “sem alterações” em T1).

Para avaliar a fiabilidade teste-reteste estimou-se o coeficiente de correlação intraclasse do tipo 2,1 [CCI (2,1)]. Decidiu-se pela aplicação do modelo 2, porque

aspira-se que quer os observadores, quer os sujeitos sejam efeitos aleatórios, com duas particularidades: no caso da QBPDS não há um observador externo; e a randomização dos sujeitos é apenas teórica. Apesar disso, acredita-se que os indivíduos escolhidos representam a população de interesse e que os resultados deste estudo podem ser generalizados para outros sujeitos. Para o modelo 2 é considerado um teste ANOVA a dois fatores (*Two-way random*).

O segundo número do tipo de CCI representa a forma utilizada, que neste caso é igual a 1 porque se tratam de medições singulares (*single measures*).

#### **3.7.4. Estimativa do EPM e da DMD**

Estas estimativas consideraram as amostras emparelhadas T0-T2. A amostra T2 foi constituída por todos os indivíduos que completaram 6 semanas de tratamento, e que foram avaliados no momento T2. Para efeito destas estimativas, foi necessário previamente:

a) Analisar a amostra T2 e os resultados ao nível da incapacidade funcional (T0-T2), por análise estatística descritiva;

b) Dividir a amostra nos dois grupos, identificados a *posteriori* (*post-hoc*), de acordo com a PGIC (“sem alterações” e “melhoria clínica”);

c) Garantir a estabilidade da condição funcional (pontuações da QBPDS em T0 e T2 estáveis), no grupo de indivíduos que se percebeu estável no espaço de 7 semanas (pela PGIC em T2); Optou-se pela aplicação de um teste paramétrico de comparação de médias para amostras emparelhadas. Não sendo verificada a normalidade dos dados, foi aplicada a alternativa não paramétrica: Teste Mann Whitney ( $p \leq 0,05$ ).

Pretendia-se testar a hipótese de que no grupo “sem alterações”, as variáveis (Incapacidade média em T0” e Incapacidade média em T2”), não diferiam de forma estatisticamente significativa entre os dois momentos de avaliação, representando assim que o grupo “sem alterações”, se mantinha estável em relação ao momento inicial (T0). Pelo contrário, perspetivava-se que o grupo “melhoria clínica” não se mantivesse

estável, ou seja, que tivessem ocorrido alterações significativas na pontuação de QBPDS de T0 para T2.

d) Comprovar a equivalência dos grupos “sem alterações” e “melhoria clínica”, relativamente às características dos indivíduos, no ponto de partida. No sentido, de assegurar que as diferenças observadas ao nível da percepção de melhoria, dos grupos, não se devem a diferenças sócio-demográficas ou clínicas dos elementos que a constituem.

As variáveis quantitativas (pontuações da QBPDS em T0; e a intensidade da dor em T0) foram comparadas através do teste T-student para amostras não emparelhadas, ou através da alternativa não paramétrica, o teste Mann Whitney; E as restantes variáveis sócio-demográficas e clínicas (género, IMC, estado civil, etc.), que são variáveis qualitativas nominais, foram aferidas através do teste Qui-Quadrado para um valor  $p > 0,05$ . As variáveis são testadas para verificar se diferenças nas variáveis no ponto de partida estariam associadas a diferenças nos grupos (“melhoria clínica”/ “sem alterações”).

e) Uma vez verificada a estabilidade da condição clínica dos participantes do grupo “sem alterações” no intervalo de 7 semanas (entre T0 e T2), conforme a alínea c) e d), estimou-se a fiabilidade teste-reteste das pontuações da QBPDS (intervalo de 7 semanas) para esse grupo.

Depois de concluídos estes procedimentos, calculou-se EPM, EPM repetição e DMD, de acordo com as fórmulas propostas pela literatura mais recente:

a)  $EPM = DP_{\text{médio}} \times \sqrt{1 - r}$  (Davidson & Keating, 2002), com as medidas relativas ao grupo estável;

b)  $EPM = DP_{\text{médio}} \times \sqrt{1 - r}$  (Davidson & Keating, 2002)

c)  $DMD = 1,65 \times \sqrt{2} \times EPM$  (Davidson & Keating, 2002; Fritz & Irrgang, 2001; Hicks & Manal, 2009). Em que 1,65 é o valor z do percentil 90 de uma distribuição normal.

### **3.7.5. Determinação dos efeitos de “chão” e de “teto”**

Com base na DMD, calculou-se a amplitude válida da QBPDS: Limite inferior (DMD) a Limite superior (100 – DMD). Os efeitos de “chão” e de “teto” são considerados presentes, se mais de 15% dos participantes, alcançam as pontuações mais inferiores (efeito de “chão”) ou superiores (efeito de “teto”) possíveis (Terwee *et al.*, 2009).

### **3.7.6. Estimar Poder de Resposta**

Para estimar o poder de resposta da QBPDS foram preferidos os métodos baseados na âncora em relação aos métodos baseados na distribuição, uma vez que estes últimos não têm em conta a “importância” da mudança (Terwee *et al.*, 2009) e este estudo está particularmente interessado na definição de poder de resposta enquanto capacidade de um instrumento detetar mudanças clinicamente relevantes.

O índice do poder de resposta foi calculado com recurso a métodos baseados na âncora, nomeadamente através da correlação das pontuações da QBPDS com a PGIC, e da curva ROC.

#### **3.7.6.1. Método “Coeficiente correlação âncora”**

Este método compara a magnitude dos coeficientes de correlação das mudanças na pontuação da QBPDS com as pontuações da PGIC, para determinar qual o índice de validade da mudança (Stratford, Binkley & Riddle, 1996).

Neste estudo, utilizou-se a medida de associação não paramétrica e alternativa ao coeficiente de correlação de Pearson, ou seja, o coeficiente de correlação de Spearman.

### **3.7.6.2. Método Curva ROC: área abaixo da curva**

A aplicação do método da curva ROC requer que os indivíduos sejam classificados em dois grupos, que representem o “sucesso” e o “insucesso” da condição que está a ser testada. Neste caso a condição a ser testada foi a melhoria da condição clínica, nomeadamente ao nível de incapacidade funcional. Assumiu-se a dicotomização da amostra em “sucesso” (grupo “melhoria clínica” onde PGIC =5,6,7) vs “insucesso” (grupo “sem alterações” onde PGIC =1,2,3,4).

Assim, solicitou-se no SPSS a análise da curva ROC, considerando que a variável da condição (adquirida ou não) é “Categorias da percepção de Melhoria” e a variável a ser testada é a “Diferença nas pontuações de QBPDS T0-T2”.

A AAC reflete a capacidade do teste de discriminar indivíduos que melhoraram (sucesso) dos que se mantêm inalterados (insucesso) (Stratford, Binkley & Riddle, 1996; Streiner & Norman, 2003). O valor de 1 para a AUC representa o rigor perfeito (100%), enquanto o valor 0,5 representa apenas o acaso.

### **3.7.7. Determinação da Diferença Mínima Clinicamente Importante**

#### **3.7.7.1. Diferença Média de Pontuação**

Calculou-se a DMCI através do método originalmente proposto por Jaeschke *et al.* (1989): “subtrair a diferença média da pontuação QBPDS T0-T2 do grupo “sem alterações” à diferença média do sub-grupo de utentes que se identificou com o nível da escala de percepção global de mudança, mais baixo para ser considerado melhoria clínica”. Para efeitos deste trabalho, desenvolvido com a PGIC- VP, considerou-se que esse nível correspondia a PGIC=5.

### 3.7.7.2. Método Curva ROC: ponto ótimo de corte

Pela análise da curva ROC, estimou-se o ponto mais próximo do topo esquerdo da curva ROC (*Receiver Operating Characteristic*), que se designa o ponto ótimo de corte e representa o valor de DMCI.

Procedeu-se à análise da curva ROC, considerando que a variável da condição (adquirida ou não) é “Categorias da percepção de Melhoria” e a variável a ser testada é a “Diferença nas pontuações de QBPDS T0-T2”.

Na variável da condição, a condição de “sucesso” foi atribuída aos utentes que reportaram uma pontuação na PGIC= 5,6,e 7 (grupo “melhoria clínica”) e a condição “insucesso” foi atribuída aos utentes com PGIC= 1,2,3 e 4 (grupo “sem alterações);

Uma problemática associada à estimativa da DMCI, conforme foi discutida na literatura, refere-se à sensibilidade da DMCI às pontuações iniciais da QBPDS (Kovacs *et al.*, 2007; Demoulin *et al.*, 2010). Para ultrapassar esse problema nas estimativas da DMCI da QBPDS, optou-se por duas estratégias:

a) Primeiro decidiu-se averiguar a DMCI em termos percentuais, porque a sua estimativa permitiu colmatar esse problema com uma única referência de DMCI, em percentagem de melhoria;

Procedeu-se a uma nova análise da curva ROC, assumindo a mesma dicotomização de “sucesso” vs “insucesso”, mas considerando que a variável a ser testada era a “Diferença relativa nas pontuações de QBPDS T0-T2”. Dado que os resultados da diferença relativa, que expressam uma diminuição na pontuação da QBPDS de T0 e T2 (ou seja, uma melhoria da incapacidade funcional) são negativos, nesta análise decidiu-se uma direção de teste oposta à anterior (opção da curva ROC): “Resultados de teste inferiores, indicam teste mais positivo”. A interpretação do ponto ótimo de corte deve ser que um valor igual ou inferior a essa percentagem de diferença de pontuação de T0 para T2 (valor negativo), representa uma diferença clinicamente importante.

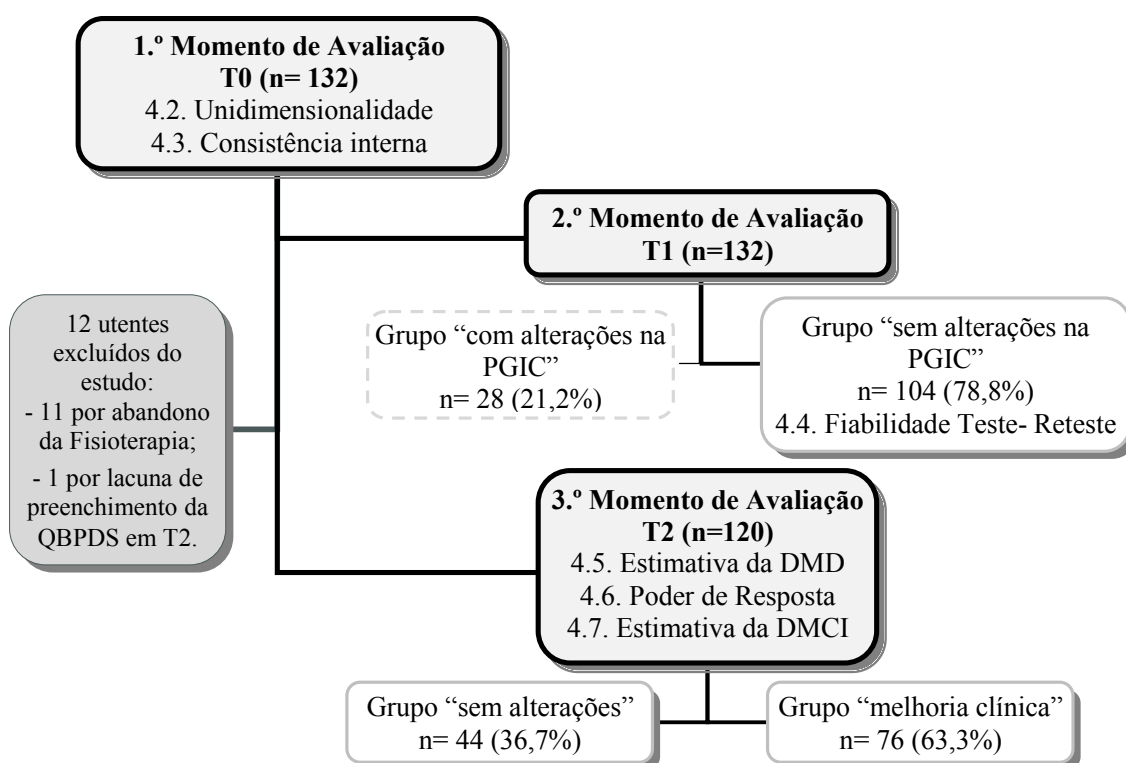
b) Posteriormente, recorreu-se à estimativa da DMCI identificada por utentes com pontuações da QBPDS no momento inicial, baixas e altas. Na literatura são referidos dois critérios de decisão de pontuações baixas e altas: pela mediana da amostra (Demoulin *et al.*, 2010) ou pelo primeiro e terceiro tercil da amostra (van der Roer *et al.*, 2006). Optou-se por dividir toda a amostra em dois subgrupos: “pontuações baixas” (utentes com pontuação da QBPDS em T0 < à mediana) e “pontuações altas” (utentes com pontuação da QBPDS em T0 ≥ à mediana).

Foi solicitada a curva ROC, para ambos os subgrupos, em que a variável a ser testada foi a “Diferença nas pontuações de QBPDS T0-T2”. Foi assumida novamente a mesma dicotomização de “sucesso” vs “insucesso”.

## 4. APRESENTAÇÃO DOS RESULTADOS

A apresentação dos resultados deste estudo encontra-se organizada por subcapítulos que correspondem aos objetivos propostos, em termos de contributo para a validação da QBPDS – VP e do estudo do seu poder de resposta e interpretabilidade. Inicia-se com uma análise descritiva para caracterização da amostra e prossegue-se com a avaliação da unidimensionalidade, consistência interna, fiabilidade teste-reteste, estimativa da DMD, poder de resposta e estimativa da DMCI.

O estudo compreende dados recolhidos nos 3 momentos de avaliação, definidos como, T0, T1 e T2, no estudo de *coorte* prospetivo realizado com utentes com DLC. Na figura 5, observa-se o número de indivíduos que participaram em cada momento de avaliação e o número de indivíduos que constituíram cada um dos grupos: em T1 - grupo “sem alterações na PGIC” e grupo “com alterações na PGIC”; e em T2 - grupo “sem alterações” (cl clinicamente estáveis) e grupo “melhoria clínica”. Através da Figura 5 também são identificadas as amostras utilizadas (T0, T1 e T2) para cada um dos propósitos do estudo (abordados nos subcapítulos que se seguem).



**Figura 5** – Constituição da amostra em cada momento de avaliação (T0,T1,T2).

#### 4.1. Caracterização da amostra T0

O momento inicial de avaliação (T0) integra 132 utentes com DLC. Todos os participantes que foram convidados a participar no estudo, por se encontrarem em lista de espera para receber tratamento de Fisioterapia e reunirem os critérios de inclusão, aceitaram participar no mesmo. Na medida, em que não houve lacunas no preenchimento da QBPDS em T0, a amostra do primeiro momento de avaliação (Amostra T0) é constituída pela totalidade dos utentes (n=132).

Na Tabela 8, são descritas as principais variáveis sócio-demográficas e clínicas, que caracterizam a amostra no ponto de partida do estudo de *coorte* (Apêndice G). Esta amostra (T0) foi utilizada para a avaliação da unidimensionalidade e consistência interna da QBPDS.

Pela observação da Tabela 8, pode-se aferir que a amostra T0, constituída por indivíduos com idades compreendidas entre os 18 (mínimo) e os 65 anos (máximo), contempla toda a amplitude de idades admitida pelos critérios de inclusão; e que a idade média dos indivíduos é de 46,58 ( $\pm 12,67$ ) anos. A amostra caracteriza-se pelo predomínio do género feminino, com 96 participantes desse género (72,7%), sendo os restantes 36 do género masculino (27,3%).

O nível de escolaridade é diverso, permitindo uma amostra abrangente dos vários níveis de escolaridade. Verifica-se uma representação equilibrada dos níveis de escolaridade das extremidades, em que 25% dos participantes têm um baixo nível de escolaridade (ensino primário) e 28% dos participantes possuem estudos ao nível do ensino superior; e constata-se um ligeiro predomínio do nível de escolaridade médio (ensino básico e ensino secundário), que constitui 47% da amostra. Na situação profissional atual, a maioria dos indivíduos estão no ativo (65,9%) e dos restantes, 12% estão reformados, 7,6 % estão desempregados e apenas 3,8% estão incapazes de trabalhar devido ao seu problema.

Do ponto de vista clínico, a amostra é constituída maioritariamente por utentes cuja condição de DLC se prolonga há mais de 24 meses (65,9%). Relativamente à localização da dor, observa-se uma prevalência aproximadamente idêntica de indivíduos que não apresentavam dor referida para o membro inferior (47%) e de indivíduos cuja dor se prolongava para o membro inferior (53%). Do mesmo modo, na variável

medicação, aproximadamente metade dos indivíduos não toma medicação para a sua dor (54,5%), enquanto a outra metade dos indivíduos toma medicação (45,5%).

**Tabela 8** – Caracterização sócio-demográfica e clínica da amostra T0.

<b>Amostra T0 (n= 132)</b>			
<b>Variáveis sócio-demográficas e clínicas</b>	<b>n (% válida)</b>	<b>Média (DP)</b>	<b>Min. – Máx.</b>
<b>Idade (Anos)</b>	132	46,58 (±12,67)	18-65
<b>Género</b>			
Masculino	36 (27,3%)		
Feminino	96 (72,7%)		
<b>Nível de escolaridade</b>			
Ensino Primário	33 (25%)		
Ensino Básico	25 (18,9%)		
Ensino Secundário incompleto	22 (16,7%)		
Ensino Secundário completo	15 (11,4%)		
Ensino Superior incompleto	6 (4,5%)		
Ensino Superior completo	31 (23,6%)		
<b>Situação profissional atual</b>			
A trabalhar a tempo inteiro	80 (60,6%)		
A trabalhar a tempo parcial	7 (5,3%)		
Incapaz de trabalhar	5 (3,8%)		
Desempregado (a)	10 (7,6%)		
Reformado (a)	16 (12,1%)		
<b>Duração da dor</b>			
3-6 meses	18 (13,6%)		
6-12 meses	12 (9,1%)		
12-24 meses	15 (11,4%)		
Mais de 24 meses	87 (65,9%)		
<b>Localização da dor</b>			
Sem dor referida para o membro inferior	62 (47%)		
Com dor referida para o membro inferior	70 (53%)		
<b>Medicação</b>			
Não	70 (54,5%)		
Sim	60 (45,5%)		
<b>Baixa Remunerada (ano anterior)</b>			
Não	102 (77,3%)		
Sim	30 (22,7%)		

A maioria dos participantes que constituem esta amostra, não esteve de baixa remunerada no último ano devido a DLC (77,3%). Apenas 30 indivíduos estiveram de baixa remunerada (22,7%).

## 4.2. Unidimensionalidade da QBPDS

Os procedimentos para a averiguação da unidimensionalidade da QBPDS preveem uma AFE dos resultados da QBPDS em T0 (n=132) (Apêndice H). Essa técnica de análise exploratória exige a verificação prévia de uma correlação suficientemente elevada entre as variáveis. O Teste de Esfericidade de Bartlett e a medida de adequação da amostragem de KMO, apresentados na Tabela 9, constituem uma forma de avaliação dessa qualidade dos dados.

Analisando a Tabela 9, o Teste de Bartlett tem associado um valor  $p = 0,000$ , ou seja, inferior a  $\alpha$ , para um nível de significância de 0,05. Este facto, leva à rejeição da hipótese nula ( $H_0$ : a matriz das correlações é uma matriz diagonal, onde as correlações entre as variáveis são iguais a zero), mostrando assim, que existe correlação entre as variáveis (Hill & Hill, 2005).

Uma medida de KMO no valor de 0,917, representa um nível de recomendação excelente para prosseguir com a AFE (Marôco, 2011).

**Tabela 9** – Resultados do Teste de Esfericidade de Bartlett e KMO

<b>Teste de Esfericidade de Bartlett</b>	Aprox. Qui-Quadrado	1886,085
	Valor p	0,000
<b>KMO</b>		0,917

### 3.2.1. Estrutura Fatorial da QPBDQ

A Análise Fatorial de Componentes Principais (AFCP), determina a solução fatorial inicial da estrutura do questionário com base nas estimativas iniciais dos pesos fatoriais na diagonal principal (Marôco, 2011). A Tabela 10, apresenta os resultados relativos à solução inicial da QBPDS (para explicar as correlações entre os 20 itens do questionário), onde são indicados os “valores próprios” para cada fator e a percentagem de variância explicada. Assim, a AFCP indica uma solução fatorial inicial de quatro fatores comuns, que no seu conjunto explicam 70,1% da variância total no conjunto dos 20 itens analisados.

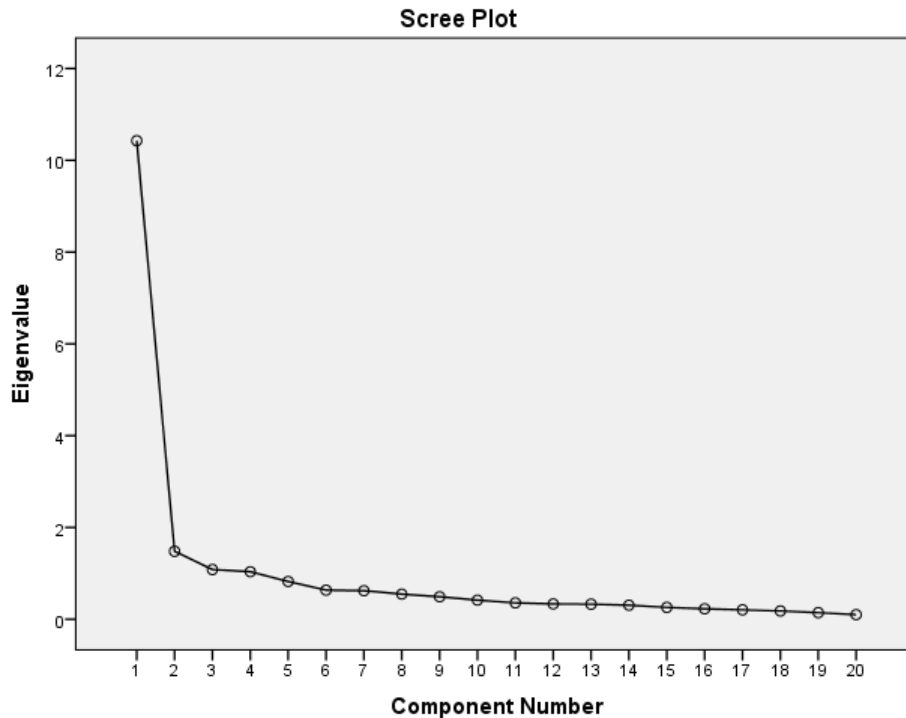
**Tabela 10** – Total de variância explicada pela solução fatorial inicial da QBPDS

Componentes	<i>Eigenvalues</i> iniciais ("valores próprios")			"Extraction Sums of Squared Loadings"		
	Total	% de variância	% acumulada	Total	% de variância	% acumulada
1	10,428	52,138	52,138	10,428	52,138	52,138
2	1,481	7,404	59,542	1,481	7,404	59,542
3	1,083	5,416	64,959	1,083	5,416	64,959
4	1,036	5,180	70,139	1,036	5,180	70,139
5	0,823	4,114	74,253			
6	0,635	3,173	77,425			
7	0,622	3,112	80,538			
8	0,547	2,737	83,275			
9	0,490	2,448	85,723			
10	0,417	2,084	87,807			
11	0,356	1,782	89,589			
12	0,334	1,668	91,257			
13	0,328	1,642	92,899			
14	0,306	1,531	94,429			
15	0,258	1,290	95,720			
16	0,228	1,142	96,862			
17	0,205	1,024	97,885			
18	0,180	0,900	98,785			
19	0,143	0,717	99,503			
20	0,099	0,497	100,000			

A decisão do número de fatores (ou componentes principais) a reter segundo o critério de Kaiser (ou regra do "*eigenvalue* superior a 1"), que pode ser verificado pela Tabela 10 (na coluna "*Eigenvalues* iniciais"), sugere que fiquem retidos 4 fatores comuns. No entanto, os restantes critérios considerados para a decisão do número de fatores a reter, estão em conformidade e indicam apenas um fator comum:

1. Na representação gráfica do *Scree plot* (gráfico Y), que ilustra a importância relativa de cada fator na explicação da variância total das variáveis originais, verifica-se que a curva de inflexão (ou cotovelo) ocorre entre o 1º e o 2º fator.

Segundo o critério do *Scree plot*, deve reter-se o número de fatores até aquele em que se observa a inflexão da curva. Deste modo, o gráfico Y indica que se mantenha apenas 1 fator comum, porque a partir do 2º fator, o ganho da variância total já não é considerável.



**Figura 6 – Scree plot**

2. Em consonância, o critério de que os itens devem ter peso fatorial igual ou superior a 0,5 (i.e., *factor loadings*  $\geq 0,50$ ), apoia a decisão de reter apenas um fator comum. A Tabela 11, apresenta o peso fatorial de cada item para cada fator. Segundo a tabela, verifica-se que os 20 itens do questionário apresentam um peso fatorial superior a 0,5 para o primeiro fator comum, e que não são obtidos pesos fatoriais de tal grandeza para os restantes fatores comuns.

3. Por último, e de acordo com a percentagem da variância explicada pelos fatores retidos (Tabela 10), confirma-se que o primeiro fator é predominante (*eigenvalue* de 10,428, e explica 52,1% da variância), e que os outros fatores são relativamente fracos (explicam menos de 7,5% da variância), pelo que devem ser eliminados.

Os resultados obtidos com a AFPCP são expressivos e nesse sentido dispensam a análise da rotação. Os resultados sugerem uma estrutura relacional dos itens da QBPDS explicada por um fator comum predominante e que reflete a unidimensionalidade do questionário.

**Tabela 11** – Peso fatorial dos itens para cada fator

	<b>Componentes</b>			
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Q1	0,690	0,457	-0,054	0,138
Q2	0,580	0,438	-0,329	0,340
Q3	0,682	0,325	-0,148	0,343
Q4	0,679	0,200	0,155	0,171
Q5	0,616	-0,023	0,429	-0,018
Q6	0,512	-0,036	0,717	0,085
Q7	0,792	-0,084	0,063	0,279
Q8	0,763	-0,071	-0,019	0,278
Q9	0,757	-0,244	0,049	0,246
Q10	0,761	-0,327	0,078	-0,021
Q11	0,766	-0,177	-0,002	-0,055
Q12	0,702	-0,395	-0,113	0,138
Q13	0,685	0,272	0,008	-0,442
Q14	0,748	0,342	0,024	-0,373
Q15	0,715	0,294	0,150	-0,170
Q16	0,783	0,084	0,047	-0,210
Q17	0,773	0,084	-0,201	-0,199
Q18	0,779	-0,278	-0,304	-0,204
Q19	0,817	-0,284	-0,155	-0,074
Q20	0,764	-0,343	-0,140	-0,121

Uma vez verificada a unidimensionalidade da QBPDS – VP, prossegue-se com a avaliação a consistência interna, considerando que todos os itens do questionário medem o mesmo construto.

#### **4.3. Consistência Interna da QBPDS**

A consistência interna foi avaliada de acordo com o coeficiente de  $\alpha$  de Cronbach da QBPDS e pela análise dos itens (Apêndice I).

O coeficiente de  $\alpha$  de Cronbach e a análise dos itens foram calculados, com os resultados da QBPDS da amostra T0 (n=132 utentes com DLC). O valor de  $\alpha$  de Cronbach da QBPDS obtido, de 0,95 (Tabela 12), é considerado elevado e revela uma excelente consistência interna do questionário (Hill & Hill, 2005).

**Tabela 12** – Resultado da estimativa de  $\alpha$  de Cronbach

<b>Número de itens</b>	<b><math>\alpha</math> de Cronbach</b>
20	0,950

Da análise dos itens resultam as Tabelas 13 e 14, que apresentam os valores de correlação item- item e correlação item-total, respetivamente. Os valores de correlação item- item (Tabela 13) são todos positivos, compreendidos entre 0,157 e 0,802, sendo predominantes os valores entre 0,3 e 0,7, que expressam uma correlação entre itens moderada a forte. Além da intensidade da associação desses itens, averiguou-se se as correlações eram significativas através de um coeficiente de correlação.

Previamente à aplicação do coeficiente de correlação, verificou-se se os dados das variáveis, ou seja dos 20 itens da QBPDS, se ajustavam a uma distribuição normal. Os resultados desses testes de normalidade, estão disponíveis no Apêndice I. Pela aplicação do teste Kolmogorov-Smirnov, foi encontrado um valor  $p \leq 0,005$  para todos os itens, de onde se depreende que as variáveis não seguem uma distribuição normal. Prossegue-se com o coeficiente de correlação de Spearman, cujos resultados encontram-se no Apêndice I. Obtêm-se correlações significativas entre os itens, para um nível de significância de 0,05. A única exceção é a correlação da questão 2 com a questão 6, com um valor  $p$  de 0,76. A correlação destas duas questões foi precisamente a que pontuou menor correlação item-item, no valor 0,157. Todavia, não se atribui nenhuma particular relevância a este caso singular.

Para complementar, as correlações item-total corrigidas (Tabela 14) têm valores intermédios e relativamente homogêneos, que variam entre 0,474 e 0,791. Esses valores revelam que as correlações entre cada item e a pontuação total do questionário são moderadamente fortes e indicam que os itens contribuem de forma relativamente igual para avaliar a incapacidade funcional.

Verifica-se ainda, pela Tabela 14, que a eliminação de qualquer questão não altera significativamente o coeficiente de  $\alpha$  de Cronbach, ou seja, não há nenhum item cuja exclusão aumente substancialmente a consistência interna da escala. Estes resultados, corroboram com os resultados da AFE, na medida que revelam argumentos para não excluir nenhuma questão da escala.

**Tabela 13 – Correlações item-item**

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
Q1	1,000	0,563	0,635	0,491	0,418	0,282	0,502	0,534	0,437	0,440	0,420	0,321	0,471	0,598	0,570	0,563	0,511	0,429	0,438	0,370
Q2	0,563	1,000	0,634	0,476	0,248	0,157	0,437	0,425	0,362	0,325	0,390	0,323	0,387	0,465	0,374	0,361	0,468	0,352	0,376	0,347
Q3	0,635	0,634	1,000	0,421	0,384	0,307	0,544	0,495	0,491	0,391	0,422	0,450	0,408	0,469	0,498	0,531	0,493	0,424	0,479	0,424
Q4	0,491	0,476	0,421	1,000	0,411	0,399	0,585	0,522	0,463	0,425	0,488	0,400	0,430	0,505	0,528	0,475	0,509	0,389	0,509	0,417
Q5	0,418	0,248	0,384	0,411	1,000	0,455	0,489	0,401	0,487	0,459	0,424	0,336	0,394	0,469	0,405	0,488	0,338	0,422	0,486	0,456
Q6	0,282	<b>0,157*</b>	0,307	0,399	0,455	1,000	0,421	0,361	0,413	0,476	0,407	0,329	0,341	0,343	0,396	0,377	0,313	0,188	0,303	0,323
Q7	0,502	0,437	0,544	0,585	0,489	0,421	1,000	0,728	0,634	0,575	0,596	0,574	0,425	0,482	0,517	0,560	0,536	0,604	0,586	0,562
Q8	0,534	0,425	0,495	0,522	0,401	0,361	0,728	1,000	0,690	0,563	0,603	0,479	0,459	0,460	0,449	0,498	0,566	0,533	0,639	0,503
Q9	0,437	0,362	0,491	0,463	0,487	0,413	0,634	0,690	1,000	0,549	0,522	0,625	0,391	0,420	0,487	0,520	0,497	0,586	0,655	0,618
Q10	0,440	0,325	0,391	0,425	0,459	0,476	0,575	0,563	0,549	1,000	0,741	0,630	0,423	0,479	0,436	0,507	0,544	0,643	0,645	0,631
Q11	0,420	0,390	0,422	0,488	0,424	0,407	0,596	0,603	0,522	0,741	1,000	0,583	0,541	0,513	0,440	0,504	0,651	0,644	0,582	0,495
Q12	0,321	0,323	0,450	0,400	0,336	0,329	0,574	0,479	0,625	0,630	0,583	1,000	0,316	0,345	0,411	0,538	0,524	0,572	0,615	0,645
Q13	0,471	0,387	0,408	0,430	0,394	0,341	0,425	0,459	0,391	0,423	0,541	0,316	1,000	0,716	0,523	0,563	0,651	0,505	0,489	0,439
Q14	0,598	0,465	0,469	0,505	0,469	0,343	0,482	0,460	0,420	0,479	0,513	0,345	0,716	1,000	0,663	0,632	0,624	0,532	0,531	0,526
Q15	0,570	0,374	0,498	0,528	0,405	0,396	0,517	0,449	0,487	0,436	0,440	0,411	0,523	0,663	1,000	0,687	0,519	0,472	0,508	0,426
Q16	0,563	0,361	0,531	0,475	0,488	0,377	0,560	0,498	0,520	0,507	0,504	0,538	0,563	0,632	0,687	1,000	0,559	0,569	0,643	0,637
Q17	0,511	0,468	0,493	0,509	0,338	0,313	0,536	0,566	0,497	0,544	0,651	0,524	0,651	0,624	0,519	0,559	1,000	0,670	0,553	0,532
Q18	0,429	0,352	0,424	0,389	0,422	0,188	0,604	0,533	0,586	0,643	0,644	0,572	0,505	0,532	0,472	0,569	0,670	1,000	0,773	0,713
Q19	0,438	0,376	0,479	0,509	0,486	0,303	0,586	0,639	0,655	0,645	0,582	0,615	0,489	0,531	0,508	0,643	0,553	0,773	1,000	0,802
Q20	0,370	0,347	0,424	0,417	0,456	0,323	0,562	0,503	0,618	0,631	0,495	0,645	0,439	0,526	0,426	0,637	0,532	0,713	<b>0,802**</b>	1,000

\*Valor inferior; \*\* Valor superior

**Tabela 14 – Correlações item-total**

	<b>Média da escala se o item for eliminado</b>	<b>Variância da escala se o item for eliminado</b>	<b>Correlações item-total corrigidas</b>	<b>“Squared Multiple Correlation”</b>	<b><math>\alpha</math> de Cronbach se o item for eliminado</b>
Q1	34,09	306,587	0,648	0,615	0,948
Q2	34,02	308,496	0,532	0,525	0,950
Q3	34,05	306,853	0,646	0,602	0,948
Q4	34,30	306,610	0,641	0,534	0,948
Q5	33,41	307,862	0,579	0,445	0,949
Q6	33,23	311,551	<b>0,474*</b>	0,455	0,951
Q7	34,07	298,293	0,765	0,707	0,946
Q8	34,44	298,279	0,727	0,725	0,947
Q9	33,49	294,908	0,729	0,661	0,947
Q10	34,15	295,290	0,732	0,706	0,947
Q11	34,56	299,241	0,734	0,705	0,947
Q12	33,37	292,815	0,667	0,636	0,948
Q13	34,95	308,615	0,641	0,614	0,948
Q14	34,16	299,997	0,706	0,701	0,947
Q15	34,04	300,495	0,674	0,630	0,948
Q16	33,48	297,824	0,750	0,678	0,947
Q17	34,88	304,886	0,737	0,685	0,947
Q18	33,98	294,977	0,742	0,777	0,947
Q19	33,62	294,695	<b>0,791**</b>	0,805	0,946
Q20	32,95	295,632	0,735	0,757	0,947

\* Valor inferior; \*\* Valor superior

#### 4.4. Fiabilidade teste–reteste

Para o efeito do estudo da fiabilidade do tipo estabilidade temporal, pretendido neste subcapítulo, apresentam-se os resultados referentes à estimativa fiabilidade teste-reteste pré intervenção da Fisioterapia (Intervalo de 4 a 7 dias) (Apêndice J)

Dos 132 participantes que completaram corretamente o preenchimento de T0 e T1 (4 a 7 dias de intervalo), houve 104 indivíduos (79%) que reportaram que a sua condição de DLC se mantinha estável (PGIC em T1= 1, 2, 3 ou 4). Esses participantes constituíram o grupo “sem alterações na PGIC” em T1, a partir do qual se desenvolveu o estudo da fiabilidade teste-reteste. A Tabela 15 resume a estatística descritiva das pontuações totais de QBPDS obtidas em T0 e T1 para esses mesmos participantes.

**Tabela 15** – Resultados da QBPDS T0 e QBPDS T1 do grupo “sem alterações na PGIC”

<b>Grupo insucesso T1</b>	<b>QBPDS T0</b>	<b>QBPDS T1</b>
N	104	104
Média (DP)	34,45 ( $\pm$ 17,275)	34,02 ( $\pm$ 13,611)
Mediana	33,00	33,50
Moda	33	26
Mínimo - Máximo	4 - 80	8 - 80

Os resultados obtidos nas duas medições de T0 e T1, revelam pontuações totais médias da QBPDS muito próximos, de 34,45 e 34,02, respetivamente. Para estudar a fiabilidade do tipo estabilidade temporal da QBPDS (com uma semana de intervalo) procedeu-se ao cálculo do CCI segundo o teste ANOVA (*two-way random*). Os resultados estão disponíveis na Tabela 16.

O CCI para o grupo “sem alterações na PGIC” (entre T0 e T1) é de 0,696 (IC 95%: 0,581–0,783), o que representa uma fiabilidade teste-reteste razoável (Hill & Hill, 2005). Segundo Terwee *et al.* (2007), um CCI, de pelo menos 0,70, recebe uma classificação positiva. Assim, o valor de CCI obtido é satisfatório.

**Tabela 16** – Resultado da estimativa de CCI (T0-T1) para o grupo “sem alterações na PGIC”

	CCI	Intervalo de confiança 95%	
		Limite inferior	Limite superior
Medições individuais	0,696	0,581	0,783

#### **4.5. Estimativa do EPM e DMD**

##### **4.5.1. Amostra T2 e resultados ao nível da incapacidade para os grupos “sem alterações” e “melhoria clínica”**

A amostra T2 foi constituída por todos os indivíduos que completaram 6 semanas de tratamento e que foram avaliados no momento T2. Assim, dos 132 indivíduos que constituíram a amostra inicial (T0), 11 foram excluídos porque abandonaram os tratamentos de Fisioterapia antes das 6 semanas de intervenção e o

restante foi excluído por lacunas no preenchimento da QBPDS em T2 (6 itens não respondidos).

A Tabela 17 resume os resultados da QBPDS em T0 e em T2, para cada um dos níveis de Percepção Global de melhoria (em T2) e para cada grupo (“sem alterações” e “melhoria clínica”), apresentando ainda o número de participantes que constituíram cada nível e cada grupo (Apêndice K).

**Tabela 17** – Resultados da QBPDS no momento inicial (T0) e após 7 semanas (T2) para os diferentes níveis de PGIC e em cada um dos grupos.

PGIC (T2)	QBPDS_T0	QBPDS_T2	Diferença	Diferença relativa
	Média (± DP)	Média (± DP)	QBPDS T0-T2 Média (± DP)	QBPDS T0-T2 % Média (± DP)
1 (n=1)	55 (± 0)	80 (± 0)	-25 (± 0)	45,5%
2 (n= 8)	27,9 (± 19,3)	29,3 (± 21,4)	- 1,4 (± 5,6)	1% (± 25,7%)
3 (n=22)	32,4 (± 17)	30,7 (± 18,7)	1,7 (± 11,1)	-1,9% (± 39,7%)
4 (n=13)	42,2 (± 21,2)	34,6 (± 16,4)	7,5 (± 11,2)	-11,9% (± 33,7%)
5 (n=35)	37,9 (± 15,4)	27,5 (± 15,5)	10,3 (± 9,4)	-27,9% (± 26,7%)
6 (n= 31)	40,3 (± 21,3)	24,1 (± 13,8)	16,2 (± 15,9)	- 38% (± 24,7%)
7 (n=10)	28 (± 12,9)	17,5 (± 15,4)	10,5 (± 13,0)	-33,9% (± 50,9%)
<b>Total (n=120)</b>	36,6 (± 18,4)	27,7 (± 17,1)	8,9 (± 13,4)	-22% (± 35,1%)
Grupo “Sem alterações”(n=44)	35 (±19,1)	32,7 (± 19,4)	2,3 (± 11,3)	-3,3% (± 35,7%)
Grupo “Melhoria clínica” (n=76)	37,5 (± 18)	24,8 (± 15)	12,7 (± 13,0)	-32,8% (± 30%)

Na QBPDS os resultados mais elevados indicam maior nível de incapacidade. A diferença nos resultados no QBPDS (T0-T2) foi calculada subtraindo os resultados da pré-intervenção pelos da pós-intervenção, para cada indivíduo. Assim, as diferenças na pontuação absoluta de T0 para T2, que sejam positivas, significam que houve uma redução na pontuação total da QBPDS entre os dois momentos, ou seja, uma redução no nível de incapacidade funcional reportado. Sendo que, valores positivos correspondem a melhorias ao nível da incapacidade funcional, e valores negativos, representam um agravamento da condição clínica.

Por outro lado, uma diferença relativa na pontuação de QBPDS de T0 para T2, que seja positiva, subentende um aumento percentual na pontuação total da QBPDS de

T2 em relação a T0, o que significa um aumento da incapacidade funcional entre esses momentos (agravamento da condição clínica). Pelo contrário, uma diferença relativa da QBPDS, negativa, identifica uma melhoria clínica ao nível da incapacidade funcional.

Pela análise da Tabela 17, observa-se que na totalidade da amostra (n=120), após as 7 semanas de intervalo, o nível de incapacidade funcional reportado, pela pontuação da QBPDS, decresceu de uma média inicial de 36,6 ( $\pm$  18,4), para uma média de 27,7 ( $\pm$  17,1). Ou seja, na globalidade, os utentes experienciaram uma melhoria no nível de incapacidade funcional reportada (uma diminuição média na QBPDS de 8,9 pontos,  $\pm$  13,4).

Os resultados indicam que os utentes, após as 7 semanas (T2), dividem-se pelos vários níveis da perceção global de melhoria e que houve preponderância do número de indivíduos que se autotestou no nível 5 e 6 da PGIC (n= 35 e n=31, respetivamente). Apenas um indivíduo assinalou a opção 1 da PGIC: "sem alterações (ou a condição piorou)", sendo-lhe correspondida uma diferença média na pontuação da QBPDS T0-T2, de 25 pontos negativos, que podem representar um eventual agravamento da incapacidade funcional.

A formação, *a posteriori*, dos grupos "sem alterações" (PGIC=1, 2, 3 e 4) e "melhoria clínica" (PGIC= 5, 6 e 7), conforme estava definido na metodologia, resultou no grupo "sem alterações" com 44 indivíduos, que representa 36,7% da amostra e no grupo "melhoria clínica" constituído pelos restantes 76 indivíduos (63,3% da amostra).

Ao nível dos grupos, a diferença média da QBPDS T0-T2 foi positiva para ambos, no entanto é notável que a diferença sentida no grupo "melhoria clínica" (12,7 pontos,  $\pm$  13,0) é substancialmente maior do que a diferença média do grupo "sem alterações" (2,3 pontos,  $\pm$  11,3).

#### **4.5.2. Comparação dos resultados de incapacidade funcional em T0 e T2, para ambos os grupos**

Contextualizada na estimativa do EPM e DMD, a comparação das pontuações da QBPDS em T0 e em T2, surge no sentido de confirmar que no grupo que se autoreportou "sem alterações", as pontuações da QBPDS não diferiram significativamente entre os dois momentos de avaliação; e pelo contrário, verificar-se

uma diferença estatisticamente significativa entre a QBPDS de T0 e de T2, no grupo “melhoria clínica”.

Este resultado, contribui para averiguar que a formação dos grupos “sem alterações” e “melhoria clínica” (com base na resposta simultânea da PGIC) também comporta a estabilidade vs instabilidade das pontuações da QBPDS em relação ao ponto de partida, respetivamente. O mesmo resultado serve ainda para aumentar a segurança relativa à estabilidade da condição clínica (inclusive ao nível da incapacidade funcional reportada) do grupo considerado para estimar o EPM e DMD.

Todos os resultados necessários para este ponto do trabalho, estão disponíveis no Apêndice K. A aplicação de um teste paramétrico de comparação de médias para amostras emparelhadas, pressupõe que se verifique, em primeiro lugar, a normalidade dos dados. Analisa-se a distribuição normal das variáveis “QBPDS T0” e “QBPDS T2”, em cada um dos grupos. Para o grupo “sem alterações” (n<50), utilizaram-se os resultados do teste Shapiro-Wilk e para o grupo “melhoria clínica” (n>50) recorreu-se aos resultados do Kolmogorov-Smirnov (Tabela 18):

**Tabela 18** – Resultados dos testes de normalidade das variáveis “QBPDS T0” e “QBPDS T2” (nos grupos “sem alterações” e “melhoria clínica”).

Variáveis		Kolmogorov-Smirnov		Shapiro-Wilk	
		Estatística de teste	Valor p	Estatística de Teste	Valor p
Grupo “sem alterações” (n=44)	QBPDS T0	---	---	0,955	0,087
	QBPDS T2	---	---	0,944	0,033*
Grupo “melhoria clínica” (n=76)	QBPDS T0	0,143	0,001*	---	---
	QBPDS T2	0,088	0,200	---	---

\* valor  $p \leq (\alpha = 0,05)$

Para qualquer um dos grupos foi encontrado, pelo menos um valor  $p \leq 0,05$ , de modo que não se verifica o pressuposto de que as variáveis seguem uma distribuição normal, necessário à aplicação de um teste paramétrico de comparação de médias para amostras emparelhadas.

Em alternativa ao teste T para amostras emparelhadas, recorre-se ao Teste de Wilcoxon ( $p \leq 0,05$ ). Testa-se a hipótese nula de que as variáveis (QBPDS em T0 e QBPDS em T2) seguem uma distribuição idêntica em tendência central. Os resultados são apresentados na Tabela 19.

**Tabela 19** – Resultado do teste estatístico de Wilcoxon das distribuições “QBPDS T0” e “QBPDS T2” (nos grupos “sem alterações” e “melhoria clínica”).

Variáveis	Grupo	Teste estatístico de Wilcoxon (Z)	Valor p
QBPDS T0 - T2	Grupo “sem alterações” (n=44)	-0,992	0,321
	Grupo “melhoria Clínica” (n=76)	-6,749	0,000*

\* valor  $p \leq (\alpha = 0,05)$

Os resultados obtidos com a aplicação do teste Wilcoxon às variáveis emparelhadas do grupo “melhoria clínica” (estatística de teste  $Z = -0,992$ , valor  $p = 0,000$ ), permitem concluir que há diferenças significativas na pontuação de QBPDS nas duas medições (T0 e T2), para um nível de significância de 0,05.

Por outro lado, o teste de Wilcoxon aplicado às variáveis do grupo “sem alterações” (estatística de teste  $Z = -6,749$ , valor  $p = 0,321$ ) revela um valor  $p$  superior ao nível de significância fixado ( $\alpha = 0,05$ ). Deste modo, assume-se que as duas distribuições são iguais em média.

No global, os resultados confirmam que no grupo “melhoria clínica”, a pontuação total da QBPDS sofreu alterações significativas entre T0 e T2, ao passo que, no grupo “sem alterações”, essa pontuação total se manteve estável entre T0 e T2.

#### **4.5.3. Equivalência dos grupos “sem alterações” e “melhoria clínica”, no ponto de partida**

Ainda no âmbito da estimativa do EPM e DMD, considera-se importante confirmar que as diferenças encontradas ao nível da QBPDS nos dois grupos, não se

devem a diferenças das variáveis no ponto de partida, dos indivíduos que constituem cada um dos grupos. Este procedimento é igualmente um contributo para dar legitimidade à dicotomização da amostra (“sem alterações” e “melhoria clínica”) ao nível do construto “incapacidade funcional”, e para salvaguardar que o EPM e DMD são calculados através de um grupo que se mantém estável em relação ao ponto de partida, precisamente porque não houve alterações na incapacidade funcional (não devendo-se a outras características do grupo).

Todos os resultados deste ponto do trabalho estão disponíveis no Apêndice K.

#### 4.5.3.1. Equivalência da pontuação da QBPDS em T0 e da Intensidade da dor em T0

De acordo a análise descritiva (Tabela 20), os dois grupos apresentam pontuações médias da QBPDS T0, relativamente idênticas, respetivamente, uma média de 35 pontos ( $\pm 19,1$ ) no grupo “sem alterações” e uma média ligeiramente superior no grupo “melhoria clínica”, de 37,5 pontos ( $\pm 18$ ).

A análise do ajustamento das variáveis “QBPDS T0”, de cada um dos grupos, a uma Distribuição Normal (que foi anteriormente realizada - Tabela 18), aferiu que não é satisfeita a normalidade das duas variáveis.

**Tabela 20** – Resultados da QBPDS T0 e EVA T0 nos grupos “sem alterações” e “melhoria clínica” e resultados da estatística do teste de Mann-Whitney.

Variável quantitativa	“Sem alterações” (n= 44)	“Melhoria Clínica” (n= 76)	Teste Mann-Whitney	Valor p
<b>QBPDS T0</b>	35 ( $\pm 19,1$ )	37,5 ( $\pm 18$ )	1517,000	0,398
<b>EVA T0</b>	5,75 ( $\pm 2,0$ )	5,59 ( $\pm 2,1$ )	1633,500	0,832

Do mesmo modo, pela Tabela 20, é possível observar que a perceção de intensidade da dor em T0 (derivada dos resultados da EVA), reportada pelos dois grupos é semelhante: no grupo “sem alterações”, com uma média de 5,75 ( $\pm 2,0$ ), e no grupo “melhoria clínica”, com um valor médio de 5,59 ( $\pm 2,1$ ).

Os Testes de normalidade das variáveis: EVA T0 no grupo “sem alterações” (teste Shapiro-Wilk), e EVA T0 no grupo “melhoria clínica” (Kolmogorov-Smirnov), obtiveram valor p de 0,019 e 0,003, respectivamente (Tabela 21). Para um nível de significância de 0,05, afere-se que não é satisfeita a normalidade das duas variáveis.

**Tabela 21** – Resultados dos testes de normalidade das variáveis: EVA T0 no grupo “sem alterações” e EVA T0 no grupo “melhoria clínica”.

Variáveis		Kolmogorov-Smirnov		Shapiro-Wilk	
		Estatística de teste	Valor p	Estatística de Teste	Valor p
EVA T0	Grupo “sem alterações”	---	---	0,937	0,019
	Grupo “melhoria clínica”	0,129	0,003	---	---

Uma vez que não se verifica a normalidade das variáveis, prossegue-se com o teste não paramétrico, o teste Mann Whitney ( $p \leq 0,05$ ), que permite testar a igualdade do valor médio de QBPDS T0 nos grupos: “sem alterações “ e “melhoria clínica”, e a igualdade do valor médio de EVA T0, nos mesmos grupos.

Na Tabela 20 encontram-se os resultados obtidos com o teste de Mann-Whitney. Para a variável QBPDS T0 obteve-se um valor  $p = 0,398$ , e para a variável EVA T0, um valor  $p = 0,832$ , ou seja, em ambos, verificaram-se valores p superiores ao nível de significância estipulado ( $\alpha=0,05$ ), de modo que não se rejeita  $H_0$ . Assume-se que ambas as variáveis, QBPDS T0 e EVA T0, apresentam duas distribuições (nos grupos: “sem alterações “ e “melhoria clínica”) com valores de tendência central idênticos; e como tal, conclui-se que não há diferenças significativas nem nas pontuações de QBPDS T0, nem na intensidade da dor em T0, entre os dois grupos.

#### 4.5.3.2. Equivalência das variáveis sócio-demográficas e clínicas em T0

O Teste do Qui-Quadrado, testa se a frequência com que os elementos das duas amostras independentes se repartem pelas classes de uma variável qualitativa é ou não aleatória. Uma exigência deste teste é a organização dos dados em tabelas de

contingência, que permitam cruzar os dados das duas variáveis (cada qual com um máximo de 3 categorias). A variável categorias de percepção de melhoria dispõe duas categorias: “melhoria clínica” e “sem alterações”; contudo, as variáveis sócio-demográficas e clínicas, tiveram que ser recodificadas, de forma a assumirem apenas 2 ou 3 categorias.

As classes das variáveis foram agregadas de acordo com a estatística descritiva, nomeadamente pelas frequências absolutas (por exemplo: a duração da dor) e medianas dos resultados (por exemplo: a idade). O Apêndice L resume as categorias das variáveis sócio-demográficas e clínicas, que necessitaram de sofrer essa transformação, e faz corresponder a cada categoria, as classes agregadas que estão na sua origem.

Na Tabela 22, as colunas “sem alterações” e “melhoria clínica” disponibilizam as frequências com que os elementos de cada grupo se repartem pelas categorias de cada variável. De acordo com essa informação, é possível fazer uma comparação inicial e grosseira (por percentagens) das duas amostras. Na Tabela 22, apresenta-se ainda os valores obtidos com o teste Qui-quadrado de Pearson e respetivo valor p, para cada cruzamento das variáveis: Categoria de percepção de melhoria e cada variável sócio-demográfica ou clínica.

**Tabela 22** – Caracterização da amostra T2 em dois grupos (“sem alterações” e “melhoria clínica”) e resultado da estatística de Teste Qui-quadrado de Pearson.

Variáveis qualitativas	“Sem alterações” (n=44)	“Melhoria Clínica” (n= 76)	Qui-quadrado de Pearson	Valor p
<b>Idade</b>	46,68 (± 12,64)	47,46 (± 12,25)	0,107	0,849*
≤ 49 anos	22 (50%)	39 (51,3%)		
> 49 anos	22 (50%)	37 (48,7%)		
<b>Género</b>			1,451	0,288*
Masculino	9 (20,5%)	24 (31,6%)		
Feminino	35 (79,5%)	52 (68,4%)		
<b>IMC em 2 categorias</b>	(n= 42; 95,5%)		2,094	0,168*
Peso normal	13 (31%)	32 (42,1%)		
Excesso de peso	29 (69%)	44 (57,9%)		
<b>Estado civil</b>			1,831	0,234*
Com suporte familiar	32 (72,7%)	45 (59,2%)		
Sem suporte familiar	12 (27,3%)	31 (40,8%)		

<b>Nível de escolaridade</b>			3,013	0,222*
Ensino Primário	15 (34,1%)	16 (21,1%)		
Ensino Básico e Secundário	21 (47,7%)	40 (52,6%)		
Ensino Superior	8 (18,2%)	20 (26,3%)		
<b>Atividade Profissional</b>		N = 74 (97,4%)	3,057	0,217*
Intelectual	6 (13,6%)	21 (28,4%)		
Técnico	21 (47,7%)	28 (37,8%)		
Operacional	17 (38,6%)	25 (33,8%)		
<b>Situação profissional atual</b>			1,783	0,410*
Ativo	31 (70,5%)	45 (59,2%)		
Não ativo	10 (27,7%)	20 (26,3%)		
Doméstica (o)	3 (6,8%)	11 (14,5%)		
<b>Duração da dor</b>			0,015	1,000*
≤ 24 meses	15 (34,1%)	26 (34,2%)		
> 24 meses	29 (65,9%)	50 (65,8%)		
<b>Localização da dor</b>			0,855	0,446*
Sem dor referida para o membro inferior	18 (40,9%)	39 (51,3%)		
Com dor referida para o membro inferior	26 (59,1%)	37 (48,7%)		
<b>Medicação</b>			0,166	0,707*
Não	33 (75%)	40 (52,6%)		
Sim	11 (25%)	36 (47,4%)		
<b>Faltas ao trabalho</b>			0,623	0,534*
Não	33 (75%)	51 (67,1%)		
Sim	11 (25%)	25 (32,9%)		
<b>Total de faltas</b>	<b>(n=11; 25%)</b>	<b>(n=24; 31,6%)</b>	---	---
1 vez	2 (18,2%)	5 (20,8%)		
2 vezes	2 (18,2%)	2 (8,3%)		
3 vezes	2 (18,2%)	4 (16,7%)		
Mais de 3 vezes	5 (45,5%)	13 (54,2%)		
<b>Duração das faltas (n=34)</b>	<b>(n=11; 25%)</b>	<b>(n= 23; 30,3%)</b>	---	---
1 dia	1 (9,1%)	2 (8,7%)		
2 dias	2 (18,2%)	2 (8,7%)		
3 dias	2 (18,2%)	5 (21,7%)		
1 semana	1 (9,1%)	3 (13,0%)		
Mais de 1 semana	5 (45,5%)	11 (47,8%)		
<b>Baixa Remunerada</b>			0,202	0,815*
Não	36 (81,8%)	59 (77,6%)		
Sim	8 (18,2%)	17 (22,4%)		

\* valor p > ( $\alpha= 0,05$ )

O teste do Qui-quadrado, testa a hipótese nula de que a variável Categoria de percepção de melhoria e a variável sócio-demográfica ou clínica não estão relacionadas uma com a outra. Por exemplo, que a distribuição da idade é independente do grupo (“sem alterações” ou “melhoria clínica”).

Uma vez que, o valor p obtido em todas as variáveis sócio-demográficas e clínicas é superior a 0,05, então, não se rejeita a Hipótese nula e conclui-se pela independência das duas variáveis (Categoria de percepção de melhoria e a variável sócio-demográfica ou clínica).

Pode-se inferir que não há diferenças significativas entre os dois grupos, ao nível das características do ponto de partida, dos utentes que constituem cada um dos grupos. Deste modo, as diferenças na pontuação da QBPDS, entre T0 e T2, que surgem entre os grupos, são aparentemente resultado da ocorrência, ou não ocorrência, de mudança na incapacidade funcional dos utentes com DLC.

#### **4.5.4. Fiabilidade teste-reteste, pré-pós intervenção da Fisioterapia (intervalo de 7 semanas)**

A metodologia estabelecida para este estudo da fiabilidade, define que apenas são incluídos os participantes que se autoclassificaram como não tendo percecionado alterações na sua condição (pontuação da PGIC em T2= 1,2,3,e 4) e que constituíram o grupo “sem alterações” T2 (n= 44). Porém, pela observação da Tabela 17, nota-se que o único indivíduo que se reportou no nível 1 da PGIC, obteve uma diferença na pontuação da QBPDS T0-T2, de 25 pontos negativos. Essa diferença nas pontuações é substancial e correlacionada com a descrição da PGIC =1: “sem alterações (ou a condição piorou)”, que aponta para um possível agravamento da incapacidade funcional.

Para salvaguardar a estabilidade da condição clínica, exigida no cálculo da fiabilidade teste-reteste, foi constituído um sub-grupo nomeado “PGIC=2,3 e 4” (n=43), que excluía o único indivíduo que se reportou no nível 1 da PGIC. Com base nesse sub-grupo foi estimada a fiabilidade teste-reteste (T0-T2b), que foi comparada com a estimativa da fiabilidade teste-reteste obtida para o grupo “sem alterações” (T0-T2). Os resultados de ambas as estimativas são apresentados no Apêndice K.

A estimativa do CCI, com base no sub-grupo “PGIC 2,3,4”, revela um acréscimo de 0,01 em relação à estimativa do CCI do grupo “sem alterações”. Conclui-se que a remoção desse único participante, não representa uma alteração significativa no CCI. Uma vez que os resultados do subgrupo “PGIC 2,3,4” foram idênticos à estimativa da fiabilidade teste-reteste para o grupo “sem alterações” (PGIC=1,2,3,e 4), decidiu-se que, no âmbito da estimativa da DMD, seria utilizado o resultado da fiabilidade teste-reteste para o grupo “sem alterações” (PGIC=1,2,3,e 4).

A fiabilidade teste-reteste da QBPDS (T0-T2) foi estimada através do CCI (tipo 2,1), calculado com os resultados da QBPDS de T0 e T2, dos 44 indivíduos do grupo “sem alterações”. O valor de CCI obtido de 0,827 (IC 95%: 0,705-0,902) (Tabela 23), exprime um bom nível de fiabilidade do tipo estabilidade temporal (Hill & Hill, 2005).

**Tabela 23** – Resultado da estimativa de CCI (T0-T2) para o grupo “sem alterações”

	CCI	Intervalo de confiança 95%	
		Limite inferior	Limite superior
Medições individuais	0,827	0,705	0,902

#### 4.5.5. Cálculo do EPM e DMD

Para calcular o EPM (T0-T2) associado à repetição destas medições de QBPDS em T0 e T2, utilizou-se a fórmula estatística:  $EPM = DP_{médio} \times \sqrt{1 - r}$ , em que:

- $DP_{médio}$ , refere-se à média dos desvios padrões do momento inicial (T0), e após 7 semanas (T2), dos indivíduos clinicamente estáveis (grupo “sem alterações”). Pela observação da Tabela 17, têm-se que  $DP_{médio} = (19,1 + 19,4)/2 = 19,25$ ; e,
- R, é o coeficiente de correlação utilizado, neste caso o CCI (T0-T2), previamente calculado, de 0,827 (IC 95%: 0,705-0,902)

Assim:

$$1) EPM = 19,25 \times \sqrt{1 - 0,827} \left[ IC\ 95\% : (19,25 \times \sqrt{1 - 0,902}) - (19,25 \times \sqrt{1 - 0,705}) \right] = 8 \ (IC\ 95\%: 6 - 10,5)$$

$$2) EPM_{repetição} = 8 \times \sqrt{2} [IC\ 95\%: (6,0 \times \sqrt{2}) - (10,5 \times \sqrt{2})] = 11,3 (IC\ 95\%: 8,4 - 14,9)$$

Obtém-se um EPM no valor de 8 pontos. Este valor é considerado para o cálculo da DMD, que foi estimada para um intervalo de confiança de 90%. Assim:

$$DMD_{90} = 1,65(\text{valor } z \text{ para o percentil } 90) \times \sqrt{2} \times 8 = 18,7$$

$$DMD_{90} = 18,7(IC\ 90\%: 14 - 24,5)$$

Uma  $DMD_{90}$  de 18,7 significa que é necessária uma diferença de 19 pontos na pontuação total da QBPDS, respondida em dois momentos diferentes, para traduzir uma diferença real.

#### 4.6. Determinação dos efeitos de “chão” e de “teto”

Depois de conhecido a  $DMD_{90}$  (19 pontos) é possível avaliar a amplitude da escala válida para detetar mudanças reais na pontuação total da QBPDS ao longo do tempo. Neste caso a amplitude válida da QBPDS (pontuação total) é de 19 a 81 pontos, sendo que 19 e 81 correspondem ao limite inferior e superior, respetivamente, da pontuação da QBPDS no momento inicial, que permite uma mudança na pontuação na próxima reavaliação, pelo menos igual à  $DMD_{90}$ .

Relativamente à presente amostra, calculou-se a proporção de indivíduos (de  $n=120$ ) que reportaram uma pontuação inicial na QBPDS inferior ao limite inferior válido da QBPDS ( $QBPDS\ T0 < 19$ ), para os quais não foi possível detetar uma melhoria real da incapacidade funcional; e a percentagem de indivíduos cuja pontuação inicial na QBPDS foi superior ao limite superior válido da QBPDS ( $QBPDS > 81$ ), para os quais não foi possível detetar um agravamento real da incapacidade funcional. As frequências absolutas e percentagens, relativas ao número de participantes com pontuações totais de QBPDS  $<19$ ,  $19-81$  e  $>81$ , estão disponíveis na Tabela 24 (Apêndice M).

**Tabela 24** – Resultados da frequência de QBPDS T0 dentro dos limites válidos da escala.

QBPDS T0 (n=120)	Frequência	Porcentagem
<19	19	15,8
19-81	100	83,3
>81	1	0,8

Os resultados revelam que 15,8% dos participantes tem uma pontuação total de QBPDS, em T0, inferior ao limite inferior da amplitude disponível da QBPDS (efeito de “chão”). Por outro lado, apenas um participante tem uma pontuação total de QBPDS superior ao limite superior válido, o que corresponde a 0,8% da amostra. De modo que no momento inicial de avaliação, 83,3% da amostra apresenta pontuações da QBPDS compreendidas na amplitude da escala que é válida.

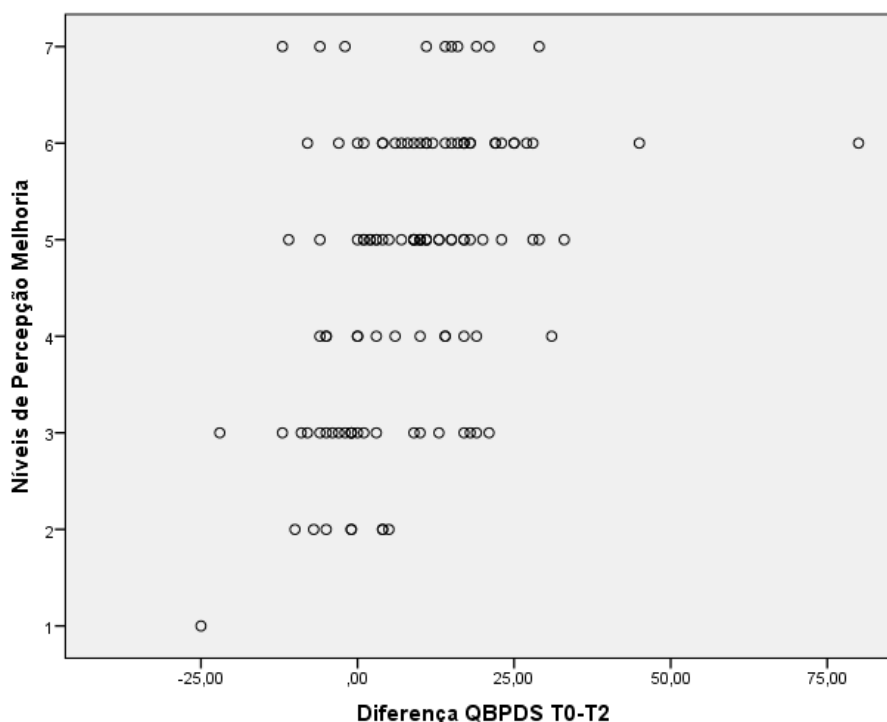
De acordo com a metodologia proposta por Terwee *et al.* (2009), os questionários com mais de 15% dos participantes a pontuar acima do limite superior ou com mais de 15% dos participantes a pontuar abaixo do limite inferior, da amplitude válida da escala, denunciam efeito de “chão” e/ou de “teto”. Na respectiva amostra, os utentes com pontuações da QBPDS em T0, inferiores ao limite inferior, ultrapassam apenas 0,8% dos 15% estipulados, de modo que não se considera uma violação do critério de Terwee *et al.* (2009). Concluiu-se que nesta amostra, não são identificados efeitos de “chão” nem de “teto” da QBPDS -VP.

#### **4.7. Poder de resposta da QBPDS**

Este estudo está particularmente interessado na definição de poder de resposta enquanto capacidade de um instrumento detetar mudanças clinicamente relevantes. O índice do poder de resposta foi calculado com recurso a dois métodos baseados na âncora.

##### **4.7.1. Método “Coeficiente correlação âncora”**

O diagrama de dispersão ilustra a relação entre as variáveis “Diferença de QBPDS T0-T2” e “Níveis de Perceção de Melhoria (T2)” (Figura 7).



**Figura 7** – Diagrama de dispersão entre a diferença QBPDS T0-T2 e os níveis de percepção de melhoria.

Os resultados obtidos revelam uma dispersão nas observações e a presença de *outliers* que podem inflacionar as correlações. Assim, conclui-se que não existe uma tendência linear na associação entre estas duas variáveis.

A ausência de uma relação do tipo linear, associado ao facto de uma das variáveis em estudo ser uma variável qualitativa ordinal (Níveis de Percepção de Melhoria em T2), dispensa a verificação da normalidade das variáveis, no sentido de decidir qual análise que se deve efetuar. As condições indicam a utilização de medida de associação não paramétrica e alternativa ao coeficiente de correlação de Pearson, ou seja, o coeficiente de correlação de Spearman (Tabela 25).

**Tabela 25** - Resultado do coeficiente de correlação de Spearman entre a diferença QBPDS T0-T2 e os níveis de percepção de melhoria.

Variáveis correlacionadas	Coefficiente de correlação de Spearman	Valor p
Diferença QBPDS T0-T2 - Níveis de Percepção Melhoria	0,426	0,000

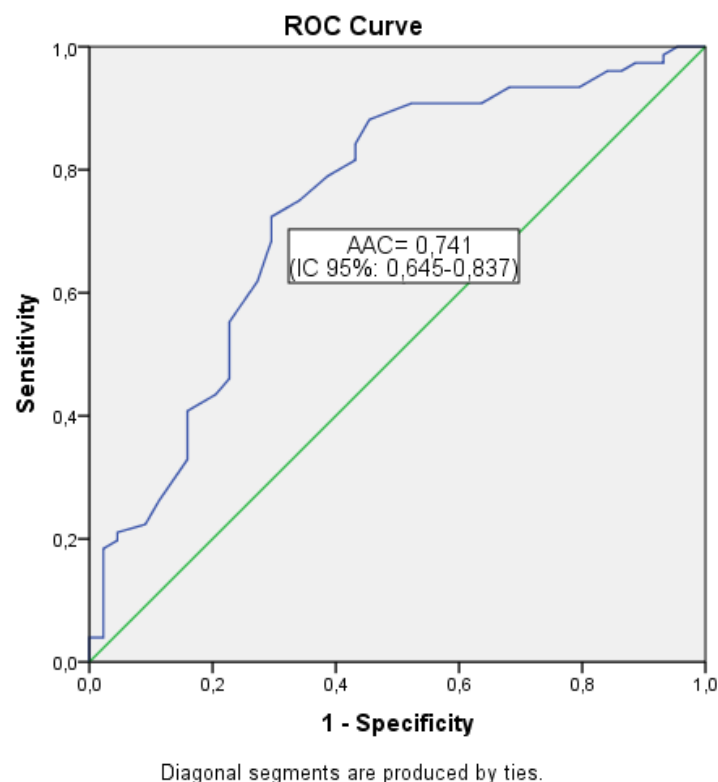
O coeficiente de  $\rho$  de Spearman obtido, é de 0,426, o que sugere a existência de uma correlação positiva e moderada entre as variáveis, ou seja, estas variam no mesmo

sentido e estão moderadamente associadas. Sendo o valor  $p = 0,000 (\leq 0,05)$ , conclui-se que esta correlação é estatisticamente significativa para um nível de significância de 0,05.

Podemos aferir que há uma relação associativa moderada entre as duas variáveis, de modo que um indivíduo que identifique uma maior diferença da pontuação da QBPDS entre T0 e T2, tende a reportar maior nível de percepção de melhoria global.

#### 4.7.2. Método Curva ROC: Área abaixo da curva

Para esta análise, assumiu-se a dicotomização da amostra em “sucesso” (grupo “melhoria clínica” onde  $PGIC = 5,6,7$ ) vs “insucesso” (grupo “sem alterações” onde  $PGIC = 1,2,3,4$ ). A Figura 8 apresenta o gráfico da curva ROC para esse critério de sucesso. Utilizou-se uma abordagem não paramétrica à curva ROC, uma vez que a variável diferença das pontuações da QBPDS T0-T2, não apresentava uma distribuição normal<sup>12</sup>.



**Figura 8** – Curva ROC da QBPDS – VP: estimativa da AAC

<sup>12</sup> Dado que as variáveis QBPDS em T0 e QBPS em T2, não seguem uma distribuição normal (Tabela 18), a diferença destas duas variáveis também não segue uma distribuição normal.

A área abaixo da curva representa o poder de resposta atribuído à QBPDS nesta análise. Neste caso, foi determinada uma AAC de 0,741 (IC 95%: 0,645 – 0,837) (ANEXO N – Tabela Área abaixo da curva). Como é um valor intermédio entre 0,5 e 1, revela uma capacidade moderada de discriminar os indivíduos com sucesso dos indivíduos sem sucesso.

#### **4.8. Determinação da Diferença Mínima Clinicamente Importante**

Neste estudo a DMCI foi determinada por dois métodos baseados em âncoras: o resultado médio dos indivíduos que apresentaram melhorias mínimas mas de forma considerada clinicamente importante (PGIC de 5) e a identificação do valor de mudança mais próximo do topo esquerdo da curva ROC (*Receiver Operating Characteristic*).

##### **4.8.1. Diferença média das pontuações nos indivíduos com PGIC=5**

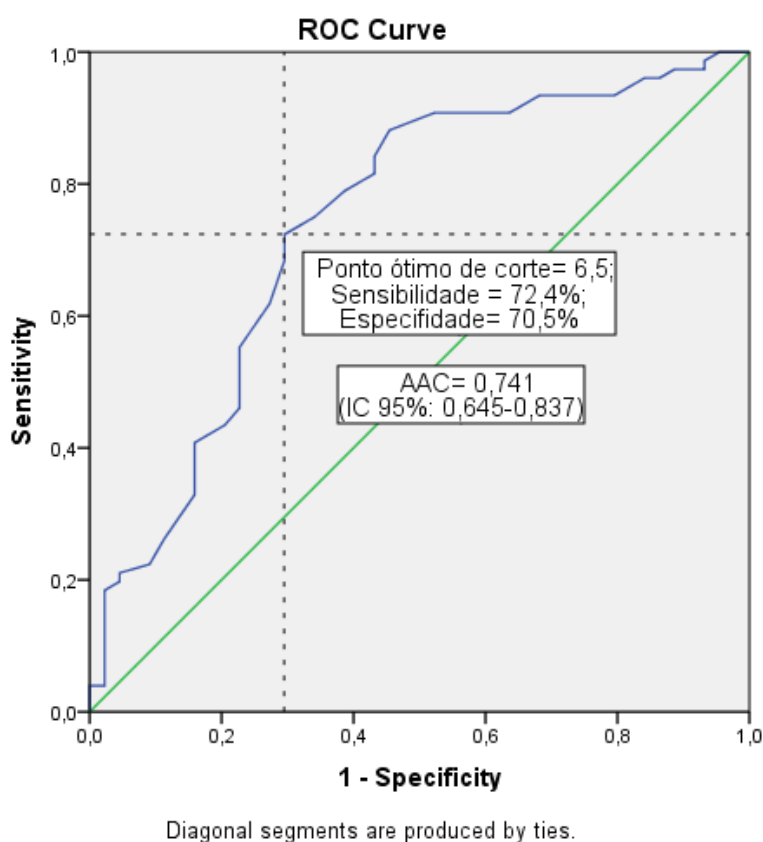
A DMCI foi calculada através do método descrito por Jaeschke *et al.* (1989). Utilizou-se a diferença média na pontuação da QBPDS, de T0 para T2, de todos os indivíduos que se classificaram com uma pontuação “5” na PGIC T2 (o limite inferior do sucesso), ou seja, a diferença média na pontuação dos que percecionam uma melhoria significativa na sua condição clínica, ainda que ligeira. De acordo com a Tabela 17 (página 77), a diferença absoluta da QBPDS, de T0 para T2, no grupo de utentes que reportaram um nível 5 da PGIC, é de 10,3 pontos ( $\pm 9,4$ ); e a diferença absoluta da QBPDS, de T0 para T2, no grupo “sem alterações” em T2, é de 2,25 pontos ( $\pm 11,3$ ). Assim, DMCI é definida pela diferença:  $10,3 - 2,25$ , ou seja 8,05 pontos.

Uma diferença na pontuação de QBPDS de um primeiro momento para um segundo momento de avaliação, igual ou superior a 8 pontos, é considerada uma DMCI.

#### 4.8.2. Método Curva ROC: ponto ótimo de corte

##### a) Ponto ótimo de corte (n=120)

Da análise da curva ROC (Figura 9) e da tabela das coordenadas da curva (no Anexo N), o ponto ótimo de corte da QBPDS, que minimiza o erro geral de classificação, é de 6,5 pontos, com sensibilidade de 72,4% e especificidade de 70,5%.



**Figura 9** – Curva ROC da QBPDS – VP: estimativa da AAC e do ponto ótimo de corte.

Este ponto ótimo de corte significa que, uma diferença na pontuação da QBPDS, de T0 para T2, de 6,5 pontos é o valor de corte que reúne a melhor combinação de sensibilidade e especificidade de todos os outros pontos de corte (ANEXO N - Tabela Coordenadas da curva).

Em termos clínicos, observou-se que, 72,4% dos indivíduos que relataram uma melhoria clínica (grupo do sucesso) tiveram uma diferença na pontuação da QBPDS, de T0 para T2, superior a 6,5 pontos; e que, 70,5% dos indivíduos do grupo insucesso

verificaram uma diferença na pontuação inferior a 6,5 pontos. Assim, uma diferença na pontuação da QBPDS, de T0 para T2, igual ou superior a 6,5 pontos, permitiu identificar corretamente 72,4% dos casos de “melhoria clínica”; E uma diferença na pontuação da QBPDS, de T0 para T2, inferior a 6,5 pontos, permitiu identificar 70,5% dos indivíduos do grupo “sem alterações”.

Tendo em conta, a proximidade dos valores da sensibilidade e especificidade, de 72,4% e 70,5%, respetivamente, subentende-se que há um equilíbrio na discriminação do sucesso e do insucesso.

Dado que o ponto ótimo de corte é uma estimativa da DMCI, de acordo com estes resultados, uma diferença na pontuação da QBPDS, entre duas medições, igual ou superior a 7 pontos, representa uma DMCI.

#### **b) Ponto ótimo de corte expresso em percentagem (n=120)**

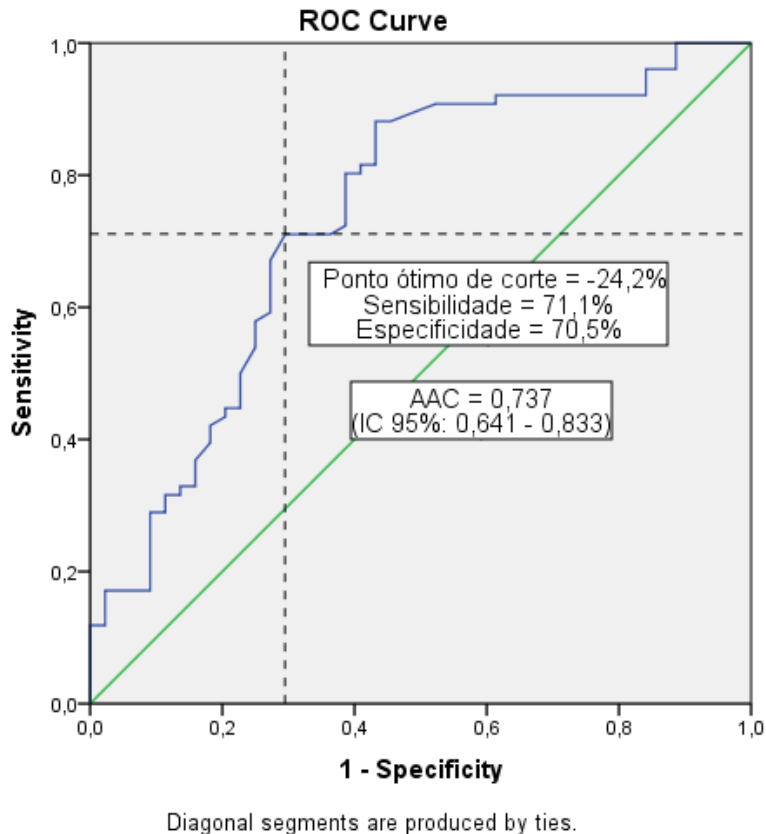
Para a estimativa da DMCI em valor percentual de diferença da pontuação total da QBPDS, de um momento para outro, testou-se a curva ROC com a variável diferença relativa da QBPDS T0-T2 (%) (Figura 10 e Apêndice N).

O ponto ótimo de corte encontrado para a diferença relativa da QBPDS T0-T2, foi de -24,2%, ou seja, identifica que uma redução de 24% da pontuação da QBPDS, no segundo momento, em relação à avaliação anterior, é considerada uma DMCI.

Este ponto ótimo de corte é acompanhado de valores de sensibilidade (71,1%) e especificidade (70,5%) moderados.

Tendo em conta a direção de teste decidida para a curva ROC (na metodologia, ponto 3.7.7.2.), o ponto ótimo de corte de -24,2% deve ser interpretado de forma antagónica à que é mais convencional. Desta forma, depreende-se que os utentes que descrevem uma diferença relativa igual, ou inferior, a -24,2%, experimentam uma melhoria clinicamente importante.

Pelo valor da AAC, de 0,737 (IC 95%: 0,641 – 0,833), ou seja superior a 0,50, pode-se inferir que o modelo apresentado tem uma capacidade moderada a elevada de discriminar os sujeitos que melhoraram e os que não melhoraram.



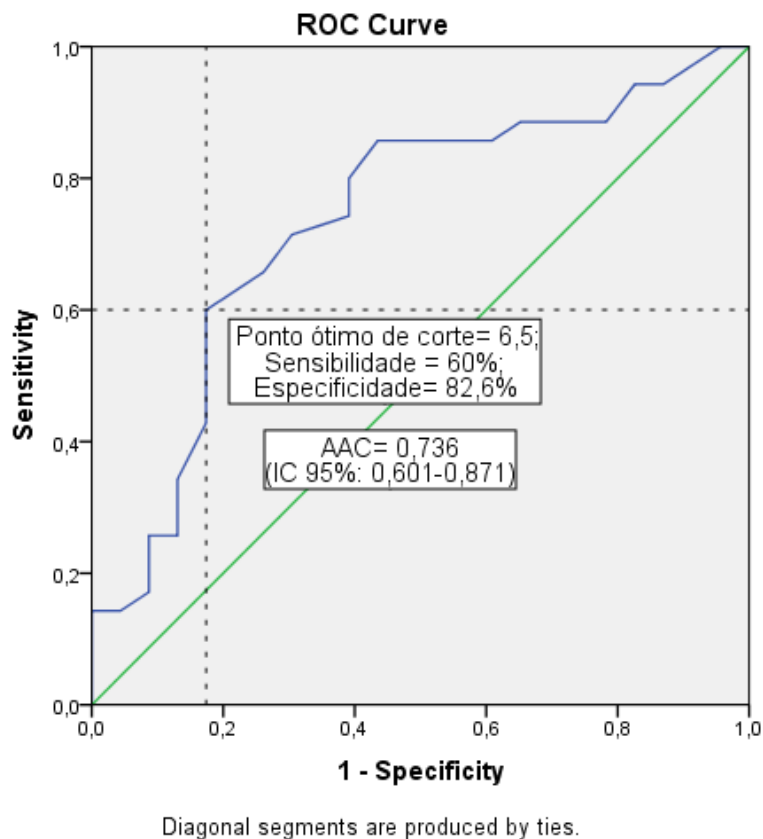
**Figura 10** – Curva ROC da QBPDS – VP para diferenças relativas de QBPDS T0-T2: estimativa da AAC e do ponto ótimo de corte.

**c) Ponto ótimo de corte em função das pontuações da QBPDS no momento inicial:**

- **Pontuações iniciais baixas (QBPDS T0 < 34):**

A Curva ROC, obtida com base na amostra de utentes com pontuações QBPDS baixas no ponto de partida (n=58), representada na Figura 11 (Apêndice N), mostra que o ponto ótimo de corte é de 6,5 pontos, com valores de sensibilidade (60%) e especificidade (82,6%) moderados.

A Curva ROC revela uma AAC de 0,736 (IC 95%: 0,601 – 0,871), que atribui ao modelo proposto uma capacidade moderada a elevada de discriminar indivíduos que melhoraram, dos que não melhoraram.



**Figura 11**– Curva ROC da QBPDs – VP para pontuações iniciais baixas (QBPDs T0< 34): estimativa da AAC e do ponto ótimo de corte.

Conclui-se que, no grupo de indivíduos com baixos níveis de incapacidade funcional iniciais (pontuações na QBPDs T0< 34), uma diferença na pontuação na QBPDs de 7 pontos foi considerada uma DMCI.

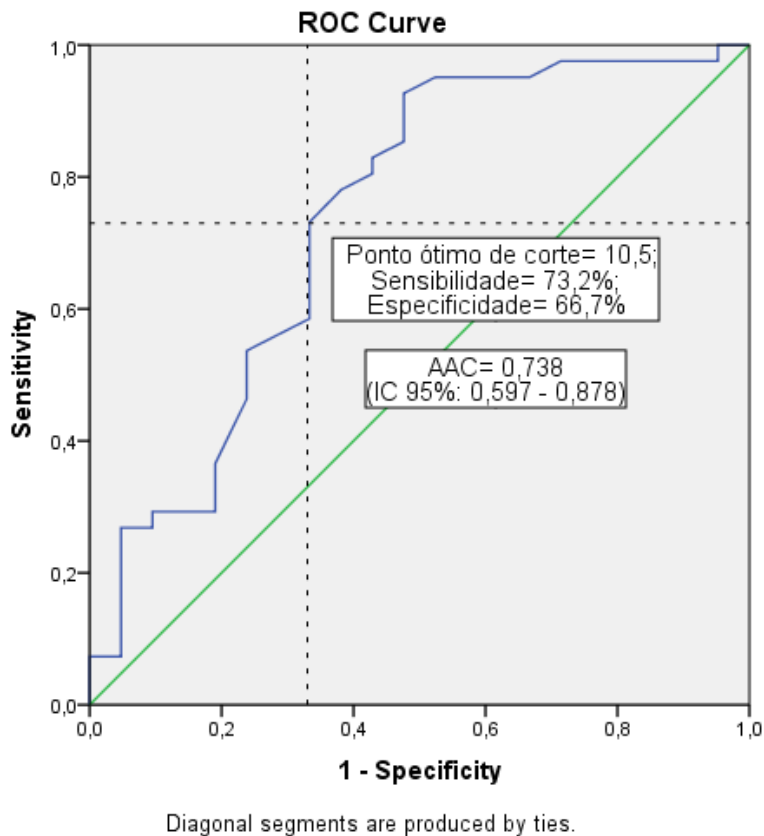
- **Pontuações iniciais altas (QBPDs T0 ≥ 34);**

A amostra de utentes com pontuações QBPDs altas, no ponto de partida, foi constituída por 62 utentes com DLC (n=62).

Com a aplicação da curva ROC (Figura 12 e Apêndice N) foi encontrado um ponto ótimo de corte de 10,5, acompanhado de níveis de sensibilidade (73,2%) e especificidade (66,7%) moderados.

Tal como foi aferido na amostra antagónica (pontuações iniciais baixas), também nesta amostra, obteve-se uma curva ROC com uma AAC moderada a elevada.

O valor de AAC de 0,738 (IC 95%: 0,597 – 0,878), atribui, ao modelo proposto, uma capacidade moderada a elevada, de discriminar os indivíduos que melhoraram, dos que não melhoraram.



**Figura 12**– Curva ROC da QBPDS – VP para pontuações iniciais altas (QBPDS T0  $\geq$  34): estimativa da AAC e do ponto ótimo de corte.

No grupo de indivíduos com elevados níveis de incapacidade funcional iniciais (pontuações na QBPDS T0  $\geq$  34), é necessária uma diferença na pontuação da QBPDS entre medições de 11 pontos para que seja considerada uma DMCI.

#### 4.8.3. DMCI - comparação de estimativas

A Tabela 26, reúne as estimativas de DMCI identificadas neste estudo.

**Tabela 26** – Resultados das estimativas da DMCI da QBPDS – VP.

<b>Método</b>	<b>Diferença média de pontuação</b>	<b>Curva ROC: ponto ótimo de corte (sensibilidade; especificidade)</b>
<b>Estimativa da DMCI</b>	8,05	DMCI = 6,5 (72; 71) DMCI% = -24,2% (71; 71) DMCI <sub>QBPDS &lt;34</sub> = 6,5 (60; 83) DMCI <sub>QBPDS ≥34</sub> = 10,5 (73; 67)

Pelo método “Diferença média de pontuação”, têm-se uma DMCI de 8 pontos e pelo método curva ROC uma DMCI de 7 pontos ou, expressa em percentagem, de - 31%. Ou seja, é necessária uma diferença da pontuação da QBPDS de T0 para T2 de 7/8 pontos ou uma redução de 31% em relação à QBPDS inicial, para ser considerada uma diferença clinicamente significativa.

Da análise da sensibilidade da DMCI às pontuações iniciais da QBPDS, verificou-se que, para pontuações iniciais baixas (QBPDS <34), continua a ser expectável uma diferença de 7 pontos para ter significado clínico. Porém, para utentes com pontuações iniciais altas (QBPDS ≥ 34), o valor de DMCI aumenta, sendo necessária uma diferença de 11 pontos para ser considerada uma mudança clinicamente importante.

#### **4.9. Proporção de DMD e DMCI**

As proporções de indivíduos que excederam a DMD e a DMCI estimadas, encontram-se na Tabela 27. Para a DMCI, admitiu-se o resultado do ponto ótimo de corte da curva ROC, de 7 pontos.

A DMD é uma diferença, que pode ser interpretada para diferenças positivas, entre os dois momentos de avaliação (que expressam redução da incapacidade funcional), como para diferenças negativas (que revelam aumento da incapacidade funcional). Nesta amostra, 17,5% dos indivíduos revelaram uma diferença na QBPDS, pelo menos igual ou superior a 19, revelando assim uma melhoria real; e 1,7%, revelaram uma diferença na QBPDS, pelo menos igual ou inferior a - 19, revelando

assim um agravamento real. Ao todo e de acordo com o valor de DMD, obtêm-se que 19,2% sofreram uma diferença na QBPDS entre T0 e T2, estatisticamente significativa.

No entanto, a proporção dos utentes que perceberam uma melhoria clinicamente significativa foi superior, neste caso, de 57%.

Esta relação da DMD e DMCI é discutida no próximo capítulo de discussão de resultados.

**Tabela 27** – Proporção de indivíduos que excedeu a DMD e a DMCI.

<b>Diferença na QBPDS T0-T2 ≥ DMD (19 pontos)</b>		<b>Diferença na QBPDS T0-T2 ≥ DMCI (7 pontos)</b>
Diferenças positivas (≥ 19)	21/120 = 0,175 (17,5%)	
Diferenças negativas (≤ -19)	2/120 = 0,0167 (1,7%)	68/120 = 0,57 (57%)
Total	23/120 = 0,1917 (19,2%)	

## 5. DISCUSSÃO DOS RESULTADOS

Neste capítulo são discutidos os resultados, que foram apresentados anteriormente, pela mesma ordem dos objetivos propostos. A discussão faz-se em função da evidência anteriormente publicada sobre a QBPDS e do conhecimento científico mais atual.

Este trabalho de investigação englobou um estudo metodológico, multicentro, baseado num *coorte* prospetivo de indivíduos com DLC, com 3 momentos de avaliação (T0, T1 e T2). Nesse estudo participaram 132 indivíduos com DLC que estavam nas listas de espera para receber tratamentos de Fisioterapia. Essa amostra, no ponto de partida (T0), constituída pelos 132 participantes, manteve-se em igual número no segundo momento de avaliação (T1), mas sofreu uma redução de 12 participantes no terceiro momento de avaliação (T2) na consequência de abandono da Fisioterapia (11 participantes) e da lacuna de preenchimento da QBPDS em T2 (1 participante). Assim, da amostra inicial prevaleceram 120 participantes até ao final do estudo.

### 5.1. Caracterização da amostra T0-T1

Para ir ao encontro do primeiro objetivo deste estudo, de avaliação da unidimensionalidade e fiabilidade da QBPDS – VP, são aplicados os resultados da amostra de T0 e T1 (n=132). A caracterização desta amostra cingiu-se aos aspetos vulgarmente analisados nos estudos análogos de outras versões linguísticas da QBPDS (Schoppink *et al.*, 1996, Davidson & Keating 2002; Yvanes-Thomas *et al.*, 2002; Bicer *et al.*, 2005): o género, a idade, nível de escolaridade, situação profissional atual, a duração da dor, localização da dor, medicação e baixa remunerada.

Atualmente, não são se conhecem dados relativos à população de DLC em Portugal, com os quais a amostra deste estudo possa ser comparada. Assim, desconhece-se se esta amostra é representativa da população com DLC em Portugal. Em alternativa, prossegue-se com uma apreciação crítica dos resultados desta amostra, por comparação à amostra do estudo de adaptação e validação da RMDQ para a Língua Portuguesa (n= 112) (Monteiro, Faisca, Nunes & Hipólito, 2010) e às amostras dos estudos anteriores

da QBPDS que tiveram uma amostra de utentes com DLC e de dimensão suficientemente grande (Schoppink *et al.*, 1996, Bicer *et al.*, 2005; van der Roer *et al.*, 2006; Demoulin *et al.*, 2010).

A distribuição da amostra por idade e género é similar à dos estudos anteriores. A idade, em todos os estudos anteriores rondou uma média de 40 (Schoppink *et al.*, 1996) a 45 anos (Monteiro *et al.*, 2010). No presente estudo, foi ligeiramente superior, com uma média de 47 anos ( $\pm 12,67$ ).

Quanto ao género, nesta amostra predomina o feminino, correspondendo a 72,7% da amostra. Este resultado é semelhante ao reportado por Monteiro *et al.* (2010) e Bicer *et al.* (2005) com representações do género feminino nas amostras de 73% e 75,9%, respetivamente. Contudo, há estudos com menor representação, na ordem de: 40% (Schoppink *et al.*, 1996); 47,2% (Demoulin *et al.*, 2010) e 58,7% (van der Roer *et al.*, 2006).

Relativamente ao nível de escolaridade, a amostra em estudo teve uma percentagem de indivíduos com habilitações de nível superior (28%) maior que a verificada por Bicer *et al.* (2005), que foi de 18%. A distribuição do nível de escolaridade, com 25% da amostra de nível básico e 28% da amostra de nível superior, apesar de contribuir para uma amostra equilibrada (com representações dos níveis de escolaridade extremos semelhantes), não é representativa da realidade em Portugal. De acordo com os resultados provisórios dos censos de 2011, em Portugal, a população que possui o ensino superior representa cerca de 12% de população total (Instituto Nacional de Estatística, 2011b) e o ensino básico 1º ciclo é concluído por 25% da população. Curiosamente nesta amostra, o ensino primário teve uma representação idêntica (25%), mas constata-se que há uma representação do nível de instrução universitário superior à da população Portuguesa em geral.

Ao nível da situação profissional atual, a percentagem de indivíduos com trabalho remunerado revela que a amostra encontrava-se maioritariamente ativa (65,9%), equiparável à do estudo de van der Roer *et al.* (2006), em que 75,4% estava ativa, mas oposta à dos estudos de Bicer *et al.* (2005) e Demoulin *et al.* (2010), que tem uma minoria dos utentes no ativo.

Em termos clínicos, a duração da dor teve uma mediana de 24 meses, igual à reportada por Demoulin *et al.* (2010), revelando que nesta amostra prevaleciam as

condições de DLC de longa duração.

Para a análise da baixa remunerada, da medicação e da localização da dor não estão disponíveis amostras de DLC com essas características que possam ser comparadas. Constatou-se apenas que 23% da presente amostra esteve o ano anterior de baixa remunerada devido à sua DLC, que cerca de metade da amostra referia dor para a perna (53%) e também, que metade da amostra toma medicação (45,5%).

Face aos resultados, as características sócio-demográficas desta amostra, de indivíduos com DLC, revelam-se, no geral, dentro dos parâmetros estimados nos restantes estudos de investigação com amostras de DLC. Tratando-se de uma amostra não probabilística, não se podem realizar inferências sobre a população de DLC em Portugal.

Assim, parecem estar reunidas condições para discutir os resultados que advêm desta amostra, relativamente à unidimensionalidade, consistência interna e fiabilidade teste-reteste.

## **5.2. Unidimensionalidade**

A AFE, mostrou que todos os itens da QBPDS – VP, são manifestações do mesmo construto e que eles formam uma escala unidimensional. Conjuntamente, a análise dos itens, revelou que todos os 20 itens estão correlacionados entre si e com a pontuação total. Assim, estes resultados apoiam que a QBPDS é baseada num modelo reflexivo, em que todos os itens são indicadores de efeito. Ou seja, espera-se que todas as pontuações dos itens mudem quando há mudanças no construto (Mokkink et al., 2010b).

De acordo com a bibliografia disponível, o único estudo que precedeu uma AFE da QBPDS, foi o estudo de conceptualização da QBPDS (Kopec *et al.*, 1996), que permitiu o desenvolvimento da escala de QBPDS “reduzida” (20 itens). Nos restantes estudos de adaptação linguística da QBPDS que foram consultados, nenhum investigou a unidimensionalidade da Escala. Apesar de não o terem feito, neste estudo, considerou-se essa etapa como uma condição primordial para avançar com a investigação das restantes propriedades psicométricas. A demonstração de unidimensionalidade potencia

a validade de todas as conclusões seguintes do estudo.

### 5.3. Consistência Interna

A estimativa da consistência interna da QBPDS – VP ( $\alpha$  de Cronbach= 0,95), corrobora com os resultados verificados noutras versões linguísticas da QBPDS (Kopec *et al.*, 1995; Schoppink *et al.*, 1996; Bicer *et al.*, 2005; Mousavi *et al.*, 2006; Melikoglu *et al.*, 2009; Rodrigues *et al.*, 2009). Todos estes estudos, obtiveram coeficientes de  $\alpha$  de Cronbach superiores a 0,90, que indicam uma elevada consistência interna.

Por comparação, na Tabela 28, citam-se apenas os estudos desenvolvidos com amostra de DLC e suficientemente grandes. Confirma-se que o coeficiente estimado no presente estudo (Vieira *et al.*, 2012) está em consonância com os coeficientes estimados anteriormente, em circunstâncias semelhantes.

Embora o valor obtido seja, aparentemente, demasiado elevado, implicando que se coloque a possibilidade de redundância dos itens, deve-se considerar a informação anexa das correlações item-item. Segundo Kopec *et al.* (1996), para evitar a redundância, se 2 ou 3 dos itens candidatos se correlacionam acima de 0,80, apenas um dos itens deve ser selecionado.

De acordo com os resultados deste estudo, apenas umas das correlações item-item foi ligeiramente superior a 0,8, respetivamente 0,802. Assim sendo, no presente caso, em que as correlações têm valores intermédios, é razoável concluir que os itens medem alguma entidade em comum com todos os outros itens - presumivelmente, o construto “nível de incapacidade funcional de utentes com DL”; e, por outro lado, algum aspeto que os outros itens não medem. Presumivelmente, os itens medem diferentes níveis de incapacidade em diferentes atividades funcionais.

Da literatura disponível, o único estudo de validação da QBPDS, que faz alusão à correlação Item-item é o estudo de Yvanes- Thomas *et al.* (2002), que reporta valores de correlação inferiores a 0,80. Esta coerência entre os resultados, dá legitimidade aos resultados apurados. Conclui-se que o questionário QBPDS tem uma excelente consistência interna, que é transversal a todas as suas versões linguísticas, sem que os seus itens sejam redundantes.

**Tabela 28** – Síntese da estimativa de  $\alpha$  de Cronbach da QBPDS - VP neste estudo e da QBPDS em estudos anteriores.

	n (DLC)	Consistência interna ( $\alpha$ de Cronbach)
Schoppink <i>et al.</i> (1996)	120	0,95 (em T0) 0,96 (em T2)
Bicer <i>et al.</i> (2005);	83	0,93
Mousavi <i>et al.</i> (2006);	100	0,92
Melikoglu <i>et al.</i> (2009)	100	0,94 (em T0) 0,95 (em T1)
Vieira <i>et al.</i> (2012)	132	0,95

#### 5.4. Fiabilidade Teste Reteste

Ao longo deste trabalho foi exposta e argumentada a necessidade de estimar a fiabilidade teste-reteste da QBPDS – VP para dois intervalos diferentes, com duas finalidades distintas.

A estimativa do ICC entre o momento inicial (T0) e 4 a 7 dias depois (T1), foi considerada para avaliar a fiabilidade do tipo estabilidade temporal, na medida que esse intervalo teste reteste é o que apresenta argumentos mais válidos, conforme já foi referido. No mesmo sentido, para garantir a estabilidade da condição clínica, apenas foram incluídos os indivíduos do grupo “sem alterações na PGIC” (em T1).

Os resultados revelam o impacto da informação simultânea da PGIC de T1, na determinação da amostra para o estudo da fiabilidade teste-reteste. Os questionários foram administrados a 132 indivíduos com DLC que responderam a T0 e T1, dos quais, apenas 104 (79%) se mantiveram estáveis num intervalo de uma semana e, por essa razão, foram incluídos na estimativa da ICC. Esta percentagem de casos estáveis num intervalo de uma semana, é relativamente baixa para aquilo que era expectável, dada a caracterização deste intervalo, como um presumível indicador de poucas alterações, e dada a natureza crónica da condição. Na literatura não há nenhuma referência, que permita comparar, para uma semana de intervalo, o número de utentes que se mantiveram estáveis.

O CCI calculado para uma semana de intervalo e restrito ao grupo “sem alterações na PGIC”, de 0,696 (IC 95%: 0,581–0,783), embora classificado como razoável (Hill & Hill, 2005) e satisfatório (Terwee *et al.*, 2007), foi inferior às estimativas dos outros estudos de fiabilidade teste–reteste da QBPDS, realizados apenas com amostras de utentes com DLC. A Tabela 29 exhibe os resultados desses estudos, que variaram entre 0,86 e 0,92.

**Tabela 29** – Síntese da estimativa de CCI da QBPDS –VP neste estudo e da QBPDS em estudos anteriores.

	N (DLC)	Intervalo	CCI
Schoppink <i>et al.</i> (1996)	89	1 semana	0,90
Mousavi <i>et al.</i> (2006)	31	1 dia	0,86
Melikoglu <i>et al.</i> (2009)	100	1 dia	0,92
Vieira <i>et al.</i> (2012)	104	4 a 7 dias	0,70

O resultado de fiabilidade teste reeste obtido, foi particularmente inferior ao atingido pelo estudo de Schoppink *et al.* (1996), que decorreu também com 1 semana de intervalo (ICC = 0,90). Esse estudo anterior, decorreu sem qualquer informação simultânea da perceção de melhoria dos utentes e incluiu todos os utentes para a estimativa da fiabilidade teste-reteste (n=89).

De acordo com as bases teóricas da fiabilidade, para um determinado erro, quanto maior a variância total entre as pontuações observadas, maior será a fiabilidade. Neste estudo, o procedimento adjuvante de utilizar a resposta simultânea da PGIC para decidir os indivíduos incluídos na estimativa fiabilidade teste-reteste, parece ter prejudicado ligeiramente essa estimativa. Embora, possa ter contribuído para garantir que a estimativa da fiabilidade teste-reteste decorria de uma amostra de indivíduos que mantinham a sua condição clínica estável, por outro lado, reduziu a variância total das pontuações. Assim, mesmo que as pontuações se mantivessem estáveis entre as duas medições, a diminuição da variância total, diminui a proporção da variância total, que é representada pela variância verdadeira das pontuações e aumentou a proporção da variância total que é atribuída ao erro.

Em termos conceptuais, o CCI de 0,70, significa que 70% da variação nas pontuações observadas pode ser atribuída a variação verdadeira nas pontuações. Os restantes 30% são atribuídos a erros de medição.

Em suma, a metodologia adotada, apesar de condicionar uma ligeira diminuição da fiabilidade teste-reteste, pelo menos permitiu um controlo e despiste das verdadeiras alterações clínicas, que são contraproducentes às bases teóricas da estimativa da fiabilidade teste-reteste. A estimativa da Fiabilidade teste-reteste da QBPDS obtida revela uma estabilidade do tipo temporal satisfatória.

### **5.5. Fiabilidade Teste-Reteste, EPM e DMD**

No âmbito da interpretabilidade da QBPDS, decidiu-se calcular a DMD por ser o ponto de partida para definir mudanças de pontuações ao longo do tempo, conforme está representada no referencial contínuo, proposto por Portney & Watkins (2009).

A DMD representa a quantidade de mudança que tem de ser adquirida para refletir uma mudança verdadeira. A estimativa deste limiar advém da estatística da estabilidade da resposta, sendo o EPM o mais frequentemente utilizado.

Para este fim, estimou-se a outra medida da Fiabilidade teste-reteste da QBPDS, neste caso, para um intervalo de 7 semanas e referente ao grupo estável (grupo “sem alterações”). Essa medida representa um coeficiente de correlação das pontuações da QBPDS, nesse grupo, entre a pré intervenção e a pós intervenção da Fisioterapia, necessária ao cálculo do EPM e subsequente DMD.

Antes de prosseguir com a estimativa da fiabilidade teste-reteste, verificou-se a validade das medições obtidas com a estratégia da definição do grupo “sem alterações”, tendo por base a PGIC, através de 3 métodos: a) Análise da estatística descritiva das diferenças das pontuações da QBPDS de T0 para T2, verificada para cada nível da PGIC; b) Verificação de que o grupo “sem alterações” é estável na pontuação da QBPDS entre T0 e T2; c) Verificação da equivalência dos grupos “sem alterações” e “melhoria clínica”, no ponto de partida. Assim:

a) Em primeiro lugar, examinou-se as diferenças médias das pontuações da QBPDS para os 7 níveis da percepção global de mudança (PGIC em T2) . Os resultados

obtidos (Tabela 17) seguem uma direção e magnitude de diferença média da pontuação para cada um dos 7 níveis da PGIC, que estão relativamente de acordo com as expectativas. As únicas inconsistências foram:

- Os indivíduos que se reportaram “melhores” (nível 6 na PGIC) tiveram uma diferença média de 16,2 ( $\pm$  15,9), enquanto os indivíduos que se reportaram “muito melhores” (nível 7 na PGIC) tiveram uma diferença média de apenas 10,5 ( $\pm$  13,0). Esta inconsistência deve-se provavelmente ao pequeno número de indivíduos que se autotransferiram “muito melhores” (n=10).

- Ao nível dos primeiros quatro níveis (PGIC =1,2,3 e 4), considerados para identificar, *a posteriori*, o grupo “sem alterações”, verificou-se que os indivíduos que se classificaram nos níveis 1 e 2 da PGIC, obtiveram uma diferença média das pontuações negativas. Essas diferenças negativas, de -25 e -1,4 ( $\pm$  5,6), respectivamente, obrigaram a considerar a possibilidade de um agravamento da incapacidade funcional.

Relativamente à diferença média da QBPDS dos indivíduos que se reportaram no nível 2 da PGIC, que foi de pequena magnitude (-1,4 pontos) e acompanhada de uma descrição nesse nível - “Quase na mesma, sem qualquer alteração visível”, interpretou-se como outra inconsistência dos resultados, que adveio provavelmente da pequena amostra do nível 2 (n=8). No entanto, a diferença média da QBPDS dos indivíduos que se reportaram no nível 1 da PGIC, de grande magnitude (-25 pontos) e com correspondência à descrição - “Sem alterações (ou a condição piorou)”, teve que ser prudentemente interpretada como um verdadeiro agravamento da capacidade funcional. Nesse sentido, foi averiguado o valor da remoção desse único indivíduo na estimativa da fiabilidade teste-reteste. Concluiu-se que a sua presença no grupo considerado clinicamente estável (“sem alterações”) não determinava alterações relevantes na estimativa da fiabilidade teste-reteste (T0-T2).

Destes resultados depreende-se uma limitação da PGIC, ocasionada pelo fato da escala não possuir um item que descreva a percepção de mudança global “sem alterações” isoladamente. Os utentes que se identificam no nível 1, tanto pode ser porque estão sem alterações, ou porque a condição piorou. Assim sendo, dificulta a garantia da estabilidade clínica dos utentes;

b) O segundo método definido para verificar a validade das medições, permitiu confirmar que as pontuações da QBPDS, relativas ao grupo “sem alterações”, não

sofreram alterações significativas entre o momento inicial (T0) e a reavaliação (T2). Por outro lado, os resultados revelaram que havia diferenças significativas na pontuação de QBPDS entre duas medições (T0 e T2) para o grupo “melhoria clínica”

c) Em terceiro lugar, confirmou-se equivalência dos grupos “sem alterações” e “melhoria clínica”, no ponto de partida, quer ao nível das variáveis quantitativas (QBPDS T0 e EVA T0), quer ao nível das qualitativas (variáveis sócio-demográficas e clínicas).

Neste estudo foi estimada a fiabilidade teste-reteste (T0-T2) dada pelo CCI de 0,827 (IC 95%: 0,705-0,902) e por conseguinte, foi estimado o EPM no valor de 8 pontos (IC 95%: 6 – 10,5) e a DMD<sub>90</sub>, no valor de 18,7 pontos (IC 90%: 14 – 24,5). A Tabela 30 permite a comparação dos resultados deste estudo de investigação e de todos os estudos anteriores da QBPDS em que foram estimadas as medidas em causa. Estes estudos, encontram-se organizados e identificados de acordo com a duração dos sintomas dos indivíduos que constituem a amostra.

Os resultados de DMD deste estudo resultam de um intervalo teste reteste de 7 semanas, ao longo do qual, os utentes com DLC receberam tratamento de Fisioterapia (até 6 semanas de intervenção de Fisioterapia). Esta opção está de acordo com a maioria dos outros estudos, cujo intervalo teste reteste foi longo, de 4 a 12 semanas, correspondendo precisamente a um período de intervenção de Fisioterapia (Fritz & Irrgang, 2001; Davidson & Keating, 2002; van der Roer *et al.*, 2006) ou tratamento multidisciplinar (Demoulin *et al.*, 2010). O único estudo que estabeleceu um intervalo de uma semana, distingue-se também pelo facto dos indivíduos não terem sido sujeitos a nenhuma intervenção (Hicks & Manal, 2009).

A estimativa da DMD pressupõe que seja identificado um grupo de utentes considerados clinicamente estáveis, com base na interpretação da escala avaliação global de mudança. Para tal, foram utilizadas escalas avaliação global de mudança com 6 níveis (van der Roer *et al.*, 2006), 7 níveis (Davidson & Keating, 2002; Demoulin *et al.*, 2010) ou 15 níveis (Hicks & Manal, 2009; Fritz & Irrgang, 2001), que são relativamente variáveis nas descrições de mudanças atribuídas a cada nível. Neste estudo, a escala seleccionada foi a PGIC – VP, com 7 níveis de mudança. Pelo número de níveis, trata-se de uma escala intermédia, compreendida entre uma escala parcimoniosa que não tem qualquer distinção na magnitude de mudança (por exemplo: escalas de 3 níveis) e uma escala complexa, na qual os indivíduos podem encontrar

dificuldade em interpretar e completar sem assistência (por exemplo: escalas de 15 níveis) (Kamper, Maher & Mackay, 2005).

**Tabela 30** - Síntese da estimativa de ICC, EPM e DMD da QBPDS - VP neste estudo e da QBPDS em estudos anteriores.

	<b>Intervalo</b>	<b>Grupo estável (%)</b>	<b>ICC</b>	<b>EPM</b>	<b>DMD</b>
<b>DLA</b>					
Fritz & Irrgang (2001)	4 semanas	23/67 = 0,34 (34%)	0,55 (IC 95%: 0,20- 0,78)	13, 08 (IC 95%: 10, 54 – 17,47)	DMD <sub>90</sub> = 30,52
van der Roer <i>et al.</i> (2006)	12 semanas	52/304 = 0,17 (17%)	—	—	DMD <sub>95</sub> = 32,9 (IC 95%: 24,6 – 49,8)
<b>DL sem especificidade da duração</b>					
Davidson & Keating (2002)	6 semanas	Categoria “sem alterações” = 47/ 101 =0,47 (47%)	0,84 (IC 95%: 0,73 -0,91)	8 (IC 95%: 6-10) EPM <sub>repetição</sub> = 11 (IC95%: 8,5-15)	DMD <sub>90</sub> = 19 (IC 90%: 14- 24)
		Item “Aproximadamente igual” = 16/101 = 0,16 (16%)	0,89 (IC 95%: 0,72 -0,96)	7 (IC 95%: 4-11) EPM <sub>repetição</sub> = 9 (IC 95%: 6-15)	DMD <sub>90</sub> = 15 (IC 90%: 9- 24)
Hicks & Manal, 2009	1 semana (média de 11 dias); Idade > 65 A.	56/56 =1 (100%)	0,94 (IC 95%: 0,90-0,97).	4,73	DMD <sub>90</sub> = 11,04 (IC 90%)
<b>DLC</b>					
Demoulin <i>et al.</i> (2010)	10 semanas	143/212 = 0,68 (68%)	—	5,7 (calculado: √ variabilidade intraindividual)	DMD <sub>95</sub> = 15,8 (estimado com IC 95%)
van der Roer <i>et al.</i> (2006)	12 semanas	62/138 = 0,45 (45%)	—	—	DMD <sub>95</sub> = 24,6 (IC 95%: 19,9 – 32,4)
Vieira <i>et al.</i> (2012)	7 semanas (6 semanas de Fisioterapia)	44/120 = 0,37 (37%)	0,827 (IC 95%: 0,705- 0,902)	8 (IC 95%: 6 – 10,5) EPM <sub>repetição</sub> = 11,3 (IC 95%: 8,4 – 14,9)	DMD <sub>90</sub> = 18,7 (IC 90%: 14 – 24,5)

Comparativamente às escalas de avaliação global de mudança, de 6 e 7 níveis adotadas noutros estudos da QBPDS, a PGIC distingue-se das mesmas, pela ausência de níveis que sirvam exclusivamente para avaliar um agravamento clínico (frequentemente 1 ou 2 níveis). Essa característica da PGIC parece ter sido apropriada para o presente estudo, considerando que era expectável que poucos indivíduos com DLC reportassem um agravamento clínico ao longo das 6 semanas de Fisioterapia.

Na sequência da característica anteriormente referida, nos outros estudos da QBPDS foram definidas *a priori*, 3 categorias: “agravamento clínico”; “sem alterações” e “melhoria clínica”; enquanto neste estudo, a PGIC foi apenas organizada em duas categorias “sem alterações” e “melhoria clínica”. Neste caso, os níveis da escala que representam a estabilidade clínica deixam de estar a meio da escala, para estar no seu início. Apesar dessa particularidade, a categoria “sem alterações”, respeitou o mesmo sentido adotado nas outras escalas, ou seja, inclui os níveis que expressam nenhuma ou ligeira alteração da condição clínica (PGIC =1,2,3 e 4). A partir do nível 5 (PGIC=5,6 e 7), foi interpretada como uma melhoria clínica.

Neste estudo, assim como na maioria dos estudos, utilizou-se o grupo “sem alterações”, que é constituído *a posteriori*, para a estimativa da fiabilidade teste-reteste (intervalo de 7 semanas), EPM e DMD.

O valor de CCI da QBPDS (7 semanas) estimado para a presente amostra, de 0,827 (IC 95%: 0,705-0,902), é aproximadamente igual ao CCI de 0,84 (IC 95%: 0,73 - 0,91), reportado por Davidson & Keating (2002). Embora com um desenho de estudo semelhante (6 semanas de intervenção de Fisioterapia), esse estudo decorre com uma amostra de utentes com DL sem especificidade de duração dos sintomas. Dada a semelhança das estimativas de CCI, no presente estudo obteve-se um EPM de 8 pontos, igual ao reportado por Davidson & Keating (2002).

Neste estudo, à semelhança do estudo de Davidson & Keating (2002), também se decidiu por calcular a DMD baseada num intervalo de confiança de 90% (valor  $z = 1,65$ ). Desta forma, o valor de  $DMD_{90}$  de 19 pontos, encontrado no estudo atual, é igual ao reportado por Davidson & Keating (2002).

À exceção do estudo supracitado, nenhum outro estudo determinou a  $DMD_{90}$  da QBPDS com um intervalo de 6 semanas de intervenção. As outras estimativas da  $DMD_{90}$  da QBPDS decorreram de estudos com intervalo teste-reteste diferente e

amostras diferentes. A  $DMD_{90}$  da QBPDS-VP estimada neste estudo (18,7 pontos; IC 90%: 14 – 24,5) apresenta um valor intermédio das outras estimativas de  $DMD_{90}$  da QBPDS, que variaram entre 11,04 e 30,52 pontos.

O resultado de  $DMD_{90}$  da QBPDS-VP obtido neste estudo foi inferior à estimativa de  $DMD_{90}$  da QBPDS decorrente de um estudo com um intervalo teste-reteste inferior, de 4 semanas de intervenção de Fisioterapia, restrito a utentes com DLA, de 30,52 pontos (Fritz & Irrgang, 2001). Importa no entanto referir que o valor obtido para a  $DMD_{90}$  neste estudo é consequência de uma estimativa de fiabilidade teste-reteste de valor superior (0,827; IC 95%: 0,705-0,902) à reportada por Fritz e Irrgang (2001), de 0,55 (IC 95%: 0,20- 0,78). Pondera-se que essa diferença no CCI seja reflexo de uma maior variabilidade desta amostra derivada das diferenças da amostra (duração dos sintomas), e expectável, uma vez que decorre de um intervalo teste-reteste maior. Desta forma, parece que a DLC contribui para uma maior heterogeneidade da amostra comparativamente à condição aguda da DL, condicionando um CCI superior.

Ao contrário dos resultados anteriores, a estimativa de  $DMD_{90}$  da QBPDS-VP, obtido neste estudo (18,7 pontos; IC 90%: 14 – 24,5), foi superior à estimativa de  $DMD_{90}$  da QBPDS reportada por Hicks & Manal (2009), que apresenta um valor de 11,04. O estudo em comparação decorreu com uma amostra de idosos com DL, com intervalo teste-reteste de apenas 1 semana. Nesse estudo verificaram que 100% da amostra se reportou clinicamente estável e dessa forma obtiveram um valor de CCI muito alto, de 0,94 (IC 95%: 0,90-0,97), um EPM de 4,73 e uma  $DMD_{90}$  de 11,04. Uma vez que teve por base uma amostra geriátrica, que não foi sujeita a intervenção de Fisioterapia durante o intervalo de uma semana, os seus resultados são obtidos em condições muito díspares do presente estudo.

Os restantes estudos, que estimaram a DMD da QBPDS baseada num intervalo de confiança de 95%, não apresentam os valores de Fiabilidade teste-reteste precedentes ao cálculo da  $DMD_{95}$ . Os estudos derivados de amostras de DLC encontraram, para 10 e 12 semanas de intervalo teste-reteste, respetivamente, uma  $DMD_{95}$  de 15,8 (Demoulin *et al.*, 2010) e 24,6 (IC 95%: 19,9 – 32,4) (van der Roer *et al.*, 2006).

Era esperado que a estimativa de DMD deste estudo, com um intervalo de confiança de 90% ( $DMD_{90}$ ), fosse inferior às estimativas de DMD com um intervalo de 95% ( $DMD_{95}$ ), uma vez que o primeiro decorre da multiplicação por um valor z (1,65),

inferior ao valor z tabelado para um intervalo de confiança de 95% (1,96). Esta explicação matemática pode explicar em parte a inferiorização do valor de DMD<sub>90</sub> deste estudo, em relação ao valor de DMD<sub>95</sub> de 25 pontos, obtido no estudo van der Roer *et al.* (2006), mas não permite explicar o resultado de 15,8 pontos (Demoulin *et al.*, 2010). A única explicação reside no EPM que foi calculado neste estudo (de 8 pontos) e que é superior ao reportado por Demoulin *et al.* (2010) de 5,7 pontos.

De uma forma global, o resultado de DMD<sub>90</sub> obtido neste estudo para a QBPDS -VP é igual ao valor de DMD<sub>90</sub> apurado num estudo com desenho semelhante (Davidson & Keating, 2002) mas diferente dos restantes estudos, que consideram diferentes intervalos teste-reteste e diferentes amostras. Assim, o resultado deste estudo em conjunto com a evidência disponível dos estudos anteriores da QBPDS, parece contrariar, a possibilidade teórica da DMD ser interpretada como uma propriedade/característica das medições, que se mantém constante ao longo das várias amostras, e apoiar a ideia de que “a DMD pode variar em função da fiabilidade estimada e aplicada no seu cálculo” (Portney & Watkins, 2009).

A conjugação dos resultados deste estudo com os anteriores (Tabela 30) corrobora com a teoria: “quanto mais fiável for um instrumento, mais preciso é o valor de DMD” (Portney & Watkins, 2009). Uma maior fiabilidade teste-reteste, determina um menor EPM e assim um menor valor de DMD, que significa maior precisão da medição.

No caso deste estudo, a estimativa da DMD de 19 pontos não revela muita precisão da sua medição. Esse valor pode ter sido sobreestimado pela fiabilidade teste-reteste não ser muito elevada (CCI de 0,827; IC 95%: 0,705-0,902) e pelo desvio padrão da diferença média das pontuações (para o grupo “sem alterações”) ter sido relativamente grande (19,25 pontos), determinando um maior EPM (8 pontos) e assim um maior valor de DMD (19 pontos).

Em suma, baseados nos resultados deste estudo, é necessário haver uma mudança de 19 pontos na QBPDS, para estar 90% confiante de que ocorreu uma diferença real na incapacidade funcional de utentes com DLC.

Com base no valor de DMD<sub>90</sub> determinou-se o limite superior e limite inferior da pontuação inicial da QBPDS-VP que permitem detetar diferenças na pontuação pelo

menos no valor da DMD, num segundo momento de avaliação. Assim, a amplitude válida para a QBPDS – VP é de 19 a 81 pontos.

O limite inferior da escala assume particular implicação clínica, visto que a expectativa e a evolução mais provável dos utentes, é reduzirem o nível de incapacidade funcional ao longo do tempo, ou seja, reduzirem a pontuação total da QBPDS –VP.

## 5.6. Determinação dos efeitos de “chão” e de “teto”

Segundo o critério proposto por Terwee *et al.* (2009) para a identificação de efeitos de “chão” e de “teto”<sup>13</sup> e de acordo com os resultados verificados nesta amostra, não foram apurados efeitos de “chão” nem de “teto” para a QBPDS -VP.

Identificou-se uma proporção ligeiramente superior ao critério de 15%, no limite inferior da amplitude válida, com 15,8% dos indivíduos a revelarem um a pontuação inicial demasiada baixa (<19 pontos) para permitir que fosse detetado uma melhoria clínica. No entanto, pela fraca representação do valor excedente, não se concluiu por efeito de “chão”.

Os resultados deste estudo estão de acordo com os resultados anteriormente descritos por Davidson & Keating (2002), que também concluíram pela ausência de efeitos de “chão” e de “teto. Contudo, é de referir que essas conclusões não advêm da mesma definição de grupo estável.

Enquanto na presente amostra essas conclusões resultam do grupo “sem alterações” (DMD<sub>90</sub> de 19 pontos), no estudo de Davidson & Keating (2002) utilizaram o critério de estabilidade de “aproximadamente igual (DMD<sub>90</sub> =15 pontos). Nesse caso apenas 14% dos indivíduos revelavam uma pontuação inicial demasiada baixa (<15 pontos) para permitir que fosse detetada uma melhoria clínica.

Em relação ao efeito de “teto”, os resultados deste estudo, de apenas 0,8% dos indivíduos com pontuações na QBPDS no momento inicial superiores ao limite superior da amplitude válida (>81 pontos), reforçam os resultados determinados no estudo de

---

<sup>13</sup> Terwee *et al.* (2009) estabelece um critério limite de 15% dos participantes a pontuar (no momento inicial) acima do limite superior ou abaixo do limite inferior, da amplitude válida da escala. Acima desse limite, são identificados efeitos de “chão” e de “teto” na escala.

Davidson & Keating (2002), com proporções de 1% (>85 pontos). Deste modo a evidência consolida a ausência de efeito de “teto” na QBPDS.

Por tudo o que foi referido, apesar de não ter sido identificado um efeito de “chão” da QBPDS - VP, considera-se que há uma tendência dos indivíduos com DL pontuarem totais na QBPDS mais baixos do que altos, que deve ser tida em conta na aplicação clínica da QBPDS - VP. Uma pontuação inicial da QBPDS-VP inferior à DMD (<19 pontos) exige algum cuidado na interpretação de mudanças de pontuação que apontem uma melhoria clínica (uma diminuição da pontuação total da QBPDS).

### **5.7. Estimativa do Poder de resposta**

O conceito de Poder de resposta adotado neste trabalho (de Bruin *et al.*, 1997 citado por Beaton *et al.*, 2001)<sup>14</sup>, permitiu que fossem considerados, para a Revisão da Literatura, resultados de estudos anteriores da QBPDS que utilizassem, quer métodos baseados na distribuição, quer métodos baseados na âncora para a estimativa do Poder de Resposta. Contudo, e face às recomendações mais atuais propostas pelo painel de peritos da COSMIN, deve ser dada preferência aos métodos baseados na âncora. Segundo Mokkink *et al.* (2010a), 76% do painel considera que o método “Coeficiente correlação âncora” é o preferido para estimar o Poder de resposta. Este método pode ser utilizado quando o instrumento de medida em estudo e a medida de padrão de ouro são medidas numa escala contínua; e 60% do painel, considera que o método da curva ROC (AAC) é um método apropriado para avaliar o poder de resposta quando o instrumento padrão de medida está disponível de um modo dicotómico (Mokkink *et al.*, 2010a).

Como já foi referido, este estudo, que está particularmente interessado na definição de poder de resposta enquanto capacidade de um instrumento detetar mudanças clinicamente relevantes. Para a sua estimativa procedeu-se à aplicação do método “Coeficiente correlação âncora” e o método curva ROC, pela AAC. Os resultados dessas estimativas são apresentados na Tabela 31, comparativamente com os anteriores.

---

<sup>14</sup> O poder de resposta de um instrumento inclui a capacidade de detetar mudanças reais e/ou mudanças clinicamente significativas, desde que seja com precisão.

**Tabela 31** – Síntese da estimativa de poder de resposta da QBPDS - VP neste estudo e em estudos da QBPDS anteriores.

	Amostra/ Intervalo	Poder de Resposta: Métodos baseados na âncora	
		Coefficiente correlação âncora	Curva ROC: AAC
Fritz & Irrgang (2001)	n =67 DLA; 4 semanas de Fisioterapia	Coefficiente de correlação de Pearson = 0,67	0,87 ( $\pm 0,048$ );
Davidson & Keating (2002)	n= 106 DL; (para AAC: n=99) 6 semanas	_____	0,74 (IC 95%: 0,64-0,84)
Demoulin et al. (2010)	n = 212 DLC; 10 semanas de tratamento multidisciplinar	_____	0,850, (IC 95%: 0,786 – 0,914); Curva ROC (mudanças de pontuação expressa em %): AAC = 0,856 (IC 95%: 0.791 - 0.922)
Vieira & Cruz (2012)	7 semanas (6 semanas de Fisioterapia)	Coefficiente de $\rho$ de Spearman = 0,426 (p=0,000)	0,741 (IC 95%: 0,645 – 0,837)

Neste estudo, a aplicação do método “Coeficiente de Correlação âncora”, foi efetuada através da estimativa do coeficiente de correlação de Spearman entre a diferença de pontuações da QBPDS T0-T2 e o nível da PGIC.

O coeficiente de  $\rho$  de Spearman obtido, com o valor de 0,426 (valor p=0,000), reflete uma correlação positiva, moderada e significativa (para nível de significância de 0,05) entre essas duas variáveis. Esta estimativa pode ser interpretada como um poder de resposta moderado da QBPDS – VP.

Da literatura publicada, o único estudo que recorreu de forma análoga a este método de poder de resposta, foi o de Fritz e Irrgang (2001). O coeficiente de correlação do presente estudo ( $\rho = 0,426$ ) foi ligeiramente inferior ao resultado reportado por Fritz e Irrgang (2001), que estimaram um Coeficiente de correlação de Pearson de 0,67. A diferença da correlação encontrada neste estudo em relação ao supramencionado, pode-

se justificar pelo estudo presente utilizar uma escala de percepção global de mudança com 7 níveis, comparativamente ao anterior, que utilizou uma escala de 15 níveis. Assim, a redução dos itens da escala, pode ter reduzido a correlação entre as duas variáveis.

Com recurso à curva ROC, analisou-se o poder de Resposta da QBPDS -VP com base na resposta simultânea da PGIC, disponível de um modo dicotómico (sucesso vs insucesso). Definiu-se a condição sucesso para uma percepção global de mudança, de PGIC = 5,6 e 7 e a condição insucesso para PGIC = 1,2,3 e 4. Solicitou-se a curva ROC e obteve-se uma AAC de 0,741 (IC 95%: 0,645 – 0,837). Este valor, sendo superior a 0,70, representa uma capacidade discriminativa das medições da escala adequada, de acordo com o critério de Terwee *et al.* (2007).

O valor de AAC obtido neste estudo é consistente com o valor estimado por Davidson & Keating (2002), de 0,74 (IC 95%: 0,64-0,84); e ligeiramente inferior aos valores de Fritz e Irrgang (2001), de 0,87 ( $\pm 0,048$ ) e de Demoulin *et al.* (2010), de 0,850 (IC 95%: 0,786 – 0,914).

O facto destes métodos baseados na âncora terem por base a autoavaliação dos utentes relativamente à mudança global como critério de mudança clínica significativa, induz, por si só, um fator de *viés* que pode explicar essas ligeiras diferenças nos resultados.

De uma forma global, as duas estimativas de poder de respostas efetuadas neste estudo, são coerentes entre si e com a literatura anterior, revelando que a QBPDS-VP tem um poder de resposta moderado.

## **5.8. Estimativa DMCI**

A estimativa da DMCI decorreu de dois métodos baseados na âncora: o método “Diferença média de pontuação”; e o método curva ROC, mediante o ponto ótimo de corte. Na Tabela 32, são apresentados os valores de DMCI, encontrados neste estudo para a QBPDS - VP, em comparação com outros estudos que também estimaram a DMCI da QBPDS para amostras de utentes com DLC.

**Tabela 32** – Síntese da estimativa de DMCI da QBPDS - VP neste estudo e em estudos da QBPDS anteriores.

	Intervalo	Grupo sucesso (segundo PGIC)	Diferença média de pontuação	DMCI
				Curva ROC: Ponto ótimo de corte (%sensibilidade; %especificidade)
van der Roer <i>et al.</i> (2006)	12 semanas	74/138 = 0,54 (54%)	12,7 ( $\pm$ 14,8)	DMCI = 5 (78; 77) DMCI% = 18.1% (72; 85) <u>Divisão em 2 grupos:</u> DMCI <sub>QBPDS<math>\leq</math>45</sub> = 3 (76; 75) DMCI <sub>QBPDS<math>\geq</math>46</sub> = 5 (91; 70) <u>Divisão em 3 categorias:</u> DMCI <sub>QBPDS<math>\leq</math>40</sub> = 3 (79; 78) DMCI <sub>40&lt;QBPDS&lt;50</sub> = 5 (79;76) DMCI <sub>QBPDS<math>\geq</math>50</sub> = 10 (78;83)
Demoulin et al. (2010)	10 semanas	46/212 = 0,22 (22%)	—————	DMCI = 6,5 (72; 71) DMCI% = -24,2% (71; 71) DMCI <sub>QBPDS<math>\leq</math>34</sub> = 6,5 (60; 83) DMCI <sub>QBPDS<math>\geq</math>34</sub> = 10,5 (73; 67)
Vieira & Cruz (2012)	7 semanas (6 semanas de Fisioterapia)	76/120 = 0,63 (63%)	8,05	

Com base nas indicações da literatura, assumiu-se *a priori* que as 7 semanas de intervalo (com 6 semanas de intervenção de Fisioterapia) seriam o tempo suficiente para que os utentes obtivessem melhorias significativas. Os resultados apurados no presente estudo, confirmaram essas expectativas, com uma taxa de melhoria autoreportada na PGIC (PGIC =5,6, e 7) de 63,3%, após 6 semanas de Fisioterapia.

Os outros dois estudos, de van der Roer *et al.* (2006) e Demoulin *et al.* (2010), com 10 e 12 semanas de intervalo, respetivamente, não obtiveram taxas de melhoria tão

elevadas. Van der Roer *et al.* (2006), obtiveram uma taxa de melhoria de 54% após 10 semanas de Fisioterapia e Demoulin *et al.* (2010), de 22%, após 12 semanas de tratamento multidisciplinar.

A explicação destas discrepâncias nas taxas de melhoria pode ser atribuída à escala de avaliação global de mudança. Embora, esses dois estudos, tenham utilizado uma escala de avaliação global de mudança de 6 e 7 níveis respetivamente (equiparáveis à deste estudo), a sua interpretação foi pré-definida, no sentido que apenas os 2 níveis extremos fossem interpretados como melhoria clínica, diminuindo assim a probabilidade dos indivíduos se percecionarem “muito melhores” ou “completamente recuperados”; em comparação aos 3 níveis da PGIC que foram considerados neste estudo (PGIC= 5, 6 e 7): “moderadamente melhor, com mudança ligeira mas significativa”, “melhor” e “muito melhor”.

Para calcular a DMCI, pela diferença das média das pontuações, utilizou-se o método original descrito por Jaeschke *et al.* (1989) e que tem sido utilizado na grande maioria na maioria dos estudos realizados com a QBPDS ou com outros instrumentos que medem um construto similar. Assim, subtraiu-se a diferença média do grupo de utentes que se identificou com o nível de PGIC mais baixo para ser considerado melhoria clínica (PGIC=5) pela diferença média do grupo de indivíduos “sem alterações”. Obteve-se um valor de DMCI da QBPDS – VP de 8 pontos.

No entanto, outros autores argumentam que a diferença média das pontuações do grupo de utentes que se identificou com o nível de PGIC mais baixo para ser considerado melhoria clínica, não deve ser subtraída (Hays *et al.* 2005 citado por Terwee *et al.*, 2009; van der Roer *et al.*, 2006). Desta forma, van der Roer *et al.* (2006), estimaram a DMCI apenas pela diferença média das pontuações do grupo de utentes que se identificou com o nível de escala de avaliação global de mudança, mais baixo possível para ser considerado melhoria clínica, neste caso o nível 2 “muito melhor” (de uma escala de 6 níveis), no valor de 12,7 ( $\pm 14,8$ ).

O resultado de DMCI, estimado no presente trabalho, 8 pontos, foi inferior à estimativa de DMCI de van der Roer *et al.* (2006), de 13 pontos, conforme já era esperado, dada a diferença de cálculo entre os dois estudos.

Pelo método curva ROC, obteve-se um valor ótimo de corte de 6,5 pontos, com valores de sensibilidade e especificidade equilibrados, de 72% e 71% respetivamente.

Este valor ótimo de corte, apresenta-se como um valor de intermédio em relação aos que foram verificados, em estudos anteriores pelo mesmo método, que oscilaram entre 5 (Demoulin *et al.*, 2010) e 8,5 pontos (van der Roer *et al.*, 2006). A presente estimativa do ponto ótimo de corte, corrobora com ambos os valores de sensibilidade e especificidade destes estudos anteriores (que também rondam os 70%) e emerge de um modelo com AAC= 0,741 (IC 95%: 0,645 – 0,837), cujo valor é considerado adequado<sup>15</sup>.

Para complementar os resultados deste estudo, aplicou-se a curva ROC com base nas diferenças relativas da QBPDS entre T0 e T2. Obteve-se um ponto ótimo de corte, para uma diferença percentual de -24,2% em relação à pontuação inicial da QBPDS (sensibilidade de 71% e especificidade de 71%). O valor negativo do ponto ótimo de corte indica o sentido da diferença relativa da QBPDS entre T0 e T2, que neste caso representa uma diminuição das pontuações. Este facto não foi evidenciado nos outros estudos, que apresentam a diferença percentual em valor absoluto. O ponto ótimo encontrado no presente estudo, é superior (em valor absoluto) ao estimado por Demoulin *et al.* (2010), cujo ponto ótimo de corte, expresso em percentagem, foi uma diferença de 18,1% da pontuação total da QBPDS, em relação ao momento inicial (sensibilidade de 72% e especificidade de 85%). No entanto, o resultado de -24,2% é inferior (em valor absoluto) à DMCI recomendada pelo painel de peritos, de 30% (Ostelo *et al.*, 2008).

Em termos clínicos e em função dos resultados obtidos nesta amostra, a utilização desta DMCI, expressa em percentagem (-24%), representa uma alternativa aceitável ao valor de DMCI de 6,5 pontos, face aos valores de sensibilidade e especificidade igualmente elevados, e face à AAC. A capacidade discriminativa do modelo subjacente à estimativa do valor percentual de DMCI, traduzida pela AAC de 0,737 (IC 95%: 0,641 – 0,833), revela-se igualmente adequada<sup>15</sup>. Embora este valor de AAC seja ligeiramente inferior ao do modelo anterior (do ponto ótimo de corte 6,5), que foi de 0,741 (IC 95%: 0,645 – 0,837).

Em relação à análise da sensibilidade da DMCI, optou-se por dividir a QBPDS em duas categorias, de acordo com a mediana encontrada nesta amostra (QBPDS <34 e QBPDS ≥ 34). Para o grupo com pontuações iniciais baixas, foi identificado um ponto

---

<sup>15</sup> Terwee *et al.* (2007) considera que uma AAC, de pelo menos 0,70 é adequada.

ótimo de corte igual ao estimado para a amostra global, de 6,5, mas com sensibilidade de 60%, e especificidade de 83%. Porém, o grupo das pontuações iniciais altas, revelou um ponto ótimo de corte de 10,5 (sensibilidade de 73% e especificidade de 67%), sendo deste modo superior ao estimado na amostra global (n=120).

O ponto ótimo de corte para pontuações iniciais baixas (QBPDS <34), embora seja o mesmo que o obtido para a amostra global (6,5 pontos), está associado a uma sensibilidade mais baixa (60%), em comparação aos 72% do ponto ótimo de corte “global”. Essa sensibilidade de 60% significa que, 40% dos utentes reportam uma melhoria (PGIC=5,6 e 7) tem uma redução na pontuação na QBPDS inferior aos 6,5 pontos. Por outro lado, a especificidade de 72% pode ser interpretada que, 28% dos utentes que se reportam “sem alterações” (PGIC =1,2,3 e 4) apresentam uma diminuição da QBPDS superior a 6,5 pontos.

Em geral, estes resultados da sensibilidade da DMCI estão de acordo, com a tendência anteriormente reportada por van der Roer *et al.* (2006) e Demoulin *et al.* (2010). Confirma-se que indivíduos com pontuações iniciais mais altas identificam uma DMCI de maior valor absoluto. Os resultados deste estudo consolidam que a DMCI é dependente das pontuações iniciais da QBPDS, conforme já tinha sido encontrado por van der Roer *et al.* (2006) e Demoulin *et al.* (2010).

Os resultados mostram ainda, que a DMCI é dependente do método utilizado. Neste estudo, os valores de DMCI obtidos, que variam entre 7 e 8 pontos, demonstraram-se relativamente próximos. As estimativas de DMCI deste estudo, são muito inferiores às recomendações propostas pelo painel de peritos (Ostelo *et al.*, 2008), cujo valor recomendado é de 20 pontos para a QBPDS.

Uma possível explicação para que as estimativas da DMCI deste estudo sejam inferiores à convencionada pelo painel de peritos, pode residir na população alvo. Este estudo incide unicamente em pacientes com DLC, enquanto as recomendações estavam particularmente focadas nos cuidados de saúde primários (sem possibilidade de especificarem a duração dos sintomas da DL). Dado que os utentes com longa duração dos sintomas não esperam muita mudança, pequenas mudanças são importantes para eles (Demoulin *et al.*, 2010). Este fato foi reportado anteriormente por van der Roer *et al.* (2006), que encontrou valores de DMCI mais baixos em indivíduos com DLC do que em utentes com DLA.

Comparando os dois métodos baseados na âncora utilizados para estimar DMCI, atribui-se maior significado aos resultados extrapolados a partir do método curva ROC (DMCI de 7 pontos) em detrimento dos resultados derivados do método “Diferença média de pontuação” (DMCI de 8 pontos), porque os resultados deste último método (pela diferença das médias) são muito arbitrários e facilmente influenciados por *outliers*. Ao passo que, a curva ROC tem em conta a sensibilidade e especificidade das estimativas da DMCI, sendo assim mais transparente quanto aos resultados que se podem obter, quando se considera determinada DMCI.

Na aplicação clínica destes valores, recomenda-se a utilização preferencial da DMCI, segundo o ponto ótimo de corte com a maior sensibilidade e que advém do modelo com maior poder discriminativo. Neste caso, opta-se pela DMCI de 6,5 pontos (com sensibilidade de 72% e AAC=0,741), em detrimento da DMCI, expressa em percentagem, de -24,2% (com sensibilidade de 71% e AAC=0,737). Esta decisão tem em consideração as características dos utentes com uma longa história de DL, para os quais uma pequena melhoria pode se interpretada como relevante. Outro argumento a favor da DMCI de 6,5 pontos, pode ser a sua facilidade de interpretação.

### **5.9. Relação DMD e DMCI**

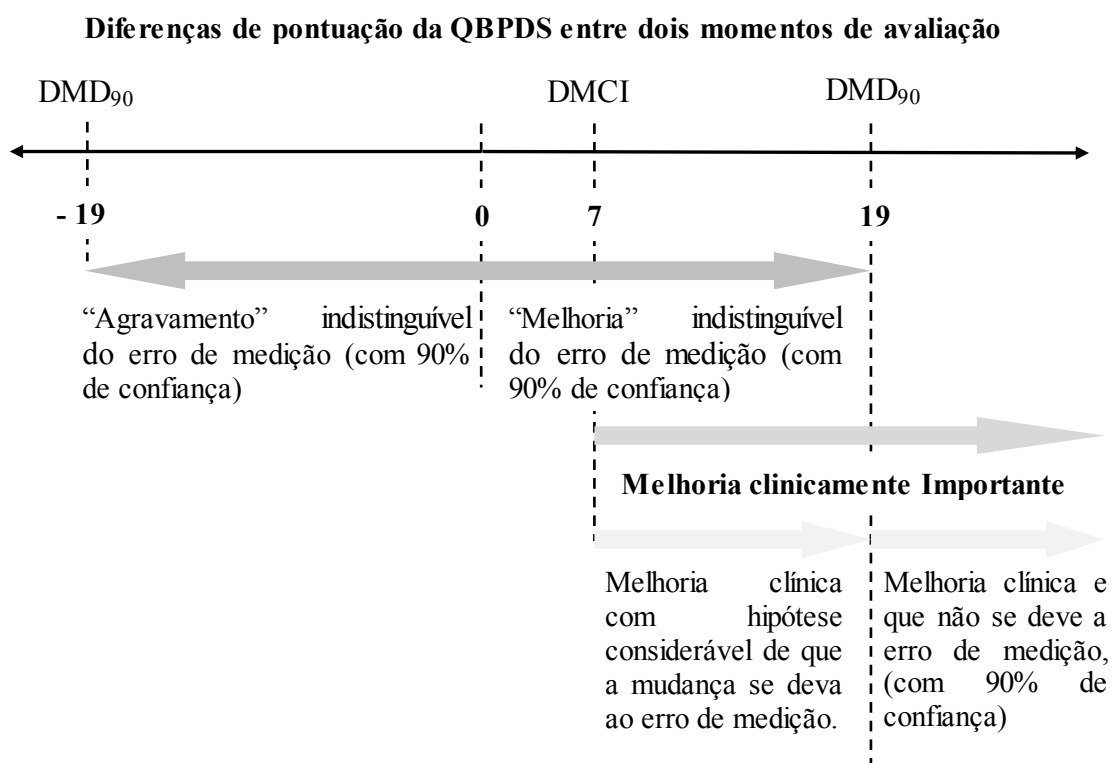
O resumo dos resultados de interpretabilidade da QBPDS- VP derivados deste estudo, é apresentado na Figura 3.

A DMD pode ser considerada uma estimativa conservadora do progresso do utente, que identifica a menor quantidade de mudança que pode ser interpretada como alguma melhoria ou agravamento real. Dessa forma, a DMD obtida para a QBPDS – VP, de 19 pontos, é representada no referencial da Figura 3, quer para diferenças na pontuação da QBPDS positivas (de melhoria de incapacidade funcional), quer para diferenças da QBPDS negativas (de agravamento da incapacidade funcional).

A utilização da DMD como critério de melhoria, deve ser interpretada pela sua elevada especificidade (evitando falsos positivos), mas baixa sensibilidade (identificando muitos falsos negativos) (Portney & Watkins, 2009). Assim a DMD deve ser interpretada preferencialmente para identificar mudanças inferiores a esse valor, porque expressa uma confiança de 90%, de que diferenças de pontuação da QBPDS que

não atingem esse limiar podem ser unicamente reflexo de erro de medição; do que propriamente para considerar mudanças acima do seu valor como verdadeiras.

Neste estudo, a definição do clinicamente importante foi associada à descrição de “melhoria clínica”. A DMCI, de 7 pontos, pela definição anterior, só pode ser interpretada como valor positivo de diferença da pontuação da QBPDS -VP, que representa uma melhoria clínica, tal como é representado na Figura 3.



**Figura 13** – Interpretação de mudanças de Pontuação da QBPDS-VP com base na DMD e DMCI estimadas

## 6. CONCLUSÃO

A DLC é uma condição músculo-esquelética que continua a ser um desafio clínico para os fisioterapeutas, médicos e outros profissionais de saúde. A literatura revela que a medição da incapacidade funcional tem merecido atenção pela comunidade científica internacional da DLC.

Em Portugal, são poucos os estudos relativos à validação de instrumentos de medida da incapacidade funcional, e não se conhecem estudos divulgados ou publicados ao nível do poder de resposta de instrumentos de medida da incapacidade funcional. Este trabalho de investigação representa um estudo pioneiro em Portugal relativamente à estimativa do poder de resposta de um desses instrumentos de incapacidade funcional, a QBPDS - VP.

O trabalho concretizou-se através de um estudo metodológico, multicentro, baseado num *coorte* prospetivo de utentes com DLC, com dois objetivos principais. O primeiro, consistiu em averiguar a unidimensionalidade da QBPDS – VP e a fiabilidade das suas medições. Os resultados da análise fatorial confirmam a unidimensionalidade da QBPDS- VP e reforçam a aceitação da única versão disponível da QBPDS de 20 itens, enquanto medida de uma única variável latente. Essa variável é, presumivelmente, o nível incapacidade funcional de utentes com DLC.

O estudo da fiabilidade englobou a estimativa da consistência interna e da fiabilidade teste-reteste. Os valores estimados corroboram com os resultados publicados na literatura referente a outras versões linguísticas, revelando uma excelente consistência interna ( $\alpha$  de Cronbach= 0,95), sem redundância dos itens, e uma fiabilidade do tipo estabilidade temporal satisfatória (CCI= 0,696; IC 95%: 0,581–0,783).

O segundo objetivo deste trabalho de investigação incidiu no estudo do Poder de resposta da QBPDS- VP e na interpretabilidade das suas medições quando aplicada a utentes com DLC. Os resultados obtidos através de métodos baseados na âncora, demonstraram que a QBPDS - VP tem um poder de resposta moderado ( $\rho = 0,426$  e AAC= 0,741; IC 95%: 0,645 – 0,837).

A estimativa de DMD<sub>90</sub> da QBPDS –VP de 19 pontos determinou uma

amplitude válida para a QBPDS – VP de 19 a 81 pontos. Não foram identificados efeitos de “chão” nem de “teto” na escala.

As estimativas da DMCI, da QBPDS–VP, quando aplicada em utentes com DLC que recebem tratamentos de Fisioterapia, variaram entre 7 (pelo método curva ROC) e 8 pontos (pelo método “Diferença média de pontuação”). Foi concluída a preferência da estimativa de 7 pontos (AAC=0,741, sensibilidade de 72%, e especificidade de 71%) obtida pelo método curva ROC.

Uma análise complementar da curva ROC, baseada nas mudanças das pontuações da QBPDS – VP, expressas em valor relativo (em relação à pontuação inicial de QBPDS), revelou um ponto ótimo de corte de -24% (AAC=0,737, sensibilidade de 71%, e especificidade de 71%).

Este estudo confirma ainda a influência das pontuações iniciais da QBPDS-VP na estimativa de DMCI. Os indivíduos com pontuação inicial da QBPDS – VP mais alta ( $\geq 34$  pontos), revelam um ponto ótimo de corte de 10,5 pontos (AAC=0,738, sensibilidade de 73%, e especificidade de 67%).

### **6.1. Limitações do estudo**

Os resultados anteriormente apresentados, devem ser interpretados cuidadosamente e de acordo com algumas limitações, principalmente no que diz respeito à sua validade externa. Os aspetos que podem ter prejudicado a validade externa prendem-se com o método de amostragem utilizado (amostragem não probabilística e seleção por conveniência), que podem ter ocasionado *viés* de seleção, e com os instrumentos de autorrelato, que potenciam *viés* de medição. Na administração deste tipo de instrumentos é necessário ter em conta a influência da desajustabilidade social dos indivíduos, aliada à cultura onde estão integrados, que leva os indivíduos a responderem aos questionários de uma forma aceitável em termos sociais, transmitindo a imagem que pretendem, podendo não corresponder à realidade percebida pelos mesmos.

Também são identificadas limitações ao nível dos procedimentos de recolha de dados, nomeadamente a participação de diferentes colaboradores (que podem ter dado

maior ou menor esclarecimento aos utentes sobre o preenchimento dos questionários) e a recolha em contextos diferentes, que podem ter promovido *viés* de medição.

Outra possível limitação é uma das maiores dificuldades deste estudo, no âmbito da interpretação de mudanças de pontuação da QBPDS-VP, prende-se com a âncora externa. Na falta de “medida de padrão de ouro” para instrumentos de medição da incapacidade funcional, usou-se uma escala de avaliação global de mudança, a PGIC-VP. Apesar deste tipo de escala ser alvo de críticas, por medir um construto global de mudança e complexo, e não, exclusivamente, o construto em causa (incapacidade funcional), constitui, na atualidade, o melhor instrumento de avaliação que pode ser considerado para âncora externa no estudo de poder de resposta de instrumentos de autorrelato.

A “fragilidade” do estudo, identificada no parágrafo anterior, não foi esquecida aquando da realização do trabalho de investigação. Ao longo do trabalho recorreram-se a estratégias que tentaram contribuir para a validação da dicotomização da amostra em dois grupos, o grupo “sem alterações” e o grupo “melhoria clínica” (da incapacidade funcional), tendo por base uma escala de avaliação global de mudança. Apesar do esforço e bons resultados obtidos com essas estratégias, não se pode garantir, que a nível individual, os utentes se mantiveram estáveis ou melhoraram ao nível da incapacidade funcional.

As restantes limitações deste estudo são atribuídas a particularidades da escala de PGIC:

a) Esta escala, que solicita aos indivíduos uma avaliação da mudança, é alvo de crítica pelo fato dos indivíduos não serem capazes de recordar com precisão os estados de saúde anteriores. Os utentes sustentam a sua perceção de mudança, preferencialmente, em função do seu estado atual, em detrimento da medição da transição, como era pretendida (Kamper, Maher & Mackay, 2005). Este facto também pode produzir *viés* de medição.

b) A atribuição de pontos de corte arbitrários para delinear importância clínica também é sujeita a crítica. O fato da divisão dos itens por cada categoria (os itens da escala que representam “melhoria” e os que representam “sem melhoria”) ser arbitrária, podendo variar de amostra para amostra (Hurst & Bolton, 2004), também pode constituir um *viés* de medição.

Para efeitos deste trabalho, assumiu-se uma divisão, fundada no significado dos itens, em que os 4 primeiros níveis da escala (PGIC =1,2,3 e 4) representavam “sem alterações” e os restantes 3 níveis (PGIC=5,6,e 7) representam “melhoria clínica”. Decorrente dessa opção, determinaram-se curvas ROC com AAC moderada a elevada (valores compreendidos entre 0,736 e 0,741) e pontos ótimos de corte, com valores de sensibilidade e especificidade moderados (à volta dos 70%), que vieram a consolidar as escolhas feitas.

Por último, é necessário esclarecer que o poder de resposta não é um atributo uniforme da escala que possa ser medido por um único resultado. Assim, o estudo do poder de resposta descrito neste trabalho valida uma aplicação em particular e não o instrumento. De forma semelhante, a DMCI estimada para um instrumento de autorrelato de resultados não é uma característica imutável e não pode ser válida para todas as aplicações em estudo (Revicki, Hays, Cella & Sloan, 2008). A DMCI pode ser influenciada pela população (por exemplo, pela gravidade da condição), pelo contexto e pelos tratamentos. Desta forma, os valores apresentados neste estudo não são uma resposta final ao problema, mas uma orientação prática e um ponto de partida para investigações futuras na QBPDS - VP.

## **6.2. Implicações do estudo**

Apesar das limitações apresentadas, perspetiva-se este estudo como um contributo importante para a investigação realizada na área da DLC em Portugal. Os resultados acrescentam valor à escala QBPDS-VP, que se julga fiável e com poder de resposta moderado na monitorização da incapacidade funcional dos utentes com DLC.

As medidas de interpretabilidade da QBPDS-VP estimadas alargam as possibilidades de ser realizada investigação em Portugal na área de DLC. Até ao momento, e do nosso conhecimento, não existia nenhuma escala validada de avaliação da incapacidade funcional de utentes com DLC, para a qual fossem conhecidas estimativas das medidas de interpretabilidade (DMD e DMCI).

O conhecimento adquirido com este trabalho, das estimativas de DMD e DMCI da QBPDS- VP, aumenta o potencial da aplicabilidade clínica da escala. As mudanças de pontuação da QBPDS- VP ganham maior interpretabilidade, consegue-se extrair

mais conclusões com a aplicação da escala e monitorizar mais eficazmente a incapacidade funcional dos utentes com DLC ao longo do processo de reabilitação.

A amplitude válida da QBPDS – VP concluída neste trabalho determina que é necessária cautela quando os indivíduos reportam pontuações iniciais da QBPDS inferiores a 19 pontos, porque a escala pode não ter possibilidade de detetar mudanças que excedam o erro de medição.

As estimativas de DMCI, por sua vez, revelam que indivíduos com pontuações iniciais da QBPDS altas ( $\geq 34$  pontos), requerem valor de DMCI maior (de 11 pontos).

### **6.3. Recomendações para estudos futuros**

A direção dos estudos futuros deve ter como perspetiva e motivação geral, contribuir para a efetividade das intervenções de Fisioterapia dirigidas a utentes com DLC (principalmente com relevância clínica para o utente), e consolidar, cada vez mais, a evidência dessas intervenções. Este trabalho de investigação, dedicado ao estudo das propriedades psicométricas da QBPDS – VP, constitui um “passo” na direção pretendida. As conclusões deste trabalho, disponibilizam uma escala de avaliação da incapacidade funcional e uma interpretação clínica dos resultados obtidos com a mesma, permitindo conhecer e reportar, a evolução do nível da incapacidade funcional de utentes com DLC.

Sendo este o primeiro estudo relativo ao poder de resposta de instrumentos de medida da incapacidade funcional em Portugal, coloca-se a dúvida: Qual o poder de resposta de outros instrumentos que medem o mesmo construto da QBPDS- VP, por exemplo, a versão Portuguesa da RMQ? Em investigações futuras, será benéfico comparar os índices de poder de resposta destes dois instrumentos de modo a identificar o instrumento que revela melhor poder de resposta.

Outra incerteza suscitada aquando a elaboração deste estudo, e que carece de ser esclarecida, prende-se com a identificação do construto que mais se relaciona com a melhoria clínica de utentes com DLC. Pode colocar-se a questão: A decisão de melhoria, reportada pelos indivíduos, está mais relacionada com a diminuição da intensidade da dor ou com a diminuição do nível de incapacidade funcional? No sentido

de esclarecer esta questão, em estudo futuros, os índices de poder de resposta de instrumentos de medida da intensidade da dor (tais como a EVA ou Escala numérica), devem ser comparados com os índices de poder de resposta dos instrumentos da incapacidade funcional. Desta forma, poder-se-á averiguar qual das medidas têm maior capacidade de discriminar os indivíduos que melhoraram e os que não melhoraram.

A estimativa da DMCI da QBPDS – VP, obtida neste estudo, para utentes com DLC que recebem tratamentos de Fisioterapia, pode constituir, em estudos futuros de prognóstico, uma forma de dicotomização da amostra (utentes que melhoraram clinicamente o nível de incapacidades funcional *vs* os que não melhoraram). A partir desses estudos, podem ser identificados fatores (sócio-demográficos e clínicos) que determinam bons ou maus resultados da Fisioterapia, ao nível da incapacidade funcional.

Por outro lado, através do resultado da estimativa de DMCI, e uma vez identificados fatores com valor preditivo de prognóstico, poder-se-á prosseguir com o estudo de intervenções específicas da Fisioterapia. Sugere-se que nesses estudos futuros seja verificada a interação da intervenção específica na relação do fator de prognóstico com o resultado de Fisioterapia (que nesse caso será o nível de incapacidade funcional dos utentes com DLC), no sentido de verificar se determinada intervenção potencia, ou não, essa relação. O objetivo final é adequar as intervenções específicas de Fisioterapia a subgrupos clínicos de DLC, que beneficiam mais com determinada intervenção.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Abrams D., Davidson, M., Harrick, J., et al. (2006). Monitoring the change: Current trends in outcome measurement usage in physiotherapy. *Manual Therapy*, 11, 46–53.
- Aguiar, P. (2007). *Guia Prático de Estatística em Investigação Epidemiológica: SPSS*. Lisboa: Climepsi.
- Airaksinen, O., Brox, J. I., Cedraschi, C., et al. (2006). European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*, 15 (2), S192–S300.
- Apeldoorn, A. T., Ostelo, R. W., van Helvoirt, H., et al. (2010). The cost-effectiveness of a treatment-based classification System for low back pain: design of a randomised controlled trial and economic evaluation. *BMC Musculoskeletal Disorders*, 11, 58.
- Beaton, D. E., Bombardier, C., Katz, J. N., & Wright, J. G. (2001). A taxonomy for responsiveness. *Journal of Clinical Epidemiology*, 54, 1204–1217.
- Bekkering, G. E., Engers, A. J., Wensing, M., et al. (2003). Development of an implementation strategy for physiotherapy guidelines on low back pain. *Australian Journal of Physiotherapy*, 49, 208-214.
- Beneddouch, I., Rostom, S., Bahiri, R., et al. (2011). Traduction, adaptation transculturelle et validation de la version marocaine de la Quebec Back Pain Disability Scale. *Annals of Physical and Rehabilitation Medicine*, 54, e263-e264.
- Bicer, A., Yazici, A., Camdeviren, H., et al. (2005). Assessment of pain and disability in patients with chronic low back pain: reliability and construct validity of the

Turkish version of the Quebec Back Pain Disability Scale and Pain Disability Index. *Journal of Back and Musculoskeletal Rehabilitation*, 18, 37- 44.

Burton, A. K., Balagué, F., Cardon, G., *et al.* (2004). European guidelines for prevention in low back pain. Consultado a 25 de outubro, 2011, através de [http://www.backpaineurope.org/web/files/WG3\\_Guidelines.pdf](http://www.backpaineurope.org/web/files/WG3_Guidelines.pdf)

Cassidy D, Carroll L, Cote P. The Saskatchewan Health and Back Pain Survey - The prevalence of low back pain and related disability in Saskatchewan adults. *Spine*, 23, 1860–1867.

Castro-Lopes, J., Saramago, P., Romão, J., & Paiva, M. (2010). Pain Proposal – A Dor Crónica em Portugal. Consultado a 10 de outubro, 2011, através de [https://www.pfizer.pt/Files/Billeder/Pfizer%20P%C3%BAblico/Not%C3%ADcias/Portugal\\_Country%20Snapshot.pdf](https://www.pfizer.pt/Files/Billeder/Pfizer%20P%C3%BAblico/Not%C3%ADcias/Portugal_Country%20Snapshot.pdf)

Cella, D., Hahn, E. A., & Dineen, K. (2002). Meaningful change in cancer-specific quality of life scores: Differences between improvement and worsening. *Qual Life Res* (11), 207-221.

Chaory, K., Rannou, F., Fermanian, J., *et al.* (2004). Impact de programmes de restauration fonctionnelle sur les peurs, croyances et conduites d'évitement du lombalgique chronique. *Annales de réadaptation et de médecine physique*, 47, 93 – 97.

Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*, 119 (2), 166.e7-166.e16.

Davidson, M., & Keating, J. L. (2002). A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther*, 82 (1), 8-24.

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.

Delaune, M. F., & Bemis-Dougherty, A. (2007). Documentation in Physical Therapy Services. *Magazine of Physical Therapy*, February, 82-84.

Demoulin, C., Ostelo, R., Knottnerus, J. A., & Smeets, R. J. E. M. (2010). Quebec Back Pain Disability Scale was responsive and showed reasonable interpretability after a multidisciplinary treatment. *Journal of Clinical Epidemiology*, 63, 1249-1255.

Domingues, L., & Cruz, E. (2011). Adaptação Cultural e Contributo para a Validação da Escala *Patient Global Impression of Change*. *Ifisionline*, 2(1), 31-37.

Du Bois M & Donceel P. (2008) A screening questionnaire to predict no return to work within 3 months for low back pain claimants. *European Spine Journal*, 17, 380-385.

Farrar, J. T., Young, J. P., Jr., La Moreaux, L., *et al.* (2001). Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*, 94(2), 149-158.

Ferreira, R. Z., Ferreira, P. H., Franco, M. R., *et al.* (2011). The effect of lumbar posture on abdominal muscle thickness during an isometric leg task in people with and without non-specific chronic low back pain. *Manual Therapy*, 16 (6), 578-84.

Finch, E., Brooks, D., Stratford, P. W., & Mayo, N. (2002). *Physical Rehabilitation Outcome Measures – A Guide to Enhanced Clinical Decision Making* (2<sup>a</sup> Ed.). Baltimore: MD: Lippincott, Williams & Wilkins.

Freburger J. K., Holmes, G. M., Agans, R. P., *et al.* (2009). The rising prevalence of chronic low back pain. *Arch Intern Med*, 169 (3), 251-258.

- Fritz, J. M., & Irrgang, J. J. (2001). A comparison of a Modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy*, 81 (2), 776-788.
- Garshasbi, A., & Faghhi Zadeh, S. (2005). The effect of exercise on the intensity of low back pain in pregnant women. *Int J Gynaecol Obstet*, 88(3), 271-275.
- Haley, S. M., & Fragala-Pinkham, M. A. (2006). Interpreting Change Scores of Tests and Measures Used in Physical Therapy. *Physical Therapy*, 86 (5), 735-743.
- Hicks, G. E. & Manal, T. J. (2009). Psychometric Properties of Commonly Used Low Back Disability Questionnaires: Are They Useful for Older Adults with Low Back Pain? *Pain Medicine*, 10 (1), 85-94.
- Hill, M. M. & Hill, A. (2005). *Investigação por Questionário (2.<sup>a</sup> Edição)*. Lisboa: Edições Sílabo.
- Hurst, H. & Bolton, J. (2004). Assessing the clinical significance of change scores recorded on subjective outcome measures. *Journal of Manipulative and Physiological Therapeutics*, 27 (1), 26-35.
- Instituto Nacional de Estatística (2011a). *Classificação Portuguesa das Profissões (Edição 2011)*. Lisboa: Instituto Nacional de Estatística, IP.
- Instituto Nacional de Estatística (2011b). *Censos 2011 – Resultados Provisórios. (Edição 2011)*. Lisboa: Instituto Nacional de Estatística, IP.
- Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*, 10(4), 407-15.
- Kamper, S. J., Maher, C. G., & Mackay, G. (2005). Global Rating of Change Scales: A Review of Strengths and Weaknesses and Considerations for Design. *The Journal of Manual & Manipulative Therapy*, 17 (3), 163-170.

- Kendall, N. A., Linton, S. J., & Main, C. J. (1997). Guide to assessing psychosocial yellow flags in acute low back: risk factors for long-term disability and work loss. Consultado a 10 de janeiro, 2012, através de <http://www.cebp.nl/media/m24.pdf>
- Koes, B. W., van Tulder, M., Lin, C. C., *et al.* (2010). An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*, 19, 2075–2094.
- Kopec, J. A., Esdaile, J. M., Abrahamowicz, M., *et al.* (1995). The Quebec Back Pain Disability Scale: Measurement Properties. *Spine*, 20 (3), 341-352.
- Kopec, J. A., Esdaile, J. M., Abrahamowicz, M., *et al.* (1996). The Quebec Back Pain Disability Scale: Conceptualization and Development. *J Clin Epidemiol*, 49 (2), 151-161.
- Kovacs, F. M., Abaira, V., Royuela, A., *et al.* (2007). Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. *Spine*, 32, 2915 – 2920.
- Krismer, M., & van Tulder, M. (2007). Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). *Best Practice & Research Clinical Rheumatology*, 21(1), 77-91.
- Lauridsen, H. H., Hartvigsen, J., Manniche, *et al.* (2006). Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskeletal Disorders*, 7, 82.
- Lin, C. C., Haas, M., Maher, C. G., *et al.* (2011). Cost-effectiveness of guideline-endorsed treatments for low back pain: a systematic review. *Eur Spine J*, versão eletrônica: DOI 10.1007/s00586-010-1676-3.

- Manchikanti, L., Singh, V., Datta, S., Cohen, S. P., & Hirsch, J. A. (2009). Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician*, 12, E35-70.
- Marôco, J. (2011). *Análise Estatística com o SPSS Statistics* (5.<sup>a</sup> Edição). Pero Pinheiro: Report Number.
- Mayer, J. M., Haldeman, S., Tricco, A. C., & Dagenais, S. (2010). Management of Chronic Low Back Pain in Active Individuals. *Current Sports Medicine Reports*, 9 (1), 60-66.
- Melikoglu, M. A., Kocabas, H., Sezer, I., et al. (2009). Validation of the Turkish version of the Quebec back pain disability scale for patients with low back pain. *Spine*, 34 (6), E219-24.
- Melloh, M., Elfering, A., Egli Presland, C., et al. (2009). Identification of prognostic factors for chronicity in patients with low back pain: a review of screening instruments. *International Orthopaedics (SICOT)*, 33, 301–313.
- Merskey, H., & Bogduk, N. (1994). Classification of Chronic Pain - Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms (2<sup>nd</sup> Edition). IASP Press.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., et al. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*, 19, 539–549.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., et al. (2010b). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *Medical Research Methodology*, 10: 22.

- Monteiro, J., Faisca, L., Nunes, O., & Hipólito, J. (2010). Questionário de incapacidade de Roland Morris: Adaptação e Validação para doentes de língua portuguesa com lombalgia. *Acta Med Port*, 23, 761-766.
- Mousavi, S. J., Parnianpour, M., Mehdian, H., *et al.* (2006). The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec Back Pain Disability Scale: translation and validation studies of the Iranian versions. *Spine*, 31 (14), E454 - 9.
- Muller, U., Roder, C., & Greenough, C. G. (2006). Back related outcome assessment instruments. *European spine journal*, 15, Suppl 1, S25-31.
- Nunes, F., Ribeiro, A., & Cruz, E. (2005). Adaptação Cultural e Validação da Quebec Back Pain Disability Scale para a População Portuguesa, com Dor Lombar Crónica. Projeto de Investigação III. Escola Superior de Saúde – Instituto Politécnico de Setúbal. Artigo não publicado.
- Ostelo, R. W. J. G., Deyo, R.A., Stratford, P., *et al.* (2008). Interpreting Change Scores for Pain and Functional Status in Low Back Pain - Towards International Consensus Regarding Minimal Important Change. *Spine*, 3(1), 90–94.
- Portney, L. G., & Watkins, M. P. (2009). *Foundations of Clinical Research: Applications to Practice* (3rd Ed). Upper Saddle River: Prentice Hall.
- Rabiais, S., Nogueira, P. J. & Falcão, J. M. (2003). A Dor Na População Portuguesa - Alguns Aspetos Epidemiológicos. Ecos. Relatório de 2002 do Observatório Nacional de Saúde. Consultado a 10 de outubro, 2011, através de <http://doentescomcancro.org/uhdc/pdfs/EstudoDorPopulacaoPortuguesa.pdf>
- Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, 61, 102-109.

- Rocchi, M. B., Sisti, D., Benedetti, P., *et al.* (2005) Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. *Eura Medicophysica*, 41 (4), 275-81.
- Rodrigues, M. F., Michel-Crosato, E., Cardoso, J. R., & Traebert, J. (2009). Psychometric properties and cross-cultural adaptation of the Brazilian Quebec Back Pain Disability Scale Questionnaire. *Spine*, 34 (13), E459–64.
- Schoppink, L. E ., van Tulder, M. W., Koes, B. W., Beurskens, S. A., & de Bie, R. A. (1996). Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Physical Therapy*, 76 (3), 268 -275.
- Slade, S. C., & Keating, J. L. (2009). Effects of preferred-exercise prescription compared to usual exercise prescription on outcomes for people with non-specific low back pain: a randomized controlled trial. *BMC Musculoskeletal Disorders*, 10, 14.
- Smeets, R. J., Vlaeyen, J. W., Hidding, A., *et al.* (2006). Active rehabilitation for chronic low back pain: cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. *BMC Musculoskelet Disord*, 7, 5.
- Smeets, R., Koke, A., Lin, C., Ferreira, M., & Demoulin, C. (2011). Measures of Function in Low Back Pain/Disorders. *Arthritis Care & Research*, 63 (S11), S158-S173.
- Soler, T., & Calderón, C. (2000). The prevalence of spondylolysis in the Spanish elite athlete. *Am J Sports Med*, 28 (1), 57-62.
- Stanton, T. R., Latimer, J., Maher, C.G., & Hancock, M. J. (2010). How do we define the condition “recurrent low back pain”? A systematic review. *Eur Spine J*, 19, 533–539.

- Steffen, T., & Seney, M. (2008). Test-Retest Reliability and Minimal Detectable Change on Balance and Ambulation Tests, the 36-Item Short- Form Health Survey, and the Unified Parkinson Disease Rating Scale in People With Parkinsonism. *Physical Therapy*, 88 (6), 733 – 746.
- Straford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the Minimum Level of Detectable Change for the Roland-Morris Questionnaire. *Physical Therapy*, 76 (4), 359-365.
- Stratford, P. W., Binkley, J. M., & Riddle, D. L. (1996). Health Status Measures: Strategies and Analytic Methods for Assessing Change Scores. *Physical Therapy*, 76 (10), 1109 -112.
- Streiner, D. L., & Norman, G. R. (2003). *Health Measurement Scales: A Practical Guide to Their Development and Use* (3<sup>rd</sup> Edition). Oxford University Press.
- Terwee, C. B., Bot, S. D. M., de Boer, M. R., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60, 34-42.
- Terwee, C. B., Roorda, L. D., Knola, D. L., et al. (2009). Linking measurement error to minimal important change of patient-reported outcomes. *Journal of Clinical Epidemiology*, 62, 1062-1067.
- van der Roer, N., Ostelo, R. W., Bekkering, G. E., van Tulder, M. W., & de Vet, H. C. (2006). Minimal clinically important change for pain intensity, functional status, and general health status in patients with non-specific low back pain. *Spine*, 31, 578 -582.
- van Middelkoop, M., Rubinstein, S. M., Kuijper, T., et al. (2011). A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J*, 20, 19–39.

- Von Korff, 1994. Studying the natural history of back pain. *Spine*, 19 (Suppl 18), 2041S-2046S.
- Von Korff, M., Ormel, J., Keefe, F. J., et al. (1992). Grading the severity of chronic pain. *Pain*; 50, 133-49.
- Waddell, G. (1987). A New Clinical Model for the Treatment of Low-Back Pain. *Spine*, 12 (7), 632 – 644.
- Walker, B. F., Muller, R., & Grant, W. D. (2004). Low Back Pain in Australian Adults - Prevalence and associated disability. *Journal of Manipulative and Physiological Therapeutics*, 27 (4), 238–244.
- WHO (1980). International Classification of Impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease. Geneva: World Health Organization.
- WHO (2000). Obesity: preventing and managing the global epidemic. Geneva: World Health Organization.
- WHO (2007). Steps to Health. A European framework to promote physical activity for health. Copenhagen: WHO regional office for Europe.
- WHO (2010). International Classification of Diseases (10<sup>th</sup> Revision) (ou ICD -10). Consultado a 10 de outubro, 2011, através de <http://apps.who.int/classifications/icd10/browse/2010/en>
- Wilhelm, F., Fayolle-Minon, I., Phaner, V., et al. (2010). Sensitivity to change of the Quebec Back Pain Disability Scale and the Dallas Pain Questionnaire. *Annals of Physical and Rehabilitation Medicine*, 53, 15–23.
- Wyrwich, K. W., Bullinger, M., Aaronson, N., et al. (2005). Estimating clinically significant differences in quality of life outcomes. *Qual Life Res*, 14, 285–295.

Young, A. E., Wasiak, R., Phillips, L., & Gross, D. P. (2011). Workers' perspectives on low back pain recurrence: "It comes and goes and comes and goes, but it's always there". *Pain*, 152, 204 - 211.

Yvanes-Thomas, M., Calmels, P., Bethoux, F., *et al.* (2002). Validity of the French-language version of the Quebec Back Pain Disability Scale in low back pain patients in France. *Joint Bone Spine*, 69, 397- 405.

**APÊNDICE A:**

**Estimativas de validade da QBPDS em diversos contextos culturais**

Este anexo, relativo às propriedades de validade da QBPDS, organiza-se de acordo com as 3 propriedades psicométricas de validade, propostas na Taxonomia de COSMIN (Mokkink *et al.*, 2010a).

### **Validade de conteúdo**

A validade de conteúdo é definida pelo grau com que o conteúdo de um instrumento reflete adequadamente o construto que pretende medir. A estimativa da validade de conteúdo é baseada no julgamento da relevância e da compreensão dos itens. Essa propriedade já foi devidamente estudada para a versão portuguesa (Nunes, Ribeiro & Cruz, 2005).

### **Validade de critério**

A validade de critério é definida como o grau com que as pontuações de um instrumento refletem adequadamente uma “medida de padrão de ouro”. O painel de peritos COSMIN chegou ao consenso que não há “medida de padrão de ouro” para “instrumentos de autorrelato de avaliação do estado de saúde” (Mokkink *et al.*, 2010a). A ausência dessa medida, impossibilita a análise da validade de critério da QBPDS.

### **Validade de construto**

A validade pressupõe verificar se um instrumento de medida avalia o construto específico para o qual foi desenvolvido. Uma clara definição do construto é o primeiro passo em qualquer avaliação da validade.

A validade de construto contém 3 aspetos: validade estrutural, que se centra nas relações internas; teste de hipóteses e validade transcultural. As 2 últimas concentram-se na relação com pontuações de outros instrumentos, ou diferenças entre grupos relevantes.

Como não há uma “medida de padrão de ouro” para a incapacidade funcional, que permita proceder à validação por critério da QBPDS, pode-se testar a validade correlacionando as pontuações com outro instrumento que meça o mesmo construto (validação do construto: teste de hipóteses) e verificar o grau com que duas medidas designadas para medir o mesmo construto estão relacionadas.

Os estudos que investigaram a validade de construto da QBPDS (na tabela seguinte), fizeram-no por correlação a outras escalas de autorrelato de incapacidade funcional, como RMDQ, o ODI e a subescala de função física do *Short Form-36* (SF-36); e ainda por correlação à intensidade da dor, através da Escala numérica (EN) e da Escala Visual Análoga (EVA).

Nas estimativas da validade de construto transversal da QBPDS, por correlação às diferentes medidas de incapacidade (Kopec *et al.*, 1995; Schoppink *et al.*, 1996; Mousavi *et al.*, 2006; Rodrigues *et al.*, 2009; Melikoglu *et al.*, 2009; Hicks & Manal, 2009), obtiveram-se coeficientes correlação compreendidos entre 0,6 a 0,91, que confirmaram uma associação moderada a forte com os vários questionários de incapacidade (RMDQ, ODI e SF-36)

Por outro lado, os coeficientes de correlação transversal da QBPDS com a intensidade da dor (Schoppink *et al.*, 1996; Yvanes- Thomas *et al.*, 2002; Bicer *et al.*, 2005; Mousavi *et al.*, 2006, Rodrigues *et al.*, 2009; Melikoglu *et al.*, 2009), revelaram valores entre 0,368 e 0,758. Através destas estimativas de correlação, depreende-se que há uma correlação fraca a moderada da QBPDS com a intensidade da dor.

O estudo de Schoppink *et al.* (1996) foi o único que analisou a validade de construto longitudinal da QBPDS. Concluíram que a validade de construto transversal com a escala RMDQ, em T0 e em T2 foi alta ( $r= 0,80$  e  $0,91$ , respetivamente) e que a validade de construto longitudinal com a escala RMDQ foi moderada ( $r= 0,60$ ). Os autores argumentam que a população em estudo, pacientes com DLC, dificultaram a avaliação da validade de construto longitudinal do questionário, uma vez que, neste tipo de população não é provável ocorrer grandes mudanças na capacidade funcional. O fato da validade de construto longitudinal ser apenas moderada, não implica necessariamente, que a escala não seja capaz de detetar mudanças ao longo do tempo.

A Tabela que se segue resume os estudos realizados sobre a validade de construto da QBPDS:

Tabela – Estimativas da Validade de construto da QBPDS em diversos contextos culturais

	<b>Versão da QBPDS</b>	<b>Amostra/ Recolha de dados</b>	<b>Validade de construto (VC)</b>
Kopec <i>et al.</i> (1995)	Inglesa	n = 242 DL/ T0= ponto de partida; T1 = 2 a 4 dias depois; T2= 2 a 6 meses depois.	VC transversal (QBPDS – RMDQ): r= 0,77
Schoppin <i>et al.</i> (1996)	Inglesa	n= 120 DLC/ T0 = ponto de partida; T1 = 1 semana; T2= 4 meses.	VC transversal em T0 (QBPDS – RMDQ): r=0,80 VC transversal em T2 (QBPDS – RMDQ): r=0,91 VC transversal em T0 (QBPDS – EN): r= 0,74 VC transversal em T2 (QBPDS – EN): r= 0,70 VC longitudinal (QBPDS – RMDQ): r= 0,60 VC Longitudinal (QBPDS – EN): r= 0,53
Yvanes-Thomas, <i>et al.</i> (2002)	Francesa	n= 32 DLC/ Um momento de avaliação	VC transversal (QBPDS – EVA): r= 0,448 (p< 0,0001);
Bicer <i>et al.</i> (2005)	Turca	n = 83 DLC/ Um momento de avaliação	VC transversal (QBPDS – EVA): r= 0,468 (p < 0,001).
Mousavi <i>et al.</i> (2006)	Iraniana	n = 100 DLC/ T0 = ponto de partida; T1 = 24horas depois (n=31);	VC transversal em T0 (QBPDS – SF-36): r =0,69 (p < 0,001). VC transversal em T0 (QBPDS – EVA): r =0,46 (p < 0,001).
Melikoglu <i>et al.</i> (2009)	Árabe	n = 100 DLC/ T0 = ponto de partida; T1 = 24horas depois;	VC transversal em T0 (QBPDS – EVA): r =0,368 (p < 0,000); VC transversal em T1 (QBPDS – EVA): r =0,441 (p < 0,000) VC transversal em T0 (QBPDS – ODI): r =0,666 (p < 0,000); VC transversal em T1 (QBPDS – ODI): r = 0,681 (p < 0,000);
Rodrigues <i>et al.</i> (2009)	Português do Brasil	n = 54 DL/ (sem informação)	VC transversal (QBPDS – RMDQ): r =0,857 VC transversal (QBPDS – EVA): r = 0,758
Bendeddouche <i>et al.</i> (2010)	Árabe	N = 64 DLC/ (sem informação)	VC transversal (QBPDS – RMDQ): r =0,644 (p< 0,001)

**APÊNDICE B:**  
**Protocolo para recrutamento dos participantes no estudo**



INSTITUTO POLITÉCNICO DE SETÚBAL - ESCOLA SUPERIOR DE SAÚDE

DEPARTAMENTO DE FISIOTERAPIA

**CONTRIBUTO PARA A VALIDAÇÃO DA VERSÃO PORTUGUESA DA QUEBEC BACK PAIN  
DISABILITY SCALE**

Ana Cristina Vieira; Eduardo Cruz; Rita Fernandes (2012)

**MANUAL PARA RECRUTAMENTO DOS  
PARTICIPANTES NO ESTUDO**

# Protocolo para recrutamento dos participantes no estudo

Este protocolo define as condições de recrutamento dos participantes no estudo, em três passos consecutivos. No final do documento inclui-se um glossário para esclarecimento de possíveis dúvidas.

## 1º PASSO – Identificar potenciais participantes

Verificar o processo clínico dos utentes enviados para a Fisioterapia e sinalizar os potenciais participantes através do diagnóstico médico de Dor Lombar: Ex: Dor Lombar, Tensão Lombar, Lombalgia, Lumbago ou Raquialgia Lombar (*International Classification of Diseases*, 2010)

## 2º PASSO – Verificar os critérios de inclusão e exclusão

Considerando um potencial participante, deve-se contactar o utente telefonicamente e verificar os critérios de inclusão e exclusão.

Nos critérios de exclusão, assume-se que a indicação para Fisioterapia inclui a verificação de situações de contraindicação ou situações com diagnóstico específico. Assim, assume-se que todos os participantes enviados para intervenção em Fisioterapia não possuem sintomas de compressão radicular, cauda equina, doença sistémica, inflamatória e/ou infecciosa, dor de origem visceral/ maligna ou fratura/risco de fratura associado a osteoporose (Smeets *et al.*, 2006). Assume-se igualmente que condições de dor lombar associadas a patologia específica, tais como, infeção, tumor, osteoporose, fratura, deformidade estrutural, doença inflamatória (ex. espondilite anquilosante), estão devidamente indicadas no registo ou processo clínico, permitindo dessa forma a exclusão destes participantes.

Assim, apenas terá que verificar **se o potencial participante:**

(coloque uma cruz no espaço apropriado para confirmar o critério):

Critérios de Inclusão	Sim
Dor localizada na região lombar e/ou associada a sintomatologia dos membros inferiores há mais de 3 meses, sem causa específica <b>ou</b> presença de recorrências de dor lombar por um período de seis meses, <b>ou</b> presença de dor em menos de metade dos dias num período de 12 meses, ocorrendo múltiplos episódios no decorrer desse período	<input type="checkbox"/>
Tem idade compreendida entre 18 e 65 anos	<input type="checkbox"/>
Sabe ler e escrever	<input type="checkbox"/>
Não realizou cirurgia lombar nos últimos 6 meses	<input type="checkbox"/>
Não realizou Fisioterapia ou outro tratamento conservador por sintomas lombares nos 3 meses prévios, com exceção da medicação para a dor	<input type="checkbox"/>
Não está grávida (Mulheres)	<input type="checkbox"/>

### 3º PASSO – Convidar o utente a participar no estudo

Concluído o processo de verificação dos critérios de inclusão e exclusão é necessário obter autorização por parte dos participantes.

Pela via telefónica e através de uma explicação sucinta da investigação, os utentes podem ser convidados a participar no estudo. Os que concordarem, são solicitados a se apresentar 30 minutos antes da marcação de fisioterapia (de modo a não interferir com o tempo dispendido para o tratamento), para lhes ser explicado o objetivo do estudo, os riscos e potenciais vantagens, os procedimentos para garantir a confidencialidade e o anonimato, assim como os procedimentos de recolha de dados (referidos no conteúdo da carta explicativa do estudo).

Os participantes informados e esclarecidos que continuarem interessados em participar, devem ser solicitados a formalizar a sua autorização para participar no estudo, mediante assinatura do formulário de consentimento informado. Depois do consentimento informado estar assinado e datado, o utente deve receber uma cópia desse documento.

# GLOSSÁRIO

## Classificação da Dor Lombar

A classificação da dor lombar (DL) em 3 categorias de acordo com as características clínicas apresentadas pelos utentes, tem sido cada vez mais utilizada no âmbito da investigação. Segundo Waddell (1987) a DL pode ser classificada da seguinte forma:

Dor com origem em patologia vertebral específica

Dor com origem radicular ou no nervo

Dor com origem não específica

## DL de Origem Não Específica e DL de Origem Específica

A DL de origem não específica diz respeito à dor lombar cuja origem não é atribuída a uma patologia específica e/ou causa conhecida, como por exemplo: infeção, tumor, osteoporose, fratura, deformidade estrutural, doença inflamatória (ex. espondilite anquilosante), síndrome radicular ou síndrome da cauda equina. A espondilose e a espondilolistesis também são classificadas como DL de origem não específica, uma vez que uma proporção considerável de utentes que as apresentam são assintomáticos (Soler & Calderón, 2000). A dor lombar de origem não específica pode ser ou não referida para o membro inferior mas sem deficit neurológico (alterações nos reflexos, sensibilidade ou força).

A DL de origem específica diz respeito à dor lombar cuja origem é atribuída a uma patologia específica e/ou causa conhecida. No entanto, a investigação tem demonstrado uma correlação baixa entre os sintomas de DL, os sinais radiológicos e a patologia, pelo que a DL de origem não específica, ou seja que não é atribuída a patologia vertebral específica ou a envolvimento neurológico, parece estar presente em 85% dos utentes que apresentam DL (Burton *et al.*, 2004)

## Justificação dos critérios de Inclusão e exclusão

- Dor localizada na região lombar e/ou associada a sintomatologia dos membros inferiores (Kendall *et al.*, 1997), com duração de pelo menos 12 semanas, sem causa específica, ou presença de recorrências de dor lombar por um período de seis meses, sem causa física específica, (Airaksinen *et al.*, 2006; Bekkering *et al.*, 2003; Krismer & van Tulder,

2007), nomeadamente doença neoplásica, infecciosa e/ou inflamatória, osteoporose, fratura, osteoporose, deformidade estrutural, síndrome da cauda equina e radicular (Waddell, 1987); Bekkering *et al.*, 2003; Krismer & van Tulder, 2007) **ou**

- Presença de dor em menos de metade dos dias num período de 12 meses, ocorrendo múltiplos episódios no decorrer desse período” (VonKorff, 1994). De acordo com VonKorff (1994), estes episódios de agudização são definidos como “períodos (normalmente uma semana ou menos) em que os sintomas da condição de dor lombar se encontram mais exacerbados do que o normal para o utente” **ou**
- Múltiplas situações de agudização dos sintomas, correspondendo assim a vários episódios de recorrência ao longo da sua condição crónica (Stanton, Latimer, Maher & Hancock, 2010);
- Idade compreendida entre os 18 e os 65 anos. A população idosa não é incluída pelo facto das alterações funcionais e biológicas associadas ao processo de envelhecimento poderem estar de alguma forma associadas à dor lombar dos participantes;
- Ausência de sintomas de compressão radicular<sup>16</sup>, cauda equina, doença sistémica, inflamatória e/ou infecciosa, dor de origem visceral/ maligna ou fratura/risco de fratura associado a osteoporose (Smeets *et al.*, 2006);
- Ausência de realização de cirurgia lombar nos 6 meses prévios e de realização de Fisioterapia ou outro tratamento conservador por sintomas lombares nos 3 meses prévios, com exceção da medicação para a dor;
- Mulheres que não se encontrem numa situação de gravidez. A dor lombar durante a gravidez é comum, tendo sido demonstrado que as mulheres grávidas experienciam algum tipo de DL durante o período de gestação, no entanto, os fatores etiológicos apontados são divergentes dos fatores associados à DCL de origem não específica (Garshasbi & Faghieh Zadeh, 2005).
- Saber ler e escrever. A condição de analfabetismo impossibilita o preenchimento dos instrumentos de avaliação da condição dos participantes.

---

<sup>16</sup> Note-se que existe uma diferenciação importante entre dor de origem específica (ex. síndrome de compressão radicular) e dor de origem não específica. Em ambos os casos pode existir dor referida para o membro inferior mas na dor lombar de origem não específica não existe deficit neurológico (alterações nos reflexos, sensibilidade ou força). É considerada a presença de sintomas de compressão radicular (origem específica) quando se verifica a presença de pelo menos dois testes neurológicos positivos: reflexos, dermatomas (sensibilidade) ou miotomas (força) (Ferreira *et al.*, 2011).

**APÊNDICE C:**  
**Caderno de instrumentos**



INSTITUTO POLITÉCNICO DE SETÚBAL - ESCOLA SUPERIOR DE SAÚDE

DEPARTAMENTO DE FISIOTERAPIA

**CONTRIBUTO PARA A VALIDAÇÃO DA VERSÃO PORTUGUESA DA QUEBEC BACK PAIN  
DISABILITY SCALE**

Ana Cristina Vieira; Eduardo Cruz; Rita Fernandes (2012)

**CADERNO DE INSTRUMENTOS**

## **Protocolo de recolha de dados**

Este protocolo destina-se apenas aos participantes no estudo que:

- **cumpriram todos os critérios de inclusão;**
- **aceitaram participar no estudo e assinaram o formulário de consentimento.**

A participação no estudo implica o preenchimento dos Instrumentos em três momentos distintos nos quais deve:

- **Garantir as mesmas condições de preenchimento nos momentos de recolha de dados, particularmente nos primeiros dois momentos;**
- **Respeitar o intervalo de tempo definido entre os momentos de recolha de dados;**
- **Respeitar a sequência de passagem dos instrumentos, caso exista;**

### **AVALIAÇÃO – MOMENTO 1**

O tempo médio de preenchimento dos instrumentos neste primeiro momento é de **6 minutos**. Solicite o preenchimento dos seguintes instrumentos, na ordem indicada:

- 1. QUESTIONÁRIO DE CARACTERIZAÇÃO SÓCIO -DEMOGRÁFICA E CLÍNICA (INCLUI ESCALA VISUAL ANÁLOGA)**
- 2. QUEBEC BACK PAIN DISABILITY SCALE - versão portuguesa**

### **AVALIAÇÃO – MOMENTO 2 (4 a 7 dias depois)**

O tempo médio de preenchimento dos instrumentos neste segundo momento é de **5 minutos**. Solicite o preenchimento dos seguintes instrumentos, na ordem indicada:

- 1. QUEBEC BACK PAIN DISABILITY SCALE - versão portuguesa**
- 2. PATIENT GLOBAL IMPRESSION CHANGE - versão portuguesa**

### **AVALIAÇÃO – MOMENTO 3 (6 semanas depois)**

O tempo médio de preenchimento dos instrumentos neste primeiro momento é de **5 minutos**. Solicite o preenchimento dos seguintes instrumentos, na ordem indicada:

- 1. QUEBEC BACK PAIN DISABILITY SCALE - versão portuguesa**
  - 2. PATIENT GLOBAL IMPRESSION CHANGE - versão portuguesa**
-

**CÓDIGO DO UTENTE** \_\_\_\_\_

**MOMENTO DE AVALIAÇÃO 1**

**(T0)**

**(Após verificação dos critérios de inclusão e assinatura de consentimento informado - Participantes em lista de espera ou na primeira semana de tratamento)**

Tempo de Preenchimento previsto apenas num momento: **6 minutos;**

**Seguir a ordem indicada:**

1. Questionário de caracterização Sócio-Demográfica e Clínica (inclui EVA)
2. Quebec Back Pain Disability Scale - versão portuguesa



DEPARTAMENTO DE FISIOTERAPIA  
INSTITUTO POLITÉCNICO DE SETÚBAL- ESCOLA SUPERIOR DE SAÚDE

### QUESTIONÁRIO DE CARACTERIZAÇÃO SÓCIO-DEMOGRÁFICA E CLÍNICA – DOR LOMBAR CRÓNICA

Um grupo de investigadores do Departamento de Fisioterapia do Instituto Politécnico de Setúbal está a realizar um estudo sobre os efeitos da Fisioterapia na dor e na capacidade funcional das pessoas. Para tal, agradecemos que respondesse com sinceridade a todas as questões que fazem parte da referida investigação. As suas respostas são confidenciais e anónimas. Obrigada.

Nome da Instituição: \_\_\_\_\_

Nº de Processo ou Código Atribuído ao Utente (a ser preenchido pelo responsável do estudo): \_\_\_\_\_

Data do preenchimento do questionário: \_\_\_\_/\_\_\_\_/\_\_\_\_

#### DADOS SOCIO-DEMOGRÁFICOS

**1. Idade** \_\_\_\_\_ **2. Sexo:** Masculino  Feminino

**3. Peso (kg):** \_\_\_\_\_ **4. Altura (cm):** \_\_\_\_\_

**5. Qual o seu Estado Civil?** (escolha uma das seguintes opções):

Solteiro(a)  Casado(a)  União de Facto  Viúvo(a)  Divorciado(a)

**6. Quais são as suas Habilitações Literárias?** (escolha uma das seguintes opções):

Ensino Primário  Ensino Básico completo (9º ano de escolaridade)  Ensino Secundário ou equivalente incompleto (12º ano de escolaridade)  Ensino Secundário ou equivalente completo (12º ano de escolaridade)  Ensino Superior incompleto (Politécnico ou Universitário)  Ensino Superior completo (Politécnico ou Universitário)

7. Qual a sua Atividade profissional/Profissão? \_\_\_\_\_

8. Fez formação Profissional específica? Sim  Não

9. Qual a sua situação profissional atual? (escolha uma das seguintes opções)

A  A trabalhar  Incapaz de  Desempregada  Reformada  Doméstica   
trabalhar a tempo parcial  trabalhar devido ao seu problema (o)  (o)  (o)

### DADOS CLÍNICOS

10. Há quanto tempo tem dor lombar? (escolha uma das seguintes opções)

3-6 meses  6-12 meses  12-24 meses  Mais de 24 meses

11. A sua dor prolonga-se para a perna?

Sim  Não

12. Atualmente toma alguma medicação para a sua dor lombar?

Sim  Não

13. No último ano faltou ao trabalho devido à sua dor?

Sim  Não

13.1. Se sim, quantas vezes?

1 vez  2 vezes  3 vezes  Mais de 3 vezes

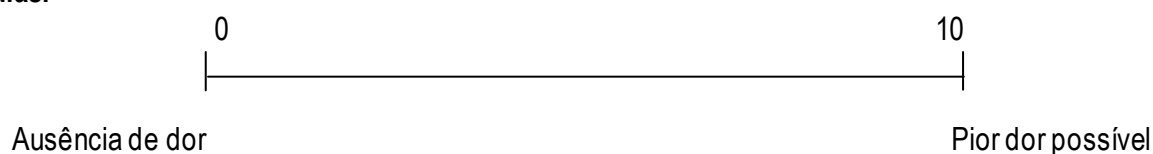
13.2. Durante quanto tempo (total de dias ou semanas que faltou no último ano)?

1 dia  2 dias  3 dias  1 semana  Mais de 1 semana

14. No último ano esteve de baixa remunerada (estado, seguros, empregador, etc)?

Sim  Não

15. Por favor, assinale com um risco vertical a intensidade media da sua dor durante os últimos 7 dias.



## Quebec Back Pain Disability Scale - versão portuguesa

Este questionário pretende saber como a sua dor nas costas afeta a sua vida no dia à dia. Pessoas com dores de costas poderão achar difícil a realização de algumas atividades diárias. Nós gostaríamos de saber se acha difícil a realização de algumas das atividades listadas abaixo, devido à sua dor de costas. Para cada atividade há uma escala de 0 a 5. Por favor escolha uma opção de resposta para cada atividade (**preencha todas as atividades**) colocando uma cruz no quadrado que corresponde à sua resposta.

**Hoje**, tem dificuldade em realizar as seguintes atividades devido à sua dor de costas?

		0 Sem dificuldade nenhuma	1 Com Um mínimo de dificuldade	2 Com alguma dificuldade	3 Com Bastante dificuldade	4 Com muita dificuldade	5 Incapaz de realizar
1	Levantar-se da cama						
2	Dormir toda a noite						
3	Virar-se na cama						
4	Andar de carro						
5	Estar de pé durante 20-30 minutos						
6	Estar sentado numa cadeira por várias horas						
7	Subir um lance de escadas						
8	Andar 300-400 metros						
9	Andar vários quilómetros						
10	Alcançar prateleiras altas						
11	Atirar uma bola						
12	Correr cerca de 100 metros						
13	Tirar comida do frigorífico						
14	Fazer a cama						
15	Calçar meias ( <i>collants</i> )						
16	Dobrar-se à frente para limpar a banheira						
17	Mover uma cadeira						
18	Puxar ou empurrar portas pesadas						
19	Carregar dois sacos de compras						
20	Levantar e carregar uma mala pesada						

Adicione todos os números para obter um **score total**: \_\_\_\_\_

**Muito obrigado pela sua colaboração**

**CÓDIGO DO UTENTE** \_\_\_\_\_

## **MOMENTO DE AVALIAÇÃO 2**

**(T1)**

**(4 a 7 dias após- Sem Intervenção ou durante a 1.<sup>a</sup> semana de tratamento)**

Tempo de Preenchimento previsto apenas num momento: **5 minutos;**

### **Seguir a ordem indicada:**

1. Quebec Back Pain Disability Scale - versão portuguesa:
2. Patient Global Impression of Change Scale- versão portuguesa

## Quebec Back Pain Disability Scale - versão portuguesa

Este questionário pretende saber como a sua dor nas costas afeta a sua vida no dia a dia. Pessoas com dores de costas poderão achar difícil a realização de algumas atividades diárias. Nós gostaríamos de saber se acha difícil a realização de algumas das atividades listadas abaixo, devido à sua dor de costas. Para cada atividade há uma escala de 0 a 5. Por favor escolha uma opção de resposta para cada atividade (**preencha todas as atividades**) colocando uma cruz no quadrado que corresponde à sua resposta.

**Hoje**, tem dificuldade em realizar as seguintes atividades devido à sua dor de costas?

		0 Sem dificuldade nenhuma	1 Com Um mínimo de dificuldade	2 Com alguma dificuldade	3 Com Bastante dificuldade	4 Com muita dificuldade	5 Incapaz de realizar
1	Levantar-se da cama						
2	Dormir toda a noite						
3	Virar-se na cama						
4	Andar de carro						
5	Estar de pé durante 20-30 minutos						
6	Estar sentado numa cadeira por várias horas						
7	Subir um lance de escadas						
8	Andar 300-400 metros						
9	Andar vários quilómetros						
10	Alcançar prateleiras altas						
11	Atirar uma bola						
12	Correr cerca de 100 metros						
13	Tirar comida do frigorífico						
14	Fazer a cama						
15	Calçar meias ( <i>collants</i> )						
16	Dobrar-se à frente para limpar a banheira						
17	Mover uma cadeira						
18	Puxar ou empurrar portas pesadas						
19	Carregar dois sacos de compras						
20	Levantar e carregar uma mala pesada						

Adicione todos os números para obter um **score total**: \_\_\_\_\_

## Patient Global Impression Change Scale - versão portuguesa

Desde o início do tratamento nesta instituição, como é que descreve a mudança (se houve) nas LIMITAÇÕES DE ATIVIDADES, SINTOMAS, EMOÇÕES E QUALIDADE DE VIDA no seu global, em relação à sua dor (selecione UMA opção):

- |  |                          |   |
|--|--------------------------|---|
| Sem alterações (ou a condição piorou)  | <input type="checkbox"/> | 1 |
| Quase na mesma, sem qualquer alteração visível                               | <input type="checkbox"/> | 2 |
| Ligeiramente melhor, mas, sem mudanças consideráveis                         | <input type="checkbox"/> | 3 |
| Com algumas melhorias, mas a mudança não representou qualquer diferença real | <input type="checkbox"/> | 4 |
| Moderadamente melhor, com mudança ligeira mas significativa                  | <input type="checkbox"/> | 5 |
| Melhor, e com melhorias que fizeram uma diferença real e útil                | <input type="checkbox"/> | 6 |
| Muito melhor, e com uma melhoria considerável que fez toda a diferença       | <input type="checkbox"/> | 7 |

**Muito obrigado pela sua colaboração**

**CÓDIGO DO UTENTE** \_\_\_\_\_

**MOMENTO DE AVALIAÇÃO 3  
(T2)  
(6 semanas após início do tratamento)**

Tempo de Preenchimento Previsto: **5 minutos**

**Seguir a ordem indicada:**

1. Quebec Back Pain Disability Scale - versão portuguesa:
2. Patient Global Impression of Change Scale- versão portuguesa (inclui EVA)

## Quebec Back Pain Disability Scale- versão portuguesa

Este questionário pretende saber como a sua dor nas costas afeta a sua vida no dia a dia. Pessoas com dores de costas poderão achar difícil a realização de algumas atividades diárias. Nós gostaríamos de saber se acha difícil a realização de algumas das atividades listadas abaixo, devido à sua dor de costas. Para cada atividade há uma escala de 0 a 5. Por favor escolha uma opção de resposta para cada atividade (**preencha todas as atividades**) colocando uma cruz no quadrado que corresponde à sua resposta.

**Hoje**, tem dificuldade em realizar as seguintes atividades devido à sua dor de costas?

		0 Sem dificuldade nenhuma	1 Com Um mínimo de dificuldade	2 Com alguma dificuldade	3 Com Bastante dificuldade	4 Com muita dificuldade	5 Incapaz de realizar
1	Levantar-se da cama						
2	Dormir toda a noite						
3	Virar-se na cama						
4	Andar de carro						
5	Estar de pé durante 20-30 minutos						
6	Estar sentado numa cadeira por várias horas						
7	Subir um lance de escadas						
8	Andar 300-400 metros						
9	Andar vários quilómetros						
10	Alcançar prateleiras altas						
11	Atirar uma bola						
12	Correr cerca de 100 metros						
13	Tirar comida do frigorífico						
14	Fazer a cama						
15	Calçar meias ( <i>collants</i> )						
16	Dobrar-se à frente para limpar a banheira						
17	Mover uma cadeira						
18	Puxar ou empurrar portas pesadas						
19	Carregar dois sacos de compras						
20	Levantar e carregar uma mala pesada						

Adicione todos os números para obter um **score total**: \_\_\_\_\_

## Patient Global Impression Change Scale- versão portuguesa

Desde o início do tratamento nesta instituição, como é que descreve a mudança (se houve) nas LIMITAÇÕES DE ATIVIDADES, SINTOMAS, EMOÇÕES E QUALIDADE DE VIDA no seu global, em relação à sua dor (selecione UMA opção):

- |  |                          |   |
|--|--------------------------|---|
| Sem alterações (ou a condição piorou)  | <input type="checkbox"/> | 1 |
| Quase na mesma, sem qualquer alteração visível                               | <input type="checkbox"/> | 2 |
| Ligeiramente melhor, mas, sem mudanças consideráveis                         | <input type="checkbox"/> | 3 |
| Com algumas melhorias, mas a mudança não representou qualquer diferença real | <input type="checkbox"/> | 4 |
| Moderadamente melhor, com mudança ligeira mas significativa                  | <input type="checkbox"/> | 5 |
| Melhor, e com melhorias que fizeram uma diferença real e útil                | <input type="checkbox"/> | 6 |
| Muito melhor, e com uma melhoria considerável que fez toda a diferença       | <input type="checkbox"/> | 7 |

**Muito obrigado pela sua colaboração**

**APÊNDICE D:**

**Pedido de autorização à administração do SESARAM, E.P.E.**

**ENTRADA**  
E.A159508 / 5 2011/11/07



HOSPITAL CENTRAL DO FUNCHAL  
DIRECÇÃO CLÍNICA

Setúbal, 7 de novembro de 2011

Exmo. Sr. Presidente do Conselho de Administração do  
SESARAM, E.P.E.,  
Dr. António João Prado Almada Cardoso

**Assunto:** Pedido de autorização para desenvolver Projeto de Investigação em Fisioterapia no âmbito de Mestrado em Fisioterapia, com a participação dos utentes do SESARAM, E.P.E.

O meu nome é Ana Cristina Pires Vieira, sou Fisioterapeuta – Técnica de Diagnóstico e Terapêutica 2.ª classe do Serviço da Saúde da Região Autónoma da Madeira, E.P.E. com o número mecanográfico 005710 e exerço funções no Setor de Fisioterapia do Centro de Saúde do Caniço. Sou estudante do Mestrado em Fisioterapia - Ramo Condições Músculo-Esqueléticas lecionado em parceria pela Escola Superior de Saúde do Instituto Politécnico de Setúbal, pela Faculdade de Ciências Médicas da Universidade Nova de Lisboa e pela Escola Nacional de Saúde Pública. De momento estou a desenvolver minha dissertação de final de curso, no âmbito da Unidade Curricular de "Trabalho de Projeto", cujo objetivo é estudar os efeitos da intervenção da Fisioterapia, ao nível da dor, capacidade funcional e da percepção global de melhoria dos indivíduos com dor lombar crónica. O presente estudo tem como orientador o Professor Eduardo Cruz e a coorientação da Professora Rita Fernandes.

Trata-se de um estudo multicentro, de coorte prospetivo, que envolve uma amostra obtida a partir de vários serviços de Medicina Física e Reabilitação de Portugal. Tratando-se de um estudo observacional não ocorrerá qualquer interferência no plano de tratamento da Fisioterapia estabelecido. Pretende-se apenas recolher dados sobre um conjunto de variáveis demográficas e clínicas dos utentes e registar o tipo de tratamento efetuado pelos Fisioterapeutas para pesquisar se há diferenças nos resultados da dor, condição funcional dos utentes, na intensidade da dor, e na percepção global de melhoria dos utentes com dor lombar crónica, quando os indivíduos são expostos a diferentes tipos de intervenção.

*Aut. do Dr. João Prado Almada Cardoso*  
*6/11/11*  
*A Comissão de Ética*  
*para*  
*10/11/2011*  
*DIRECÇÃO CLÍNICA*

O desenho de estudo (Anexo A) surge enquadrado num grupo de estudos desenvolvidos no âmbito das mesmas instituições, que visam no seu global estudar a dor lombar crónica. Para aumentar o potencial estatístico, serve-se de uma amostra comum e utiliza-se um caderno de instrumentos também comum, que vai permitir contribuir para a validação de instrumentos de avaliação dos utentes com dor lombar crónica utilizados na Fisioterapia, estudar indicadores de prognóstico para os resultados da intervenção em fisioterapia em indivíduos e os efeitos da intervenção da Fisioterapia.

Para esses fins, utilizam-se 8 questionários de autopreenchimento (disponíveis no Anexo B):

- Questionário de Caracterização Sócio-Demográfica & Clínica que inclui Escala Visual Análoga (EVA) ✓
- Quebec Back Pain Disability Scale - Versão Portuguesa ✓
- Patterns of Activity Measure – Pain - Versão Portuguesa ✓
- Questionário de Baecke - Versão Portuguesa ✓
- Roland Morris Questionnaire – Versão Portuguesa ✓
- Patient Global Impression of Change – Versão Portuguesa ✓
- Tampa Scale of Kinesiophobia - Versão Portuguesa ✓
- International Physical Activity Questionnaire – Versão Portuguesa ✓

A recolha de dados será realizada em 3 momentos pré-definidos. O primeiro momento (T0), coincidente com o início dos tratamentos de Fisioterapia, em que se recolhem dados sócio-demográficos e clínicos relativos ao utente e à sua condição. O segundo momento (T1) ocorre cerca de 3 a 7 dias depois, para fins de fiabilidade teste-reteste, e o terceiro momento (T2), após 6 semanas de intervenção, permite avaliar outras propriedades psicométricas dos instrumentos e os resultados da intervenção.

Para além da recolha de dados junto ao utente, pretende-se ainda recolher dados relativos a tipologia, frequência, e duração dos tratamentos de Fisioterapia, através da grelha pré-definida relativa aos procedimentos e modalidades terapêuticas realizadas (ANEXO C), preenchida pelos Fisioterapeutas colaboradores.

Os participantes no estudo são recrutados a partir da lista de espera dos utentes referenciados pelo médico para os setores de fisioterapia. São adultos com dor lombar crónica, que satisfazem um conjunto de critérios de inclusão e exclusão, conforme está no manual para recrutamento dos participantes no estudo (ANEXO D).

Após verificação dos critérios de inclusão/ exclusão todos os indivíduos elegíveis recebem informações claras sobre os objetivos deste estudo e o contexto em que é efetuado, mediante uma carta explicativa do estudo (ANEXO D). É clarificada a voluntariedade da sua participação e assegurada a garantia de que podem interromper a sua participação em qualquer altura, sem que esse facto implique qualquer constrangimento ou dano relativamente ao seu tratamento. É igualmente assegurado o anonimato da sua participação e de todos os dados recolhidos durante a execução deste estudo. Esse anonimato é assegurado pelos investigadores através da atribuição de um número de codificação a todos os instrumentos de avaliação utilizados. Os indivíduos que aceitarem participar assinam o termo de consentimento informado.

O tratamento e análise de dados serão realizados no software PASW versão 18. É possível que os resultados deste estudo sejam divulgados e/ou publicados no futuro, nomeadamente no âmbito da apresentação do Trabalho de Projeto do Mestrado em Fisioterapia - Ramo das Condições Músculo-Esqueléticas, nunca sendo os participantes identificados de forma individual. Uma vez apresentados os resultados, os dados originais serão destruídos.

Para tal, gostaria de solicitar **a sua autorização para distribuir questionários e recolher os dados dos utentes com dor lombar crónica, em tratamento no Departamento de Medicina Física e Reabilitação do SESARAM, E.P.E., nos vários setores de Fisioterapia do Hospital e Centros de Saúde.**

Certos que o seu contributo nos irá ajudar a desenvolver este estudo agradecemos antecipadamente a sua colaboração e disponibilidade.

Com os melhores cumprimentos,

Ana Cristina Pires Vieira

Ana Cristina Pires Vieira

**APÊNDICE E:**

**Carta explicativa do estudo e consentimento informado**

DEPARTAMENTO DE FISIOTERAPIA  
ESCOLA SUPERIOR DE SAÚDE - INSTITUTO POLITÉCNICO DE SETÚBAL  
**CONTRIBUTO PARA A VALIDAÇÃO DA VERSÃO PORTUGUESA DA QUEBEC BACK PAIN  
DISABILITY SCALE**  
Ana Cristina Veira; Eduardo Cruz; Rita Fernandes (2012)

## **CARTA EXPLICATIVA DO ESTUDO AOS UTENTES**

Foi-lhe pedido para participar num estudo de investigação no âmbito de desenvolvimento da Dissertação de Mestrado em Fisioterapia – Ramo das Condições Músculo-Esqueléticas. Estará envolvido na recolha de informação, que visa ajudar a intervenção da fisioterapia em utentes com dor lombar crónica. A participação neste estudo significa que a informação sobre si e a sua condição será recolhida e analisada juntamente com as informações recolhidas de outras pessoas com o mesmo problema de saúde. As respostas individuais que fornecer serão confidenciais.

### **QUAL É O OBJETIVO DESTE ESTUDO?**

Este estudo tem como principal objetivo contribuir para a validação da versão Portuguesa de um instrumento de autorrelato, que é frequentemente utilizado em utentes com dor lombar crónica, para avaliar e monitorizar o nível de incapacidade funcional: *Quebec Back Pain Disability Scale*.

O estudo irá recolher informação de 200 pessoas com dor lombar crónica, que frequentam serviços de Medicina Física e Reabilitação em Portugal. A informação recolhida neste estudo poderá, no futuro, ajudar a reconhecer intervenções da Fisioterapia mais eficazes para aumentar a capacidade funcional e a perceção global de melhoria dos indivíduos com dor lombar crónica.

### **O QUE É QUE ESTE ESTUDO ENVOLVE?**

Ser-lhe-á pedido para responder a 3 questionários: Questionário de caracterização sócio-demográfica e clínica, que inclui Escala Visual Análoga (EVA); *Quebec Back Pain Disability Scale* - Versão Portuguesa; e *Patient Global Impression of Change* – Versão Portuguesa;

O seu preenchimento será realizado em 3 momentos pré-definidos, com a repetição de alguns dos questionários. O primeiro momento (T0), coincidente com o início dos tratamentos de Fisioterapia, em que se recolhem dados sócio-demográficos relativos ao utente e à sua condição, pode levar cerca de 6 minutos a completar. O segundo momento (T1) ocorre cerca de 4 a 7 dias depois, com consumo de cerca

de 5 minutos, e o terceiro momento (T2), após 6 semanas de intervenção de Fisioterapia, com consumo de aproximadamente 5 minutos.

A informação recolhida será armazenada juntamente com as informações de outros utentes com dor lombar crónica.

### **A QUEM É PEDIDO PARA PARTICIPAR NESTE ESTUDO?**

Foi-lhe pedido para participar neste estudo pois tem sintomas de dor lombar crónica, com recomendação para realizar um plano de intervenção de Fisioterapia. Serão envolvidos 200 doentes de vários serviços de Medicina Física e Reabilitação em Portugal.

### **EXISTEM RISCOS NESTA PARTICIPAÇÃO?**

Não existem riscos associados ao preenchimento dos questionários para este estudo e o estudo não irá interferir no plano de intervenção.

### **EXISTEM BENEFÍCIOS POR PARTICIPAR?**

Não irá receber nenhum benefício imediato por participar neste estudo. No entanto, a informação recolhida no estudo beneficiará no futuro pessoas com dor lombar crónica que frequentam os tratamentos de Fisioterapia.

### **QUEM TERÁ ACESSO À MINHA INFORMAÇÃO?**

Não há identificação do seu nome em nenhum relatório. Todo o material recolhido, pertencente a este estudo, será codificado e tratado de forma anónima e confidencial, sendo conservado à responsabilidade de Ana Cristina Vieira. Contudo, não podemos garantir confidencialidade absoluta. A sua informação pessoal pode ser revelada se solicitada pelas vias legais. É também possível que os resultados deste estudo sejam divulgados e/ou publicados no futuro, nomeadamente no âmbito da apresentação do Trabalho de Projeto do Mestrado em Fisioterapia - Ramo das Condições Músculo-Esqueléticas, nunca sendo os participantes identificados de forma individual. Uma vez apresentados os resultados, os dados originais serão destruídos.

### **EXISTEM CUSTOS ENVOLVIDOS?**

A sua participação não envolve quaisquer encargos ou despesas da sua parte, com exceção do tempo necessário para o preenchimento dos questionários.

## QUAIS SÃO OS MEUS DIREITOS?

A sua participação neste estudo é inteiramente voluntária, pode recusar participar neste estudo ou desistir em qualquer altura, sem ter que fornecer qualquer tipo de explicação. Se decidir não participar, isto não afetará o seu futuro tratamento, ou direitos de saúde e legais.

## A QUEM POSSO CONTACTAR SE TIVER ALGUMA QUESTÃO OU PREOCUPAÇÃO?

Se tiver alguma dúvida sobre os seus direitos como participante, pode contactar uma das responsáveis por este estudo, a Fisioterapeuta Ana Cristina Pires Vieira, que exerce as suas funções no Centro de Saúde do Caniço, através dos números: 966478135 ou 291930070 (n.º de telefone do C.S. do Caniço), ou por e-mail: [ft.cristinavieira@gmail.com](mailto:ft.cristinavieira@gmail.com)



## DOCUMENTO DE CONSENTIMENTO INFORMADO

Entendo que toda a informação derivada do estudo “Contributo para a Validação da versão Portuguesa da *Quebec Back Pain Disability Scale*” é propriedade de Ana Cristina Vieira, Eduardo Cruz e Ana Rita Fernandes. Dou o meu consentimento para que dados anónimos a meu respeito possam ser guardados e processados pela Escola Superior de Saúde do Instituto Politécnico de Setúbal para fins de avaliação científica. Li (Foi-me lida) a informação mencionada acima. Entendo o significado desta informação, e as minhas perguntas foram satisfatoriamente respondidas. Tive tempo suficiente para decidir sobre a participação neste estudo. Venho por este meio consentir a minha participação e consentir na recolha, uso e revelação de informação. Irei receber uma cópia deste documento de consentimento informado assinada e datada.

### O participante

\_\_\_\_\_  
\_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de 20\_\_\_\_

### O investigador responsável pelo estudo

\_\_\_\_\_  
(Ana Cristina Pires Vieira)  
\_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de 20\_\_\_\_

**APÊNDICE F:**  
**Pedido de colaboração aos fisioterapeutas**

Setúbal, 15 de novembro de 2011

## PEDIDO DE COLABORAÇÃO NA RECOLHA DE DADOS

Caro(a) Colega

O meu nome é Ana Cristina Pires Vieira, sou estudante do Mestrado em Fisioterapia Ramo Condições Músculo-Esqueléticas lecionado em parceria pela Escola Superior de Saúde do Instituto Politécnico de Setúbal, pela Faculdade de Ciências Médicas e Escola Nacional de Saúde Pública, da Universidade Nova de Lisboa. Atualmente estou a desenvolver a minha dissertação de final de curso, no âmbito da Unidade Curricular de “Trabalho de Projeto”, cujo objetivo é contribuir para a validação da versão Portuguesa da *Quebec Back Pain Disability Scale*. O presente estudo tem como orientador o Professor Eduardo Cruz.

Neste contexto gostaria de solicitar a sua colaboração para distribuir e recolher 3 questionários de autopreenchimento:

- Questionário de Caracterização Sócio-Demográfica e Clínica, que inclui Escala Visual Análoga (EVA)
- Quebec Back Pain Disability Scale - Versão Portuguesa
- Patient Global Impression of Change – Versão Portuguesa

Estes questionários devem ser distribuídos **apenas** aos utentes que se enquadrem nos seguintes critérios de inclusão/ exclusão:

- Dor localizada na região lombar e/ou associada a sintomatologia dos membros inferiores, com duração de pelo menos três meses (Kendall et al., 1997), sem causa específica conhecida, nomeadamente doença neoplásica, infecciosa e/ou inflamatória,

osteoporose, fratura, osteoporose, deformidade estrutural, síndrome da cauda equina e radicular (Waddell, 1987); **ou**

- Presença de dor em menos de metade dos dias num período de 12 meses, ocorrendo múltiplos episódios no decorrer desse período” (Von Korff, 1994); **ou**
- Múltiplas situações de agudização dos sintomas, correspondendo assim a vários episódios de recorrência ao longo da sua condição crónica (Stanton, Latimer, Maher & Hancock, 2010);
- Idade compreendida entre os 18 e os 65 anos;
- Ausência de sintomas de compressão radicular, cauda equina, doença sistémica e/ou infecciosa, dor de origem visceral/ maligna ou fratura/ risco de fratura associado a osteoporose (Smeets et al., 2006c);
- Ausência de realização de cirurgia lombar nos 6 meses prévios e de realização de Fisioterapia ou outro tratamento conservador por sintomas lombares nos 3 meses prévios, com exceção da medicação para a dor;
- Estar a realizar fisioterapia sem acumular com outro tipo de tratamento para a Dor Crónica Lombar com exceção de medicação para a dor;
- Mulheres que não se encontrem numa situação de gravidez. A dor lombar durante a gravidez é comum, tendo sido demonstrado que as mulheres grávidas experienciam algum tipo de DL durante o período de gestação, no entanto, os fatores etiológicos apontados são divergentes dos fatores associados à DCL de origem não específica (Garshasbi & Faghieh Zadeh, 2005);
- Saber ler e escrever;

Considerando que a realização deste estudo só será possível com a sua participação, agradeço antecipadamente a sua colaboração. Caso surja alguma dúvida por favor não hesite em contactar-nos através do número 966478135 ou pelo seguinte e-mail: [ft.cristinavieira@gmail.com](mailto:ft.cristinavieira@gmail.com).

Com os melhores cumprimentos,

Ana Cristina Pires Vieira

**APÊNDICE G:**  
***Output IBM SPSS Statistics* – Caracterização sócio-demográfica e  
clínica da amostra**

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\Cristina_estudo2_caractAmos  
tra_1406.sav'.  
DATASET NAME DataSet4 WINDOW=FRONT.  
DATASET ACTIVATE DataSet1.  
DESCRIPTIVES VARIABLES=Idd  
  /STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

```
[DataSet1]  
C:\Users\Cristina\Desktop\JULHO_2012\Cristina_estudo1_caractAmostra_05  
07.sav
```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Idade (Anos)	132	18	65	46,58	12,665
Valid N (listwise)	132				

```
FREQUENCIES VARIABLES=Genero Hab_Literarias Duracao_Dor Local_Dor  
Sit_Prof Baixa  
  /ORDER=ANALYSIS.
```

## Frequencies

```
[DataSet1]  
C:\Users\Cristina\Desktop\JULHO_2012\Cristina_estudo1_caractAmostra_05  
07.sav
```

**Statistics**

	Género	Habilitações Literárias	Duração da dor	Localização da dor	Situação Profissional Atual
N	Valid	132	132	132	132
	Missing	0	0	0	0

**Statistics**

	Baixa remunerada	
N	Valid	132
	Missing	0

**Gênero**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Masculino	36	27,3	27,3	27,3
Feminino	96	72,7	72,7	100,0
Total	132	100,0	100,0	

**Habilitações Literárias**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Ensino Primário	33	25,0	25,0	25,0
Ensino Básico completo	25	18,9	18,9	43,9
Ensino Secundário incompleto	22	16,7	16,7	60,6
Ensino Secundário completo	15	11,4	11,4	72,0
Ensino Superior incompleto	6	4,5	4,5	76,5
Ensino Superior completo	31	23,5	23,5	100,0
Total	132	100,0	100,0	

**Duração da dor**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 3-6 meses	18	13,6	13,6	13,6
6-12 meses	12	9,1	9,1	22,7
12-24 meses	15	11,4	11,4	34,1
Mais de 24 meses	87	65,9	65,9	100,0
Total	132	100,0	100,0	

**Localização da dor**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Sem dor referida para a perna	62	47,0	47,0	47,0
Com dor referida para a perna	70	53,0	53,0	100,0
Total	132	100,0	100,0	

**Medicação**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	72	54,5	54,5	54,5
Valid Sim	60	45,5	45,5	100,0
Total	132	100,0	100,0	

**Situação Profissional Atual**

	Frequency	Percent	Valid Percent	Cumulative Percent
A trabalhar a tempo inteiro	80	60,6	60,6	60,6
A trabalhar a tempo parcial	7	5,3	5,3	65,9
Incapaz de trabalhar	5	3,8	3,8	69,7
Valid Desempregado (a)	10	7,6	7,6	77,3
Reformado (a)	16	12,1	12,1	89,4
Doméstica (o)	14	10,6	10,6	100,0
Total	132	100,0	100,0	

**Baixa remunerada**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	102	77,3	77,3	77,3
Valid Sim	30	22,7	22,7	100,0
Total	132	100,0	100,0	

**APÊNDICE H:**  
***Output IBM SPSS Statistics – Análise Fatorial***

```

FACTOR
/VARIABLES Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16
Q17 Q18 Q19 Q20
/MISSING LISTWISE
/ANALYSIS Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16 Q17
Q18 Q19 Q20
/PRINT INITIAL KMO EXTRACTION
/PLOT EIGEN
/CRITERIA MINEIGEN(1) ITERATE(25)
/EXTRACTION PC
/ROTATION NOROTATE
/METHOD=CORRELATION.

```

## Factor Analysis

[DataSet1] C:\Users\Cristina\Desktop\JULHO\_2012\Análise Fatorial  
QBPDQ-T0.sav

### KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	,917
Approx. Chi-Square	1886,085
Bartlett's Test of Sphericity	df
	190
	Sig.
	,000

### Communalities

	Initial	Extraction
Q1	1,000	,707
Q2	1,000	,751
Q3	1,000	,710
Q4	1,000	,554
Q5	1,000	,564
Q6	1,000	,785
Q7	1,000	,717
Q8	1,000	,665
Q9	1,000	,696
Q10	1,000	,692
Q11	1,000	,621
Q12	1,000	,680
Q13	1,000	,739
Q14	1,000	,816
Q15	1,000	,649

Q16	1,000	,666
Q17	1,000	,684
Q18	1,000	,818
Q19	1,000	,777
Q20	1,000	,735

Extraction Method: Principal  
Component Analysis.

**Total Variance Explained**

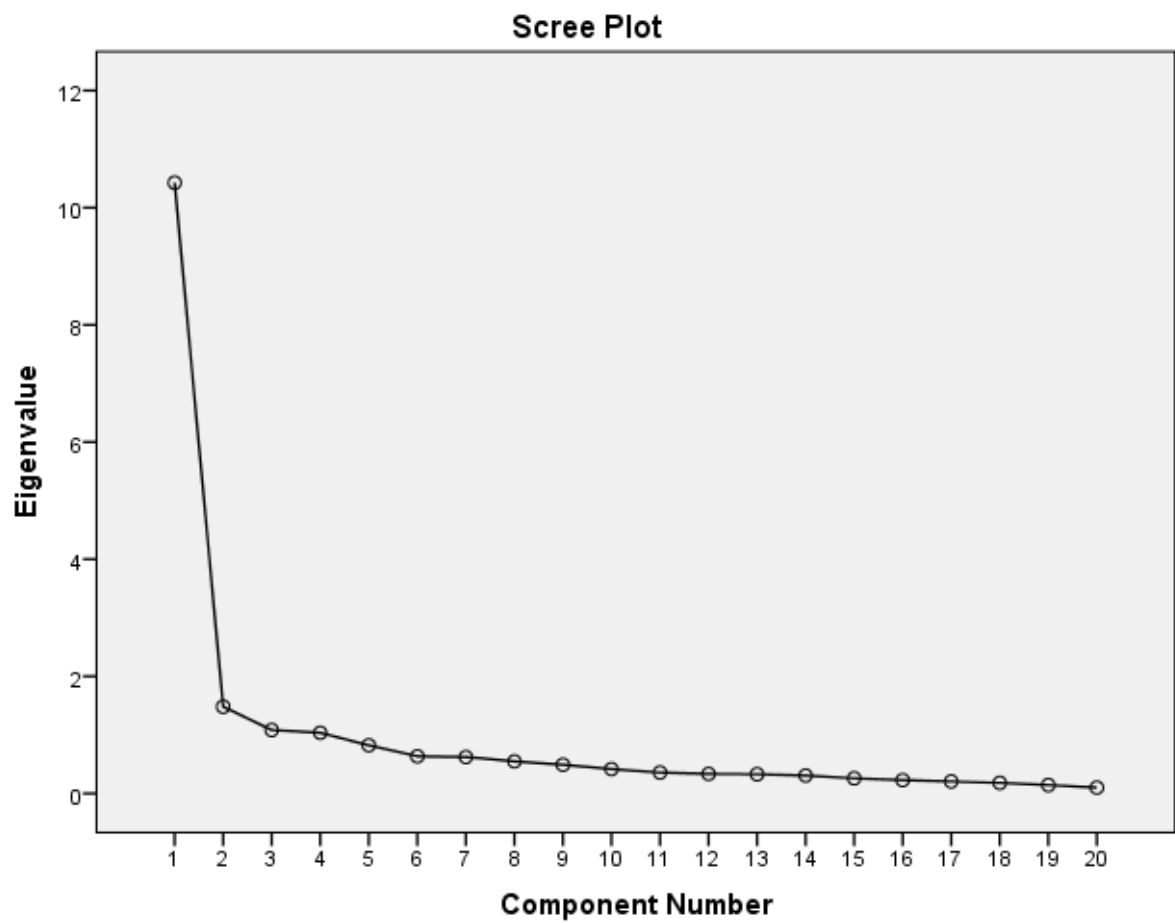
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings	
	Total	% of Variance	Cumulative %	Total	% of Variance
1	10,428	52,138	52,138	10,428	52,138
2	1,481	7,404	59,542	1,481	7,404
3	1,083	5,416	64,959	1,083	5,416
4	1,036	5,180	70,139	1,036	5,180
5	,823	4,114	74,253		
6	,635	3,173	77,425		
7	,622	3,112	80,538		
8	,547	2,737	83,275		
9	,490	2,448	85,723		
10	,417	2,084	87,807		
11	,356	1,782	89,589		
12	,334	1,668	91,257		
13	,328	1,642	92,899		
14	,306	1,531	94,429		
15	,258	1,290	95,720		
16	,228	1,142	96,862		
17	,205	1,024	97,885		
18	,180	,900	98,785		
19	,143	,717	99,503		
20	,099	,497	100,000		

**Total Variance Explained**

Component	Extraction Sums of Squared Loadings
	Cumulative %
1	52,138
2	59,542
3	64,959
4	70,139

5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

Extraction Method: Principal Component Analysis.



Component Matrix<sup>a</sup>

	Component			
	1	2	3	4
Q1	,690	,457	-,054	,138
Q2	,580	,438	-,329	,340
Q3	,682	,325	-,148	,343
Q4	,679	,200	,155	,171
Q5	,616	-,023	,429	-,018
Q6	,512	-,036	,717	,085
Q7	,792	-,084	,063	,279
Q8	,763	-,071	-,019	,278
Q9	,757	-,244	,049	,246
Q10	,761	-,327	,078	-,021
Q11	,766	-,177	-,002	-,055
Q12	,702	-,395	-,113	,138
Q13	,685	,272	,008	-,442
Q14	,748	,342	,024	-,373
Q15	,715	,294	,150	-,170
Q16	,783	,084	,047	-,210
Q17	,773	,084	-,201	-,199
Q18	,779	-,278	-,304	-,204
Q19	,817	-,284	-,155	-,074
Q20	,764	-,343	-,140	-,121

Extraction Method: Principal Component Analysis.<sup>a</sup>

a. 4 components extracted.

**APÊNDICE I:**  
***Output IBM SPSS Statistics – Consistência interna***

## Estimativa de $\alpha$ de cronbach

```
GET
  FILE='C:\Users\Cristina\Desktop\ Análise Fatorial QBPDQ-T0.sav'.

Warning.  Command name: GET FILE
SPSS Statistics data file "C:\Users\Cristina\Desktop\ Análise
Factorial QBPDQ-T0.sav" is written in a character encoding (ISO-8859-
15)
incompatible with the current LOCALE setting.  It may not be readable.
Consider changing LOCALE or setting UNICODE on.  (DATA 1721)
DATASET NAME DataSet1 WINDOW=FRONT.
RELIABILITY
  /VARIABLES=Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16
Q17 Q18 Q19 Q20
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR
  /SUMMARY=TOTAL.
```

### Reliability

[DataSet1] C:\Users\Cristina\Desktop\Análise Fatorial QBPDQ-T0.sav

### Scale: ALL VARIABLES

**Case Processing Summary**

		N	%
Cases	Valid	132	100,0
	Excluded <sup>a</sup>	0	,0
	Total	132	100,0

a. Listwise deletion based on all variables in the procedure.

**Reliability Statistics**

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,950	,951	20

## Análise dos itens

**Item Statistics**

	Mean	Std. Deviation	N
Q1	1,66	1,104	132
Q2	1,73	1,224	132
Q3	1,70	1,096	132
Q4	1,45	1,114	132
Q5	2,34	1,165	132
Q6	2,52	1,188	132
Q7	1,68	1,250	132
Q8	1,31	1,308	132
Q9	2,26	1,434	132
Q10	1,60	1,413	132
Q11	1,19	1,261	132
Q12	2,38	1,637	132
Q13	,80	1,030	132
Q14	1,59	1,278	132
Q15	1,71	1,311	132
Q16	2,27	1,289	132
Q17	,87	1,044	132
Q18	1,77	1,408	132
Q19	2,13	1,339	132
Q20	2,80	1,396	132

**Inter-Item Correlation Matrix**

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Q1	1,000	,563	,635	,491	,418	,282	,502	,534
Q2	,563	1,000	,634	,476	,248	,157	,437	,425
Q3	,635	,634	1,000	,421	,384	,307	,544	,495
Q4	,491	,476	,421	1,000	,411	,399	,585	,522
Q5	,418	,248	,384	,411	1,000	,455	,489	,401
Q6	,282	,157	,307	,399	,455	1,000	,421	,361
Q7	,502	,437	,544	,585	,489	,421	1,000	,728
Q8	,534	,425	,495	,522	,401	,361	,728	1,000
Q9	,437	,362	,491	,463	,487	,413	,634	,690
Q10	,440	,325	,391	,425	,459	,476	,575	,563
Q11	,420	,390	,422	,488	,424	,407	,596	,603
Q12	,321	,323	,450	,400	,336	,329	,574	,479
Q13	,471	,387	,408	,430	,394	,341	,425	,459
Q14	,598	,465	,469	,505	,469	,343	,482	,460

Q15	,570	,374	,498	,528	,405	,396	,517	,449
Q16	,563	,361	,531	,475	,488	,377	,560	,498
Q17	,511	,468	,493	,509	,338	,313	,536	,566
Q18	,429	,352	,424	,389	,422	,188	,604	,533
Q19	,438	,376	,479	,509	,486	,303	,586	,639
Q20	,370	,347	,424	,417	,456	,323	,562	,503

**Inter-Item Correlation Matrix**

	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
Q1	,437	,440	,420	,321	,471	,598	,570	,563
Q2	,362	,325	,390	,323	,387	,465	,374	,361
Q3	,491	,391	,422	,450	,408	,469	,498	,531
Q4	,463	,425	,488	,400	,430	,505	,528	,475
Q5	,487	,459	,424	,336	,394	,469	,405	,488
Q6	,413	,476	,407	,329	,341	,343	,396	,377
Q7	,634	,575	,596	,574	,425	,482	,517	,560
Q8	,690	,563	,603	,479	,459	,460	,449	,498
Q9	1,000	,549	,522	,625	,391	,420	,487	,520
Q10	,549	1,000	,741	,630	,423	,479	,436	,507
Q11	,522	,741	1,000	,583	,541	,513	,440	,504
Q12	,625	,630	,583	1,000	,316	,345	,411	,538
Q13	,391	,423	,541	,316	1,000	,716	,523	,563
Q14	,420	,479	,513	,345	,716	1,000	,663	,632
Q15	,487	,436	,440	,411	,523	,663	1,000	,687
Q16	,520	,507	,504	,538	,563	,632	,687	1,000
Q17	,497	,544	,651	,524	,651	,624	,519	,559
Q18	,586	,643	,644	,572	,505	,532	,472	,569
Q19	,655	,645	,582	,615	,489	,531	,508	,643
Q20	,618	,631	,495	,645	,439	,526	,426	,637

**Inter-Item Correlation Matrix**

	Q17	Q18	Q19	Q20
Q1	,511	,429	,438	,370
Q2	,468	,352	,376	,347
Q3	,493	,424	,479	,424
Q4	,509	,389	,509	,417
Q5	,338	,422	,486	,456
Q6	,313	,188	,303	,323
Q7	,536	,604	,586	,562
Q8	,566	,533	,639	,503
Q9	,497	,586	,655	,618
Q10	,544	,643	,645	,631

Q11	,651	,644	,582	,495
Q12	,524	,572	,615	,645
Q13	,651	,505	,489	,439
Q14	,624	,532	,531	,526
Q15	,519	,472	,508	,426
Q16	,559	,569	,643	,637
Q17	1,000	,670	,553	,532
Q18	,670	1,000	,773	,713
Q19	,553	,773	1,000	,802
Q20	,532	,713	,802	1,000

**Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q1	34,09	306,587	,648	,615	,948
Q2	34,02	308,496	,532	,525	,950
Q3	34,05	306,853	,646	,602	,948
Q4	34,30	306,610	,641	,534	,948
Q5	33,41	307,862	,579	,445	,949
Q6	33,23	311,551	,474	,455	,951
Q7	34,07	298,293	,765	,707	,946
Q8	34,44	298,279	,727	,725	,947
Q9	33,49	294,908	,729	,661	,947
Q10	34,15	295,290	,732	,706	,947
Q11	34,56	299,241	,734	,705	,947
Q12	33,37	292,815	,667	,636	,948
Q13	34,95	308,615	,641	,614	,948
Q14	34,16	299,997	,706	,701	,947
Q15	34,04	300,495	,674	,630	,948
Q16	33,48	297,824	,750	,678	,947
Q17	34,88	304,886	,737	,685	,947
Q18	33,98	294,977	,742	,777	,947
Q19	33,62	294,695	,791	,805	,946
Q20	32,95	295,632	,735	,757	,947

## Teste de Normalidade das variáveis: Q1- Q20

```

EXAMINE VARIABLES=Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15
Q16 Q17 Q18 Q19 Q20
/PLOT NPFPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
    
```

### Explore

[DataSet1] C:\Users\Cristina\Desktop\Análise Fatorial QBPDQ-T0.sav

#### Descriptives

		Statistic	Std. Error
Q1	Mean	1,66	,096
	95% Confidence Interval for		
	Lower Bound	1,47	
	Upper Bound	1,85	
	5% Trimmed Mean	1,62	
	Median	2,00	
	Variance	1,219	
	Std. Deviation	1,104	
	Minimum	0	
	Maximum	4	
	Range	4	
	Interquartile Range	1	
	Skewness	,262	,211
	Kurtosis	-,273	,419
Q2	Mean	1,73	,107
	95% Confidence Interval for		
	Lower Bound	1,52	
	Upper Bound	1,94	
	5% Trimmed Mean	1,69	
	Median	2,00	
	Variance	1,498	
	Std. Deviation	1,224	
	Minimum	0	
	Maximum	5	
	Range	5	
	Interquartile Range	1	
	Skewness	,310	,211
	Kurtosis	-,527	,419

Q3	Mean	1,70	,095
	95% Confidence Interval for	Lower Bound	1,52
	Mean	Upper Bound	1,89
	5% Trimmed Mean	1,67	
	Median	2,00	
	Variance	1,202	
	Std. Deviation	1,096	
	Minimum	0	
	Maximum	4	
	Range	4	

### Descriptives

		Statistic	Std. Error
Q3	Interquartile Range	1	
	Skewness	,188	,211
	Kurtosis	-,397	,419
	Mean	1,45	,097
	95% Confidence Interval for	Lower Bound	1,26
	Mean	Upper Bound	1,64
	5% Trimmed Mean	1,39	
	Median	1,00	
	Variance	1,241	
	Std. Deviation	1,114	
Q4	Minimum	0	
	Maximum	4	
	Range	4	
	Interquartile Range	1	
	Skewness	,488	,211
	Kurtosis	-,265	,419
	Mean	2,34	,101
	95% Confidence Interval for	Lower Bound	2,14
	Mean	Upper Bound	2,54
	5% Trimmed Mean	2,36	
Median	2,00		
Variance	1,356		
Q5	Std. Deviation	1,165	
	Minimum	0	
	Maximum	5	
	Range	5	
	Interquartile Range	1	
	Skewness	,010	,211
	Kurtosis	-,600	,419

	Mean		2,52	,103
	95% Confidence Interval for	Lower Bound	2,32	
	Mean	Upper Bound	2,73	
Q6	5% Trimmed Mean		2,55	
	Median		2,00	
	Variance		1,412	
	Std. Deviation		1,188	

### Descriptives

		Statistic	Std. Error	
Q6	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	1		
	Skewness	-,193	,211	
	Kurtosis	-,249	,419	
	Mean	1,68	,109	
	95% Confidence Interval for	Lower Bound	1,47	
	Mean	Upper Bound	1,90	
	5% Trimmed Mean		1,63	
	Median		1,00	
	Variance		1,562	
Q7	Std. Deviation	1,250		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,554	,211	
	Kurtosis	-,417	,419	
	Mean	1,31	,114	
	95% Confidence Interval for	Lower Bound	1,09	
	Mean	Upper Bound	1,54	
	5% Trimmed Mean		1,21	
	Median		1,00	
Variance		1,712		
Q8	Std. Deviation	1,308		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,941	,211	
	Kurtosis	,289	,419	

	Mean		2,26	,125
Q9	95% Confidence Interval for	Lower Bound	2,01	
	Mean	Upper Bound	2,50	
	5% Trimmed Mean		2,23	

### Descriptives

		Statistic	Std. Error	
Q9	Median	2,00		
	Variance	2,055		
	Std. Deviation	1,434		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,246	,211	
	Kurtosis	-,638	,419	
	Mean	1,60	,123	
	95% Confidence Interval for	Lower Bound	1,36	
	Mean	Upper Bound	1,84	
	5% Trimmed Mean		1,51	
Q10	Median	2,00		
	Variance	1,998		
	Std. Deviation	1,413		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,644	,211	
	Kurtosis	-,340	,419	
	Mean	1,19	,110	
	95% Confidence Interval for	Lower Bound	,97	
	Mean	Upper Bound	1,41	
	5% Trimmed Mean		1,07	
Q11	Median	1,00		
	Variance	1,590		
	Std. Deviation	1,261		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	1,120	,211	
	Kurtosis	,789	,419	

Q12	Mean	2,38	,142
-----	------	------	------

**Descriptives**

			Statistic	Std. Error
Q12	95% Confidence Interval for Mean	Lower Bound	2,10	
		Upper Bound	2,66	
	5% Trimmed Mean		2,37	
	Median		2,00	
	Variance		2,680	
	Std. Deviation		1,637	
	Minimum		0	
	Maximum		5	
	Range		5	
	Interquartile Range		3	
	Skewness		,251	,211
	Kurtosis		-1,087	,419
	Mean		,80	,090
	Q13	95% Confidence Interval for Mean	Lower Bound	,63
Upper Bound			,98	
5% Trimmed Mean		,68		
Median		,00		
Variance		1,060		
Std. Deviation		1,030		
Minimum		0		
Maximum		4		
Range		4		
Interquartile Range		1		
Skewness		1,343	,211	
Kurtosis		1,512	,419	
Mean		1,59	,111	
Q14		95% Confidence Interval for Mean	Lower Bound	1,37
	Upper Bound		1,81	
	5% Trimmed Mean		1,52	
	Median		2,00	
	Variance		1,633	
	Std. Deviation		1,278	
	Minimum		0	
	Maximum		5	
	Range		5	
	Interquartile Range		1	

**Descriptives**

		Statistic	Std. Error
Q14	Skewness	,585	,211
	Kurtosis	-,159	,419
	Mean	1,71	,114
	95% Confidence Interval for		
	Mean	Lower Bound	1,49
		Upper Bound	1,94
	5% Trimmed Mean	1,66	
	Median	2,00	
Q15	Variance	1,718	
	Std. Deviation	1,311	
	Minimum	0	
	Maximum	5	
	Range	5	
	Interquartile Range	1	
	Skewness	,466	,211
	Kurtosis	-,559	,419
	Mean	2,27	,112
	95% Confidence Interval for		
	Mean	Lower Bound	2,04
		Upper Bound	2,49
	5% Trimmed Mean	2,24	
	Median	2,00	
Variance	1,662		
Q16	Std. Deviation	1,289	
	Minimum	0	
	Maximum	5	
	Range	5	
	Interquartile Range	2	
	Skewness	,359	,211
	Kurtosis	-,555	,419
	Mean	,87	,091
	95% Confidence Interval for		
	Mean	Lower Bound	,69
	Upper Bound	1,05	
Q17	5% Trimmed Mean	,76	
	Median	1,00	
	Variance	1,090	
	Std. Deviation	1,044	
	Minimum	0	

**Descriptives**

		Statistic	Std. Error	
Q17	Maximum	5		
	Range	5		
	Interquartile Range	1		
	Skewness	1,406	,211	
	Kurtosis	2,075	,419	
	Mean	1,77	,123	
	95% Confidence Interval for	Lower Bound	1,52	
	Mean	Upper Bound	2,01	
	5% Trimmed Mean	1,68		
	Median	2,00		
	Variance	1,983		
Q18	Std. Deviation	1,408		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,660	,211	
	Kurtosis	-,297	,419	
	Mean	2,13	,117	
	95% Confidence Interval for	Lower Bound	1,90	
	Mean	Upper Bound	2,36	
	5% Trimmed Mean	2,09		
Median	2,00			
Variance	1,792			
Q19	Std. Deviation	1,339		
	Minimum	0		
	Maximum	5		
	Range	5		
	Interquartile Range	2		
	Skewness	,265	,211	
	Kurtosis	-,554	,419	
	Mean	2,80	,122	
Q20	95% Confidence Interval for	Lower Bound	2,56	
	Mean	Upper Bound	3,04	
	5% Trimmed Mean	2,81		
	Median	3,00		

**Descriptives**

		Statistic	Std. Error
Q20	Variance	1,950	
	Std. Deviation	1,396	
	Minimum	0	
	Maximum	5	
	Range	5	
	Interquartile Range	2	
	Skewness	,014	,211
	Kurtosis	-,909	,419

#### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Q1	,212	132	,000	,891	132	,000
Q2	,177	132	,000	,912	132	,000
Q3	,205	132	,000	,903	132	,000
Q4	,186	132	,000	,888	132	,000
Q5	,191	132	,000	,926	132	,000
Q6	,185	132	,000	,920	132	,000
Q7	,230	132	,000	,903	132	,000
Q8	,207	132	,000	,850	132	,000
Q9	,208	132	,000	,922	132	,000
Q10	,176	132	,000	,882	132	,000
Q11	,242	132	,000	,831	132	,000
Q12	,190	132	,000	,906	132	,000
Q13	,297	132	,000	,755	132	,000
Q14	,171	132	,000	,900	132	,000
Q15	,176	132	,000	,906	132	,000
Q16	,218	132	,000	,921	132	,000
Q17	,253	132	,000	,781	132	,000
Q18	,191	132	,000	,897	132	,000
Q19	,197	132	,000	,928	132	,000
Q20	,170	132	,000	,928	132	,000

a. Lilliefors Significance Correction

## Coeficiente de correlação de Spearman: Q1- Q20

```

NONPAR CORR
  /VARIABLES=Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16
Q17 Q18 Q19 Q20
  /PRINT=SPEARMAN TWOTAIL SIG
  /MISSING=PAIRWISE.
  
```

### Nonparametric Correlations

[DataSet3] C:\Users\Cristina\Desktop\JULHO\_2012\Análise Factorial  
QBPDQ-T0.sav

**Correlations**

			Q1	Q2	Q3	Q4
Spearman's rho		Correlation Coefficient	1,000	,556	,641	,464
	Q1	Sig. (2-tailed)	.	,000	,000	,000
		N	132	132	132	132
		Correlation Coefficient	,556	1,000	,637	,444
	Q2	Sig. (2-tailed)	,000	.	,000	,000
		N	132	132	132	132
		Correlation Coefficient	,641	,637	1,000	,405
	Q3	Sig. (2-tailed)	,000	,000	.	,000
		N	132	132	132	132
		Correlation Coefficient	,464	,444	,405	1,000
	Q4	Sig. (2-tailed)	,000	,000	,000	.
		N	132	132	132	132
		Correlation Coefficient	,391	,223	,358	,388
	Q5	Sig. (2-tailed)	,000	,010	,000	,000
		N	132	132	132	132
		Correlation Coefficient	,304	,155	,309	,408
	Q6	Sig. (2-tailed)	,000	,076	,000	,000
		N	132	132	132	132
		Correlation Coefficient	,473	,417	,542	,538
	Q7	Sig. (2-tailed)	,000	,000	,000	,000
		N	132	132	132	132
		Correlation Coefficient	,514	,427	,513	,469
	Q8	Sig. (2-tailed)	,000	,000	,000	,000
		N	132	132	132	132

	Correlation Coefficient	,439	,349	,466	,437
Q9	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,420	,343	,383	,377
Q10	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,435	,434	,464	,465
Q11	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,339	,336	,469	,426
Q12	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

### Correlations

		Q5	Q6	Q7	Q8
	Correlation Coefficient	,391	,304	,473	,514
Q1	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,223	,155	,417	,427
Q2	Sig. (2-tailed)	,010	,076	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,358	,309	,542	,513
Q3	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,388	,408	,538	,469
Q4	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Spearman's rho	Correlation Coefficient	1,000	,473	,470	,371
Q5	Sig. (2-tailed)	.	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,473	1,000	,439	,383
Q6	Sig. (2-tailed)	,000	.	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,470	,439	1,000	,669
Q7	Sig. (2-tailed)	,000	,000	.	,000
	N	132	132	132	132
	Correlation Coefficient	,371	,383	,669	1,000
Q8	Sig. (2-tailed)	,000	,000	,000	.
	N	132	132	132	132

	Correlation Coefficient	,440	,402	,630	,706
Q9	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,437	,469	,562	,538
Q10	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,370	,396	,592	,547
Q11	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,317	,336	,575	,536
Q12	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

### Correlations

		Q9	Q10	Q11	Q12
Spearman's rho	Correlation Coefficient	,439	,420	,435	,339
	Q1 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,349	,343	,434	,336
	Q2 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,466	,383	,464	,469
	Q3 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,437	,377	,465	,426
	Q4 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,440	,437	,370	,317
	Q5 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,402	,469	,396	,336
	Q6 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,630	,562	,592	,575
	Q7 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,706	,538	,547	,536
	Q8 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

	Correlation Coefficient	1,000	,554	,516	,633
Q9	Sig. (2-tailed)	.	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,554	1,000	,684	,625
Q10	Sig. (2-tailed)	,000	.	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,516	,684	1,000	,610
Q11	Sig. (2-tailed)	,000	,000	.	,000
	N	132	132	132	132
	Correlation Coefficient	,633	,625	,610	1,000
Q12	Sig. (2-tailed)	,000	,000	,000	.
	N	132	132	132	132

### Correlations

		Q13	Q14	Q15	Q16
	Correlation Coefficient	,494	,603	,583	,572
Q1	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,425	,439	,383	,341
Q2	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,414	,457	,496	,529
Q3	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,430	,467	,500	,468
Q4	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Spearman's rho	Correlation Coefficient	,360	,463	,364	,478
Q5	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,335	,337	,371	,407
Q6	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,386	,460	,453	,540
Q7	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,426	,417	,413	,482
Q8	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

	Correlation Coefficient	,360	,395	,438	,476
Q9	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,412	,456	,410	,486
Q10	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,542	,489	,423	,474
Q11	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,294	,328	,401	,506
Q12	Sig. (2-tailed)	,001	,000	,000	,000
	N	132	132	132	132

### Correlations

		Q17	Q18	Q19	Q20
	Correlation Coefficient	,551	,438	,422	,382
Q1	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,457	,331	,331	,326
Q2	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,507	,410	,439	,428
Q3	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,508	,361	,502	,413
Q4	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Spearman's rho	Correlation Coefficient	,331	,389	,449	,449
Q5	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,300	,211	,325	,322
Q6	Sig. (2-tailed)	,000	,015	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,519	,600	,576	,561
Q7	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,548	,516	,657	,501
Q8	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

	Correlation Coefficient	,490	,575	,634	,612
Q9	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,486	,674	,655	,659
Q10	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,599	,659	,559	,506
Q11	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,522	,570	,625	,647
Q12	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

### Correlations

		Q1	Q2	Q3	Q4
Spearman's rho	Correlation Coefficient	,494	,425	,414	,430
Q13	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,603	,439	,457	,467
Q14	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,583	,383	,496	,500
Q15	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,572	,341	,529	,468
Q16	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,551	,457	,507	,508
Q17	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,438	,331	,410	,361
Q18	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,422	,331	,439	,502
Q19	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,382	,326	,428	,413
Q20	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

**Correlations**

		Q5	Q6	Q7	Q8
Spearman's rho	Correlation Coefficient	,360	,335	,386	,426
	Q13 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q14	Correlation Coefficient	,463	,337	,460	,417
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q15	Correlation Coefficient	,364	,371	,453	,413
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q16	Correlation Coefficient	,478	,407	,540	,482
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q17	Correlation Coefficient	,331	,300	,519	,548
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q18	Correlation Coefficient	,389	,211	,600	,516
	Sig. (2-tailed)	,000	,015	,000	,000
	N	132	132	132	132
Q19	Correlation Coefficient	,449	,325	,576	,657
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q20	Correlation Coefficient	,449	,322	,561	,501
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

**Correlations**

		Q9	Q10	Q11	Q12
Spearman's rho	Correlation Coefficient	,360	,412	,542	,294
	Q13 Sig. (2-tailed)	,000	,000	,000	,001
	N	132	132	132	132
Q14	Correlation Coefficient	,395	,456	,489	,328
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q15	Correlation Coefficient	,438	,410	,423	,401
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

	Correlation Coefficient	,476	,486	,474	,506
Q16	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,490	,486	,599	,522
Q17	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,575	,674	,659	,570
Q18	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,634	,655	,559	,625
Q19	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
	Correlation Coefficient	,612	,659	,506	,647
Q20	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

#### Correlations

		Q13	Q14	Q15	Q16
Spearman's rho	Correlation Coefficient	1,000	,672	,470	,510
	Q13	Sig. (2-tailed)	.	,000	,000
		N	132	132	132
Q14	Correlation Coefficient	,672	1,000	,637	,607
	Sig. (2-tailed)	,000	.	,000	,000
	N	132	132	132	132
Q15	Correlation Coefficient	,470	,637	1,000	,663
	Sig. (2-tailed)	,000	,000	.	,000
	N	132	132	132	132
Q16	Correlation Coefficient	,510	,607	,663	1,000
	Sig. (2-tailed)	,000	,000	,000	.
	N	132	132	132	132
Q17	Correlation Coefficient	,616	,580	,488	,525
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q18	Correlation Coefficient	,445	,500	,451	,524
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q19	Correlation Coefficient	,445	,497	,500	,616
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

	Correlation Coefficient	,420	,517	,422	,624
Q20	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132

### Correlations

		Q17	Q18	Q19	Q20
Spearman's rho	Correlation Coefficient	,616	,445	,445	,420
	Q13 Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q14	Correlation Coefficient	,580	,500	,497	,517
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q15	Correlation Coefficient	,488	,451	,500	,422
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q16	Correlation Coefficient	,525	,524	,616	,624
	Sig. (2-tailed)	,000	,000	,000	,000
	N	132	132	132	132
Q17	Correlation Coefficient	1,000	,666	,552	,541
	Sig. (2-tailed)	.	,000	,000	,000
	N	132	132	132	132
Q18	Correlation Coefficient	,666	1,000	,742	,700
	Sig. (2-tailed)	,000	.	,000	,000
	N	132	132	132	132
Q19	Correlation Coefficient	,552	,742	1,000	,800
	Sig. (2-tailed)	,000	,000	.	,000
	N	132	132	132	132
Q20	Correlation Coefficient	,541	,700	,800	1,000
	Sig. (2-tailed)	,000	,000	,000	.
	N	132	132	132	132

**APÊNDICE J:**  
***Output IBM SPSS Statistics – Fiabilidade teste- resteste***

GET

FILE='C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav'.

Warning. Command name: GET FILE  
SPSS Statistics data file  
"C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav"  
is written in a character encoding (ISO-8859-15)  
incompatible with the current LOCALE setting. It may not be readable.  
Consider changing LOCALE or setting UNICODE on. (DATA 1721)  
DATASET NAME DataSet1 WINDOW=FRONT.  
FREQUENCIES VARIABLES=QBPDQ\_T0 QBPDQ\_T1  
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN MODE  
/HISTOGRAM  
/ORDER=ANALYSIS.

## Frequências

[DataSet1]  
C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav

		QBPDQ_T0	QBPDQ_T1
N	Valid	104	104
	Missing	0	0
Mean		34,45	34,02
Median		33,00	33,50
Mode		33	26
Std. Deviation		17,275	13,611
Minimum		4	8
Maximum		80	80

GET

FILE='C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav'.

Warning. Command name: GET FILE  
SPSS Statistics data file  
"C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav"  
is written in a character encoding (ISO-8859-15)  
incompatible with the current LOCALE setting. It may not be readable.  
Consider changing LOCALE or setting UNICODE on. (DATA 1721)  
DATASET NAME DataSet3 WINDOW=FRONT.  
RELIABILITY  
/VARIABLES=QBPDQ\_T0 QBPDQ\_T1  
/SCALE('ALL VARIABLES') ALL  
/MODEL=ALPHA  
/ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95 TESTVAL=0.

## Reliability

[DataSet3] C:\Users\Cristina\Desktop\junho\_2012\Fiabilidade\_(1,2,3,4)\_QBPDQ.sav

### Scale: ALL VARIABLES

#### Case Processing Summary

		N	%
Cases	Valid	104	100,0
	Excluded <sup>a</sup>	0	,0
	Total	104	100,0

a. Listwise deletion based on all variables in the procedure.

#### Intraclass Correlation Coefficient

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Single Measures	,696 <sup>a</sup>	,581	,783	5,568	103	103
Average Measures	,820 <sup>c</sup>	,735	,878	5,568	103	103

#### Intraclass Correlation Coefficient

	F Test with True Value 0 <sup>b</sup>
	Sig
Single Measures	,000 <sup>a</sup>
Average Measures	,000 <sup>c</sup>

Two-way mixed effects model where people effects are random and measures effects are fixed.

a. The estimator is the same, whether the interaction effect is present or not.

b. Type C intraclass correlation coefficients using a consistency definition -the between-measure variance is excluded from the denominator variance.

c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

**APÊNDICE K:**

***Output IBM SPSS Statistics* – Verificação de condições prévias à  
estimativa do EPM e DMD**

## Estatística Descritiva das Pontuações da QBPDS T0 e T2

```
GET
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_1.sav'
.
DATASET NAME DataSet4 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa
/STATISTICS=MEAN STDDEV MIN MAX.
```

### Descriptives

```
[DataSet4]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_1.sav
```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	1	55	55	55,00	.
Incapacidade T2	1	80	80	80,00	.
Diferença de Pontuação QBPDS (T2-T0)	1	-25,00	-25,00	-25,0000	.
Diferença relativa QBPDS T0-T2(%)	1	45,45	45,45	45,4545	.
Valid N (listwise)	1				

```
GET
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_2.sav'
.
DATASET NAME DataSet5 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa
/STATISTICS=MEAN STDDEV MIN MAX.
```

### Descriptives

```
[DataSet5]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_2.sav
```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	8	7	56	27,88	19,298
Incapacidade T2	8	6	59	29,25	21,393
Diferença de Pontuação QBPDS (T2-T0)	8	-10,00	5,00	-1,3750	5,57898
Diferença relativa QBPDS T0-T2(%)	8	-40,00	31,82	,9843	25,65487
Valid N (listwise)	8				

```

DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet4.
DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet5.
GET

```

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_3.sav'
```

```

.
DATASET NAME DataSet6 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa
  /STATISTICS=MEAN STDDEV MIN MAX.

```

## Descriptives

```

[DataSet6]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_3.sav

```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	22	12	80	32,41	16,998
Incapacidade T2	22	8	81	30,73	18,686
Diferença de Pontuação QBPD (T2-T0)	22	-22,00	21,00	1,6818	11,08146
Diferença relativa QBPD T0-T2(%)	22	-68,00	88,00	-1,8886	39,71859
Valid N (listwise)	22				

```

DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet6.
GET

```

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_4.sav'
```

```

.
DATASET NAME DataSet7 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa
  /STATISTICS=MEAN STDDEV MIN MAX.

```

## Descriptives

```

[DataSet7]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_4.sav

```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	13	7	70	42,15	21,169
Incapacidade T2	13	7	59	34,62	16,404
Diferença de Pontuação QBPD (T2-T0)	13	-6,00	31,00	7,5385	11,16255
Diferença relativa QBPD T0-T2(%)	13	-58,82	71,43	-11,9445	33,69264
Valid N (listwise)	13				

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_5.sav'  
.  
DATASET NAME DataSet8 WINDOW=FRONT.  
DESCRIPTIVES VARIABLES=QBPDO_T0 QBPDO_T2 Dif_T2_T0 Dif_Relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

[DataSet8]  
C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_5.sav

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	35	14	70	37,86	15,426
Incapacidade T2	35	5	57	27,54	15,538
Diferença de Pontuação QBPDS (T2-T0)	35	-11,00	33,00	10,3143	9,34871
Diferença relativa QBPDS T0-T2(%)	35	-84,85	42,86	-27,8603	26,67376
Valid N (listwise)	35				

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_6.sav'  
.  
DATASET NAME DataSet9 WINDOW=FRONT.  
DESCRIPTIVES VARIABLES=QBPDO_T0 QBPDO_T2 Dif_T2_T0 Dif_Relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

[DataSet9]  
C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_6.sav

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	31	4	95	40,26	21,259
Incapacidade T2	31	3	45	24,10	13,773
Diferença de Pontuação QBPDS (T2-T0)	31	-8,00	80,00	16,1613	15,91875
Diferença relativa QBPDS T0-T2(%)	31	-84,21	29,63	-37,9958	24,67746
Valid N (listwise)	31				

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_7.sav'  
.  
DATASET NAME DataSet10 WINDOW=FRONT.  
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

[DataSet10]  
C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_7.sav

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	10	6	56	28,00	12,910
Incapacidade T2	10	2	43	17,50	15,436
Diferença de Pontuação QBPD (T2-T0)	10	-12,00	29,00	10,5000	13,00641
Diferença relativa QBPD T0-T2(%)	10	-91,30	38,71	-33,9423	50,85912
Valid N (listwise)	10				

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\CorrelaçãoDifQBPDQ_T0_T2_ju  
nho.sav'.  
DATASET NAME DataSet12 WINDOW=FRONT.  
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T0_T2 Dif_relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

[DataSet12]  
C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDQ\_T0\_T2\_junho.sa  
v

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	120	4	95	36,60	18,367
Incapacidade T2	120	2	81	27,72	17,108
Diferença QBPD T0-T2	120	-25,00	80,00	8,8833	13,37324
Diferença Relativa QBPD T0-T2	120	-91,30	88,00	-21,9659	35,10438
Valid N (listwise)	120				

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_insuc  
sso.sav'.  
DATASET NAME DataSet13 WINDOW=FRONT.  
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T2_T0 Dif_Relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

```
[DataSet13]  
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_insuc  
v
```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	44	7	80	34,98	19,090
Incapacidade T2	44	6	81	32,73	19,428
Dif_T2_T0	44	-25,00	31,00	2,2500	11,32013
Diferença relativa QBPDS T0-T2(%)	44	-68,00	88,00	-3,2613	35,73074
Valid N (listwise)	44				

```
DATASET ACTIVATE DataSet3.  
DESCRIPTIVES VARIABLES=QBPDQ_T0 QBPDQ_T2 Dif_T0_T2 Dif_relativa  
/STATISTICS=MEAN STDDEV MIN MAX.
```

## Descriptives

```
[DataSet3]  
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_suc  
cesso.sav
```

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation
Incapacidade T0	76	4	95	37,54	17,997
Incapacidade T2	76	2	57	24,82	14,991
Diferença de Pontuação QBPD (T2-T0)	76	-12,00	80,00	12,7237	13,02060
Diferença relativa QBPDS T0-T2 (%)	76	-91,30	42,86	-32,7948	29,98750
Valid N (listwise)	76				

## Comparação dos resultados de incapacidade funcional em T0 e T2, para ambos os grupos

```

DATASET ACTIVATE DataSet13.
EXAMINE VARIABLES=QBPDQ_T0 QBPDQ_T2
/PLOT BOXPLOT HISTOGRAM NPLOT
/COMPARE GROUPS
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

```

### Explore

```

[DataSet13]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_insucesso.sa
v

```

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Incapacidade T0	44	100,0%	0	0,0%	44	100,0%
Incapacidade T2	44	100,0%	0	0,0%	44	100,0%

#### Descriptives

			Statistic	Std. Error
Incapacidade T0	Mean		34,98	2,878
	95% Confidence Interval for	Lower Bound	29,17	
	Mean	Upper Bound	40,78	
	5% Trimmed Mean		34,33	
	Median		33,00	
	Variance		364,441	
	Std. Deviation		19,090	
	Minimum		7	
	Maximum		80	
	Range		73	
	Interquartile Range		33	
	Skewness		,407	,357
	Kurtosis		-,778	,702

	Mean		32,73	2,929
	95% Confidence Interval for Mean	Lower Bound	26,82	
		Upper Bound	38,63	
	5% Trimmed Mean		31,63	
	Median		33,50	
	Variance		377,459	
Incapacidade T2	Std. Deviation		19,428	
	Minimum		6	
	Maximum		81	
	Range		75	
	Interquartile Range		33	
	Skewness		,503	,357
	Kurtosis		-,209	,702

#### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Incapacidade T0	,113	44	,195	,955	44	,087
Incapacidade T2	,107	44	,200 <sup>*</sup>	,944	44	,033

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

```

DATASET ACTIVATE DataSet3.
EXAMINE VARIABLES=QBPDQ_T0 QBPDQ_T2
  /PLOT BOXPLOT NPLOT
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.

```

## Explore

[DataSet3]  
C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_sucesso.sav

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Incapacidade T0	76	100,0%	0	0,0%	76	100,0%
Incapacidade T2	76	100,0%	0	0,0%	76	100,0%

### Descriptives

			Statistic	Std. Error
Incapacidade T0	Mean		37,54	2,064
	95% Confidence Interval for Mean	Lower Bound	33,43	
		Upper Bound	41,65	
	5% Trimmed Mean		36,91	
	Median		34,00	
	Variance		323,878	
	Std. Deviation		17,997	
	Minimum		4	
	Maximum		95	
	Range		91	
	Interquartile Range		24	
	Skewness		,671	,276
	Kurtosis		,486	,545
	Mean		24,82	1,720
Incapacidade T2	95% Confidence Interval for Mean	Lower Bound	21,39	
		Upper Bound	28,24	
	5% Trimmed Mean		24,39	
	Median		22,00	
	Variance		224,739	
	Std. Deviation		14,991	
	Minimum		2	
	Maximum		57	
	Range		55	
	Interquartile Range		28	
	Skewness		,384	,276
	Kurtosis		-,917	,545

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Incapacidade T0	,143	76	,001	,963	76	,024
Incapacidade T2	,088	76	,200*	,950	76	,005

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

```

DATASET ACTIVATE DataSet13.
NPAR TESTS
  /WILCOXON=QBPDQ_T0 WITH QBPDQ_T2 (PAIRED)
  /MISSING ANALYSIS.

```

## NPar Tests

```

[DataSet13]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_insucesso.sa
v

```

## Wilcoxon Signed Ranks Test

		Ranks		
		N	Mean Rank	Sum of Ranks
Incapacidade T2 -	Negative Ranks	20 <sup>a</sup>	25,35	507,00
	Positive Ranks	21 <sup>b</sup>	16,86	354,00
Incapacidade T0	Ties	3 <sup>c</sup>		
Total		44		

- a. Incapacidade T2 < Incapacidade T0
- b. Incapacidade T2 > Incapacidade T0
- c. Incapacidade T2 = Incapacidade T0

### Test Statistics<sup>a</sup>

	Incapacidade T2 - Incapacidade T0
Z	-,992 <sup>b</sup>
Asymp. Sig. (2-tailed)	,321

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

```

DATASET ACTIVATE DataSet3.
NPAR TESTS
  /WILCOXON=QBPDQ_T0 WITH QBPDQ_T2 (PAIRED)
  /MISSING ANALYSIS.

```

## NPar Tests

[DataSet3]

C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_sucesso.sav

### Wilcoxon Signed Ranks Test

		Ranks		
		N	Mean Rank	Sum of Ranks
Incapacidade T2 - Incapacidade T0	Negative Ranks	67 <sup>a</sup>	39,40	2640,00
	Positive Ranks	7 <sup>b</sup>	19,29	135,00
Ties		2 <sup>c</sup>		
Total		76		

a. Incapacidade T2 < Incapacidade T0

b. Incapacidade T2 > Incapacidade T0

c. Incapacidade T2 = Incapacidade T0

#### Test Statistics<sup>a</sup>

	Incapacidade T2 - Incapacidade T0
Z	-6,749 <sup>b</sup>
Asymp. Sig. (2-tailed)	,000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

## Equivalência dos grupos “sem alterações” e “melhoria clínica”, no ponto de partida

```

DATASET ACTIVATE DataSet13.
EXAMINE VARIABLES=EVA_T0
  /PLOT BOXPLOT HISTOGRAM NPLOT
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
  
```

### Explore

[DataSet13]  
 C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_insucesso.sav

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Intensidade Dor T0	44	100,0%	0	0,0%	44	100,0%

#### Descriptives

		Statistic	Std. Error
Intensidade Dor T0	Mean	5,75	,309
	95% Confidence Interval for Mean		
	Lower Bound	5,13	
	Upper Bound	6,37	
	5% Trimmed Mean	5,77	
	Median	5,00	
	Variance	4,192	
	Std. Deviation	2,047	
	Minimum	2	
	Maximum	9	
	Range	7	
	Interquartile Range	4	
	Skewness	,099	,357
Kurtosis	-,893	,702	

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Intensidade Dor T0	,166	44	,004	,937	44	,019

a. Lilliefors Significance Correction

```

DATASET ACTIVATE DataSet3.
EXAMINE VARIABLES=EVA_T0
  /PLOT BOXPLOT NPLOT
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
    
```

### Explore

```

[DataSet3]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_sucesso.sav
    
```

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Intensidade Dor T0	76	100,0%	0	0,0%	76	100,0%

### Descriptives

		Statistic	Std. Error
Intensidade Dor T0	Mean	5,59	,244
	95% Confidence Interval for Mean		
	Lower Bound	5,11	
	Upper Bound	6,08	
	5% Trimmed Mean	5,63	
	Median	6,00	
	Variance	4,511	
	Std. Deviation	2,124	
	Minimum	1	
	Maximum	10	
	Range	9	
	Interquartile Range	3	
	Skewness	-,165	,276
Kurtosis	-,542	,545	

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Intensidade Dor T0	,129	76	,003	,968	76	,054

a. Lilliefors Significance Correction

NPAR TESTS

```
/M-W= QBPDQ_T0 BY Sucesso_Insucesso_PGIC567(1 2)
/MISSING ANALYSIS.
```

### NPar Tests

[DataSet1]

C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDQ\_T0\_T2\_30\_06\_12.sav

### Mann-Whitney Test

#### Ranks

	Categorias de Percepção	N	Mean Rank	Sum of Ranks
	Melhoria (5,6,7)			
	Sucesso (5,6,7)	76	62,54	4753,00
Incapacidade T0	Insucesso (1,2,3,4)	44	56,98	2507,00
	Total	120		

#### Test Statistics<sup>a</sup>

	Incapacidade T0
Mann-Whitney U	1517,000
Wilcoxon W	2507,000
Z	-,844
Asymp. Sig. (2-tailed)	,398

a. Grouping Variable: Categorias de Percepção Melhoria (5,6,7)

DATASET ACTIVATE DataSet12.

NPAR TESTS

```
/M-W= EVA_T0 BY Sucesso_Insucesso_PGIC567(1 2)
/STATISTICS=DESCRIPTIVES
/MISSING ANALYSIS.
```

## NPar Tests

[DataSet12]

C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDQ\_T0\_T2\_junho.sav

### Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Intensidade Dor T0	120	5,65	2,089	1	10
Categorias de Percepção Melhoria (5,6,7)	120	1,37	,484	1	2

## Mann-Whitney Test

### Ranks

	Categorias de Percepção Melhoria (5,6,7)	N	Mean Rank	Sum of Ranks
Intensidade Dor T0	Sucesso (5,6,7)	76	59,99	4559,50
	Insucesso (1,2,3,4)	44	61,38	2700,50
	Total	120		

### Test Statistics<sup>a</sup>

	Intensidade Dor T0
Mann-Whitney U	1633,500
Wilcoxon W	4559,500
Z	-,212
Asymp. Sig. (2-tailed)	,832

a. Grouping Variable: Categorias de Percepção  
Melhoria (5,6,7)

FREQUENCIES VARIABLES=Cat\_Idade Genero IMC\_categorias Cat\_EstCivil  
 Cat\_HabLit Cat\_ActProf Cat\_SitProf Cat\_Duração\_Dor Local\_Dor Medicacao  
 Faltas\_Trabalho Baixa  
 /ORDER=ANALYSIS.

## Frequencies

[DataSet13]

C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_insucesso.sav

**Statistics**

		Cat Idade_Mediana	Género	Índice de Massa Corporal (3 categorias)	Categorias Estado Civil	Categorias Hab Literárias
N	Valid	44	44	42	44	44
	Missing	0	0	2	0	0

**Statistics**

		Categorias Atividade Profissional	Categorias Situação Profissional	Duração Da Dor	Localização da dor	Medicação
N	Valid	44	44	44	44	44
	Missing	0	0	0	0	0

**Statistics**

		Faltas ao trabalho	Baixa remunerada
N	Valid	44	44
	Missing	0	0

## Frequency Table

**Cat Idade\_Mediana**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	< ou igual a 49 anos	22	50,0	50,0
	> 49 anos	22	50,0	100,0
Total	44	100,0	100,0	

**Género**

	Frequency	Percent	Valid Percent	Cumulative Percent
Masculino	9	20,5	20,5	20,5
Valid Feminino	35	79,5	79,5	100,0
Total	44	100,0	100,0	

**Índice de Massa Corporal (3 categorias)**

	Frequency	Percent	Valid Percent	Cumulative Percent
Peso normal	13	29,5	31,0	31,0
Valid Excesso de peso	29	65,9	69,0	100,0
Total	42	95,5	100,0	
Missing 0	2	4,5		
Total	44	100,0		

**Categorias Estado Civil**

	Frequency	Percent	Valid Percent	Cumulative Percent
Com suporte familiar	32	72,7	72,7	72,7
Valid Sem suporte familiar	12	27,3	27,3	100,0
Total	44	100,0	100,0	

**Categorias Hab Literárias**

	Frequency	Percent	Valid Percent	Cumulative Percent
Ensino Primário	15	34,1	34,1	34,1
Valid Ensino Básico e Secundário	21	47,7	47,7	81,8
Ensino Superior	8	18,2	18,2	100,0
Total	44	100,0	100,0	

**Categorias Atividade Profissional**

	Frequency	Percent	Valid Percent	Cumulative Percent
Intelectual	6	13,6	13,6	13,6
Valid Técnico	21	47,7	47,7	61,4
Operacional	17	38,6	38,6	100,0
Total	44	100,0	100,0	

### categorias Situação Profissional

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Ativo	31	70,5	70,5	70,5
Não Ativo	10	22,7	22,7	93,2
Doméstica(o)	3	6,8	6,8	100,0
Total	44	100,0	100,0	

### Duração Da Dor

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid < ou igual a 24 meses	15	34,1	34,1	34,1
>24 meses	29	65,9	65,9	100,0
Total	44	100,0	100,0	

### Localização da dor

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Sem dor referida para a perna	18	40,9	40,9	40,9
Com dor referida para a perna	26	59,1	59,1	100,0
Total	44	100,0	100,0	

### Medicação

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	24	54,5	54,5	54,5
Sim	20	45,5	45,5	100,0
Total	44	100,0	100,0	

### Faltas ao trabalho

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	33	75,0	75,0	75,0
Sim	11	25,0	25,0	100,0
Total	44	100,0	100,0	

**Baixa remunerada**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	36	81,8	81,8	81,8
Valid Sim	8	18,2	18,2	100,0
Total	44	100,0	100,0	

```

DATASET ACTIVATE DataSet3.
FREQUENCIES VARIABLES=Cat_Idade Genero IMC_categorias Cat_EstCivil
Cat_HabLit Cat_ActProf Cat_SitProf Cat_Duração_Dor Local_Dor Medicacao
Faltas_Trabalho Baixa
/ORDER=ANALYSIS.
    
```

**Frequencies**

```

[DataSet3]
C:\Users\Cristina\Desktop\junho_2012\PGIC_pontuações\PGIC_sucesso.sav
    
```

**Statistics**

	Cat Idade_Mediana	Género	Índice de Massa Corporal (3 categorias)	Categorias Estado Civil	Categorias Hab Literárias
N Valid	76	76	76	76	76
Missing	0	0	0	0	0

**Statistics**

	Categorias Atividade Profissional	Categorias Situação Profissional	Duração Da Dor	Localização da dor	Medicação
N Valid	74	76	76	76	76
Missing	2	0	0	0	0

**Statistics**

	Faltas ao trabalho	Baixa remunerada
N Valid	76	76
Missing	0	0

## Frequency Table

**Cat Idade\_Mediana**

	Frequency	Percent	Valid Percent	Cumulative Percent
< ou igual a 49 anos	39	51,3	51,3	51,3
Valid > 49 anos	37	48,7	48,7	100,0
Total	76	100,0	100,0	

**Género**

	Frequency	Percent	Valid Percent	Cumulative Percent
Masculino	24	31,6	31,6	31,6
Valid Feminino	52	68,4	68,4	100,0
Total	76	100,0	100,0	

**Índice de Massa Corporal (3 categorias)**

	Frequency	Percent	Valid Percent	Cumulative Percent
Peso normal	32	42,1	42,1	42,1
Valid Excesso de peso	44	57,9	57,9	100,0
Total	76	100,0	100,0	

**Categorias Estado Civil**

	Frequency	Percent	Valid Percent	Cumulative Percent
Com suporte familiar	45	59,2	59,2	59,2
Valid Sem suporte familiar	31	40,8	40,8	100,0
Total	76	100,0	100,0	

**Categorias Hab Literárias**

	Frequency	Percent	Valid Percent	Cumulative Percent
Ensino Primário	16	21,1	21,1	21,1
Valid Ensino Básico e Secundário	40	52,6	52,6	73,7
Ensino Superior	20	26,3	26,3	100,0
Total	76	100,0	100,0	

**Categorias Atividade Profissional**

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	Intelectual	21	27,6	28,4	28,4
	Técnico	28	36,8	37,8	66,2
	Operacional	25	32,9	33,8	100,0
	Total	74	97,4	100,0	
Missing	System	2	2,6		
Total		76	100,0		

**Categorias Situação Profissional**

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	Ativo	45	59,2	59,2	59,2
	Não Ativo	20	26,3	26,3	85,5
	Doméstica(o)	11	14,5	14,5	100,0
	Total	76	100,0	100,0	

**Duração Da Dor**

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	< ou igual a 24 meses	26	34,2	34,2	34,2
	>24 meses	50	65,8	65,8	100,0
	Total	76	100,0	100,0	

**Localização da dor**

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	Sem dor referida para a perna	39	51,3	51,3	51,3
	Com dor referida para a perna	37	48,7	48,7	100,0
	Total	76	100,0	100,0	

### Medicação

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	40	52,6	52,6	52,6
Valid Sim	36	47,4	47,4	100,0
Total	76	100,0	100,0	

### Faltas ao trabalho

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	51	67,1	67,1	67,1
Valid Sim	25	32,9	32,9	100,0
Total	76	100,0	100,0	

### Baixa remunerada

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Não	59	77,6	77,6	77,6
Valid Sim	17	22,4	22,4	100,0
Total	76	100,0	100,0	

GET

FILE='C:\Users\Cristina\Desktop\junho\_2012\QBPDQ-Suc\_(5,6,7).sav'.

Warning. Command name: GET FILE

SPSS Statistics data file "C:\Users\Cristina\Desktop\junho\_2012\QBPDQ-Suc\_(5,6,7).sav" is written in a character encoding (ISO-8859-15) incompatible with the current LOCALE setting. It may not be readable. Consider changing LOCALE or setting UNICODE on. (DATA 1721)

DATASET NAME DataSet1 WINDOW=FRONT.

CROSSTABS

```
/TABLES=Cat_Idade Genero IMC_categorias Cat_EstCivil Cat_HabLit  
Cat_ActProf Cat_SitProf Cat_Duração_Dor Local_Dor Medicacao  
Faltas_Trabalho Total_Faltas Duracao_Faltas Baixa BY Sucesso_Insucesso  
/FORMAT=AVALUE TABLES  
/STATISTICS=CHISQ PHI  
/CELLS=COUNT EXPECTED ROW COLUMN TOTAL SRESID  
/COUNT ROUND CELL.
```

## Crosstabs

[DataSet1] C:\Users\Cristina\Desktop\junho\_2012\QBPDQ-Suc\_(5,6,7).sav

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Cat Idade_Mediana * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Género * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Índice de Massa Corporal (3 categorias) * Percepção Melhoria	118	98,3%	2	1,7%	120	100,0%
Categorias Estado Civil * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Categorias Hab Literárias * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Categorias Atividade Profissional * Percepção Melhoria	118	98,3%	2	1,7%	120	100,0%
Categorias Situação Profissional * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Duração Da Dor * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Localização da dor * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Medicação * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Faltas ao trabalho * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%
Total de faltas * Percepção Melhoria	35	29,2%	85	70,8%	120	100,0%
Duração das faltas * Percepção Melhoria	34	28,3%	86	71,7%	120	100,0%
Baixa remunerada * Percepção Melhoria	120	100,0%	0	0,0%	120	100,0%

### Cat Idade\_Mediana \* Percepção Melhoria

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,107 <sup>a</sup>	1	,744		
Continuity Correction <sup>b</sup>	,019	1	,891		
Likelihood Ratio	,107	1	,744		
Fisher's Exact Test				,849	,446
Linear-by-Linear Association	,106	1	,745		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 21,14.

b. Computed only for a 2x2 table

**Género \* Percepção Melhoria**

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,451 <sup>a</sup>	1	,228		
Continuity Correction <sup>b</sup>	,983	1	,322		
Likelihood Ratio	1,491	1	,222		
Fisher's Exact Test				,288	,161
Linear-by-Linear Association	1,439	1	,230		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,83.

b. Computed only for a 2x2 table

**Índice de Massa Corporal (3 categorias) \* Percepção Melhoria**

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,094 <sup>a</sup>	1	,148		
Continuity Correction <sup>b</sup>	1,558	1	,212		
Likelihood Ratio	2,135	1	,144		
Fisher's Exact Test				,168	,105
Linear-by-Linear Association	2,076	1	,150		
N of Valid Cases	118				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 15,64.

b. Computed only for a 2x2 table

### Categorias Estado Civil \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,831 <sup>a</sup>	1	,176		
Continuity Correction <sup>b</sup>	1,333	1	,248		
Likelihood Ratio	1,867	1	,172		
Fisher's Exact Test				,234	,124
Linear-by-Linear Association	1,816	1	,178		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 15,41.

b. Computed only for a 2x2 table

### Categorias Hab Literárias \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,013 <sup>a</sup>	2	,222
Likelihood Ratio	2,958	2	,228
Linear-by-Linear Association	2,569	1	,109
N of Valid Cases	120		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 10,03.

### Categorias Atividade Profissional \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,057 <sup>a</sup>	2	,217
Likelihood Ratio	3,234	2	,199
Linear-by-Linear Association	1,954	1	,162

N of Valid Cases	118		
------------------	-----	--	--

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 9,84.

### Categorias Situação Profissional \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1,783 <sup>a</sup>	2	,410
Likelihood Ratio	1,885	2	,390
Linear-by-Linear Association	1,702	1	,192
N of Valid Cases	120		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 5,02.

### Duração Da Dor \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,015 <sup>a</sup>	1	,901		
Continuity Correction <sup>b</sup>	,000	1	1,000		
Likelihood Ratio	,015	1	,902		
Fisher's Exact Test				1,000	,528
Linear-by-Linear Association	,015	1	,902		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 14,69.

b. Computed only for a 2x2 table

### Localização da dor \* Percepção Melhoria

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,855 <sup>a</sup>	1	,355		

Continuity Correction <sup>b</sup>	,539	1	,463		
Likelihood Ratio	,857	1	,354		
Fisher's Exact Test				,446	,232
Linear-by-Linear Association	,848	1	,357		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 20,43.

b. Computed only for a 2x2 table

### Medicação \* Percepção Melhoria

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,166 <sup>a</sup>	1	,684		
Continuity Correction <sup>b</sup>	,047	1	,829		
Likelihood Ratio	,166	1	,684		
Fisher's Exact Test				,707	,415
Linear-by-Linear Association	,164	1	,685		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 20,07.

b. Computed only for a 2x2 table

### Faltas ao trabalho \* Percepção Melhoria

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,623 <sup>a</sup>	1	,430		
Continuity Correction <sup>b</sup>	,338	1	,561		
Likelihood Ratio	,632	1	,427		
Fisher's Exact Test				,534	,282
Linear-by-Linear Association	,618	1	,432		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 12,90.

b. Computed only for a 2x2 table

### Total de faltas \* Percepção Melhoria

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,788 <sup>a</sup>	3	,852
Likelihood Ratio	,745	3	,863
Linear-by-Linear Association	,090	1	,764
N of Valid Cases	35		

a. 6 cells (75,0%) have expected count less than 5. The minimum expected count is 1,26.

**Duração das faltas \* Percepção Melhoria**

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,724 <sup>a</sup>	4	,948
Likelihood Ratio	,693	4	,952
Linear-by-Linear Association	,138	1	,710
N of Valid Cases	34		

a. 8 cells (80,0%) have expected count less than 5. The minimum expected count is ,97.

**Baixa remunerada \* Percepção Melhoria**

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,202 <sup>a</sup>	1	,653		
Continuity Correction <sup>b</sup>	,046	1	,830		
Likelihood Ratio	,204	1	,651		
Fisher's Exact Test				,815	,420
Linear-by-Linear Association	,200	1	,655		
N of Valid Cases	120				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,96.

b. Computed only for a 2x2 table

## Fiabilidade teste-reteste T0-T2 (grupo “sem melhorias”) vs Fiabilidade teste-reteste T0-T2b (sub grupo “PGIC 2,3,4)

```

DATASET ACTIVATE DataSet13.
RELIABILITY
/VARIABLES=QBPDQ_T0 QBPDQ_T2
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/ICC=MODEL(RANDOM) TYPE(CONSISTENCY) CIN=95 TESTVAL=0.
    
```

### Reliability

[DataSet13]  
 C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_insucesso.sav

### Scale: ALL VARIABLES

**Case Processing Summary**

		N	%
Cases	Valid	44	100,0
	Excluded <sup>a</sup>	0	,0
	Total	44	100,0

a. Listwise deletion based on all variables in the procedure.

**Intraclass Correlation Coefficient**

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Single Measures	,827 <sup>a</sup>	,705	,902	10,579	43	43
Average Measures	,905	,827	,948	10,579	43	43

**Intraclass Correlation Coefficient**

	F Test with True Value 0 <sup>b</sup>
	Sig
Single Measures	,000 <sup>a</sup>
Average Measures	,000

Two-way random effects model where both people effects and measures effects are random.

a. The estimator is the same, whether the interaction effect is present or not.

b. Type C intraclass correlation coefficients using a consistency definition - the between-measure variance is excluded from the denominator variance.

```

RELIABILITY
/VARIABLES=QBPDQ_T0 QBPDQ_T2
    
```

```

/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/ICC=MODEL(RANDOM) TYPE(CONSISTENCY) CIN=95 TESTVAL=0.

```

## Reliability

[DataSet3]  
C:\Users\Cristina\Desktop\junho\_2012\PGIC\_pontuações\PGIC\_insucesso2,3,4.s  
av

### Scale: ALL VARIABLES

**Case Processing Summary**

		N	%
Cases	Valid	43	100,0
	Excluded <sup>a</sup>	0	,0
	Total	43	100,0

a. Listwise deletion based on all variables in the procedure.

**Intraclass Correlation Coefficient**

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Single Measures	,837 <sup>a</sup>	,719	,908	11,296	42	42
Average Measures	,911	,837	,952	11,296	42	42

**Intraclass Correlation Coefficient**

	F Test with True Value 0 <sup>b</sup>
	Sig
Single Measures	,000 <sup>a</sup>
Average Measures	,000

Two-way random effects model where both people effects and measures effects are random.

a. The estimator is the same, whether the interaction effect is present or not.

b. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.

**APÊNDICE L:**  
**Recodificação das variáveis sócio-demográficas e clínicas**

Tabela – Recodificação das variáveis sócio-demográficas e clínicas

<b>Categorias</b>	<b>Classes agregadas</b>
<b>Categorias da Idade</b>	
≤ 49 anos	
> 49 anos	
<b>Categorias do IMC</b>	
Peso normal	Peso normal; Abaixo do peso normal
Excesso de peso	Excesso de peso
<b>Categorias do Estado civil</b>	
Com suporte familiar	Casado(a); União de Facto
Sem suporte familiar	Solteiro(a); Viúvo(a); Divorciado(a)
<b>Categorias do Nível de escolaridade</b>	
Ensino Primário	Ensino Primário
Ensino Básico e Secundário	Ensino Básico; Ensino Secundário incompleto; Ensino Secundário completo
Ensino Superior	Ensino Superior incompleto; Ensino Superior completo.
<b>Categorias da Atividade Profissional</b>	
Intelectual	Forças Armadas; Representantes do Poder legislativo; Especialistas das Atividades intelectuais; Estudante
Técnico	Técnicos e profissões de nível intermédio; Pessoal administrativo; Serviços pessoais;
Operacional	Qualificados da agricultura, pesca e floresta; Qualificados da indústria, construção e artífices; Operadores de instalações e máquinas; Trabalhadores não qualificados
<b>Situação profissional atual</b>	
Ativo	A trabalhar a tempo inteiro; A trabalhar a tempo parcial
Não Ativo	Incapaz de trabalhar; Desempregado (a); Reformado (a)
Doméstica (o)	Doméstica (o)
<b>Duração da dor</b>	
≤ 24 meses	3-6 meses; 6-12 meses; 12-24 meses
> 24 meses	Mais de 24 meses



**APÊNDICE M:**

***Output IBM SPSS Statistics – Verificação dos efeitos de “chão” e de  
“teto”***

```

DATASET ACTIVATE DataSet12.
FREQUENCIES VARIABLES=Amplitude
  /ORDER=ANALYSIS.

```

## Frequências

```

[DataSet12]
C:\Users\Cristina\Desktop\junho_2012\CorrelaçãoDifQBPDQ_T0_T2_junho.sa
v

```

### Statistics

Amplitude da escala

N	Valid	120
	Missing	0

### Amplitude da escala

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	19-81	100	83,3	83,3
	<19	19	15,8	99,2
	>81	1	,8	100,0
Total	120	100,0	100,0	

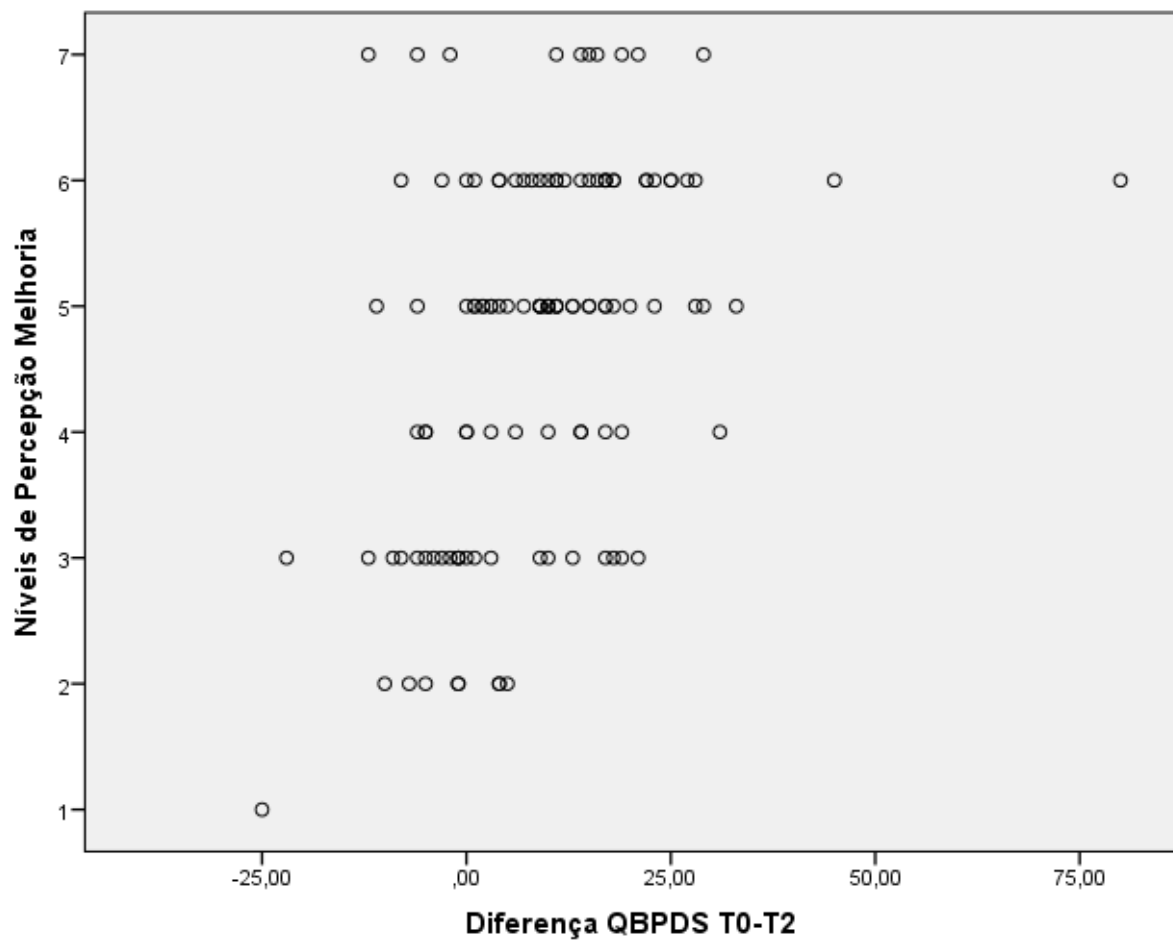
**APÊNDICE N:**  
***Output IBM SPSS Statistics – Poder de Resposta e DMCI***

## Coeficiente Correlação Âncora

```
GRAPH  
  /SCATTERPLOT(BIVAR)=Dif_T0_T2 WITH PGIC_T2  
  /MISSING=LISTWISE.
```

### Graph

```
[DataSet1]  
C:\Users\Cristina\Desktop\junho_2012\CorrelaçãoDifQBPDQ_T0_T2_30_06_12  
.sav
```



```
NONPAR CORR  
  /VARIABLES=PGIC_T2 Dif_T0_T2  
  /PRINT=SPEARMAN TWOTAIL NOSIG  
  /MISSING=PAIRWISE.
```

## Nonparametric Correlations

[DataSet12]

C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDS\_T0\_T2\_junho.sav

**Correlations**

		Níveis de Perceção Melhoria
Spearman's rho	Correlation Coefficient	1,000
	Níveis de Perceção Melhoria Sig. (2-tailed)	.
	N	120
	Correlation Coefficient	,426**
Diferença QBPDS T0-T2	Sig. (2-tailed)	,000
	N	120

**Correlations**

		Diferença QBPDS T0-T2
Spearman's rho	Correlation Coefficient	,426
	Níveis de Perceção Melhoria Sig. (2-tailed)	,000
	N	120
	Correlation Coefficient	1,000**
Diferença QBPDS T0-T2	Sig. (2-tailed)	.
	N	120

\*\* . Correlation is significant at the 0.01 level (2-tailed).

## Curva ROC

GET

```
FILE='C:\Users\Cristina\Desktop\junho_2012\CorrelaçãoDifQBPDS_T0_T2_30_06_12.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
ROC Dif_T0_T2 BY Sucesso_Insucesso (1)
  /PLOT=CURVE(REFERENCE)
  /PRINT=SE COORDINATES
  /CRITERIA=CUTOFF(INCLUDE) TESTPOS(LARGE) DISTRIBUTION(FREE) CI(95)
  /MISSING=EXCLUDE.
```

## ROC Curve

[DataSet1]

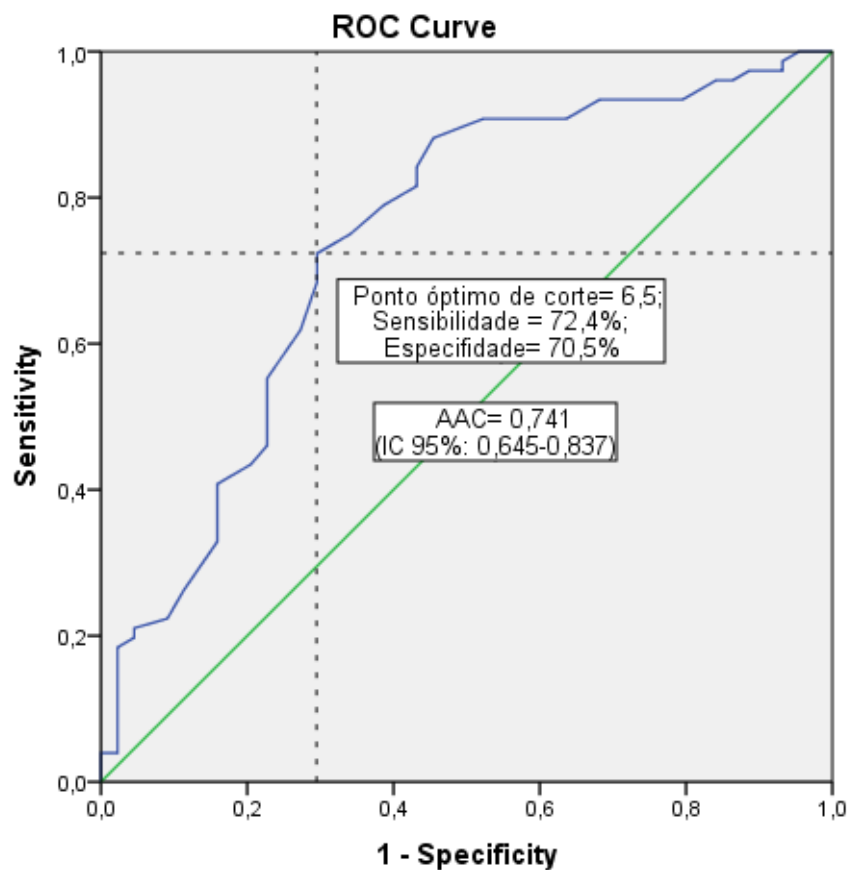
C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDQ\_T0\_T2\_30\_06\_12.sav

### Case Processing Summary

Categorias de Percepção	Valid N (listwise)
Melhoria	
Positive <sup>a</sup>	76
Negative	44

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is Sucesso (5,6,7).



Diagonal segments are produced by ties.

### Area Under the Curve

Test Result Variable(s): Diferença QBPDS T0-T2

Area	Std. Error <sup>a</sup>	Asymptotic Sig. <sup>b</sup>	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,741	,049	,000	,645	,837

The test result variable(s): Diferença QBPDS T0-T2 has at least one tie between the positive atual state group and the negative atual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

### Coordinates of the Curve

Test Result Variable(s): Diferença QBPDS T0-T2

Positive if Greater Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
-26,0000	1,000	1,000
-23,5000	1,000	,977
-17,0000	1,000	,955
-11,5000	,987	,932
-10,5000	,974	,932
-9,5000	,974	,909
-8,5000	,974	,886
-7,5000	,961	,864
-6,5000	,961	,841
-5,5000	,934	,795
-4,5000	,934	,705
-3,5000	,934	,682
-2,5000	,921	,659
-1,5000	,908	,636
-,5000	,908	,523
,5000	,882	,455
1,5000	,842	,432
2,5000	,816	,432
3,5000	,789	,386
4,5000	,750	,341
5,5000	,737	,318
6,5000	,724	,295
7,5000	,697	,295
8,5000	,684	,295

9,5000	,618	,273
10,5000	,553	,227
11,5000	,474	,227
12,5000	,461	,227
13,5000	,434	,205
14,5000	,408	,159
15,5000	,355	,159
16,5000	,329	,159
17,5000	,263	,114
18,5000	,224	,091
19,5000	,211	,045
20,5000	,197	,045
21,5000	,184	,023

**Coordinates of the Curve**

Test Result Variable(s): Diferença QBPDS T0-T2

Positive if Greater Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
22,5000	,158	,023
24,0000	,132	,023
26,0000	,105	,023
27,5000	,092	,023
28,5000	,066	,023
30,0000	,039	,023
32,0000	,039	,000
39,0000	,026	,000
62,5000	,013	,000
81,0000	,000	,000

The test result variable(s): Diferença QBPDS T0-T2 has at least one tie between the positive atual state group and the negative atual state group.

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

## Curva ROC (Diferença relativa)

```
ROC Dif_relativa BY Sucesso_Insucesso_PGIC567 (1)
/PLOT=CURVE(REFERENCE)
/PRINT=SE COORDINATES
/CRITERIA=CUTOFF(INCLUDE) TESTPOS(SMALL) DISTRIBUTION(FREE) CI(95)
/MISSING=EXCLUDE.
```

### ROC Curve

[DataSet12]

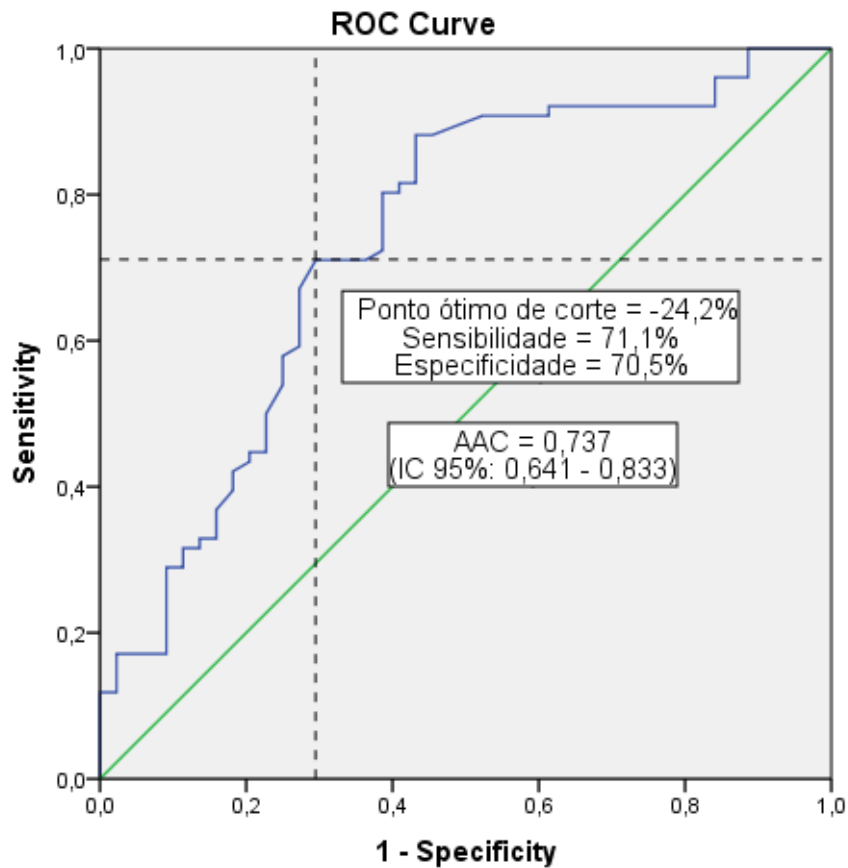
C:\Users\Cristina\Desktop\junho\_2012\CorrelaçãoDifQBPDQ\_T0\_T2\_junho.sav

#### Case Processing Summary

Categorias de Percepção	Valid N (listwise)
Melhoria (5,6,7)	
Positive <sup>a</sup>	76
Negative	44

Smaller values of the test result variable(s) indicate stronger evidence for a positive atual state.

a. The positive atual state is Sucesso (5,6,7).



Diagonal segments are produced by ties.

#### Area Under the Curve

Test Result Variable(s): Diferença Relativa QBPDS T0-T2

Area	Std. Error <sup>a</sup>	Asymptotic Sig. <sup>b</sup>	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,737	,049	,000	,641	,833

The test result variable(s): Diferença Relativa QBPDS T0-T2 has at least one tie between the positive actual state group and the negative actual state group.

Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

**Coordinates of the Curve**

Test Result Variable(s): Diferença Relativa

QBPDS T0-T2

Positive if Less Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
-92,3000	,000	,000
-88,8300	,013	,000
-85,6050	,026	,000
-84,5300	,039	,000
-83,5350	,053	,000
-80,6200	,066	,000
-77,0700	,079	,000
-73,8200	,092	,000
-70,3150	,105	,000
-68,3750	,118	,000
-67,3350	,118	,023
-66,1900	,145	,023
-64,3350	,158	,023
-62,4300	,171	,023
-60,9500	,171	,045
-59,4100	,171	,068
-58,2550	,171	,091
-57,4150	,197	,091
-55,5700	,211	,091
-53,4700	,224	,091
-52,3300	,237	,091
-51,4400	,250	,091
-50,9650	,263	,091
-50,3850	,276	,091
-49,3600	,289	,091
-48,0450	,289	,114
-46,6600	,303	,114
-45,3150	,316	,114
-44,5600	,316	,136
-44,3650	,329	,136
-44,2050	,329	,159
-44,0600	,342	,159
-42,5900	,355	,159
-40,5900	,368	,159
-39,8400	,395	,182
-39,0200	,408	,182
-37,9300	,421	,182

**Coordinates of the Curve**

Test Result Variable(s): Diferença Relativa

QBPDS T0-T2

Positive if Less Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
-36,9300	,434	,205
-35,7750	,447	,205
-35,0950	,447	,227
-34,6900	,461	,227
-33,8550	,474	,227
-32,4550	,487	,227
-31,4150	,500	,227
-30,6250	,539	,250
-29,6550	,553	,250
-28,9400	,566	,250
-28,2200	,579	,250
-27,6850	,592	,273
-27,3850	,605	,273
-27,1500	,632	,273
-26,3700	,645	,273
-25,6750	,658	,273
-25,3200	,671	,273
-24,1650	,711	,295
-22,3800	,711	,318
-21,1300	,711	,341
-20,7100	,711	,364
-19,8550	,724	,386
-19,0450	,737	,386
-18,7700	,750	,386
-17,6200	,763	,386
-16,3350	,776	,386
-15,5000	,789	,386
-12,3400	,803	,386
-9,2850	,803	,409
-8,0150	,816	,409
-6,4550	,816	,432
-4,9700	,829	,432
-4,0100	,842	,432
-3,7750	,855	,432
-3,1000	,868	,432
-2,2100	,882	,432
-,9600	,882	,455

### Coordinates of the Curve

Test Result Variable(s): Diferença Relativa

QBPDS T0-T2

Positive if Less Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
,6250	,908	,523
3,2550	,908	,545
5,6300	,908	,591
6,7500	,908	,614
7,9150	,921	,614
8,7950	,921	,636
11,5750	,921	,659
14,0900	,921	,682
14,7200	,921	,705
15,2650	,921	,727
16,0250	,921	,750
17,8600	,921	,773
19,3100	,921	,795
22,6050	,921	,818
26,4550	,921	,841
28,4500	,934	,841
29,6800	,947	,841
30,2500	,961	,841
31,2950	,961	,864
32,5750	,961	,886
36,0200	,974	,886
40,7850	,987	,886
44,1550	1,000	,886
47,7250	1,000	,909
51,0850	1,000	,932
61,8000	1,000	,955
79,7150	1,000	,977
89,0000	1,000	1,000

The test result variable(s): Diferença Relativa QBPDS T0-T2 has at least one tie between the positive atual state group and the negative atual state group.

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

## Curva ROC (pontuações iniciais baixas)

### ROC Curve

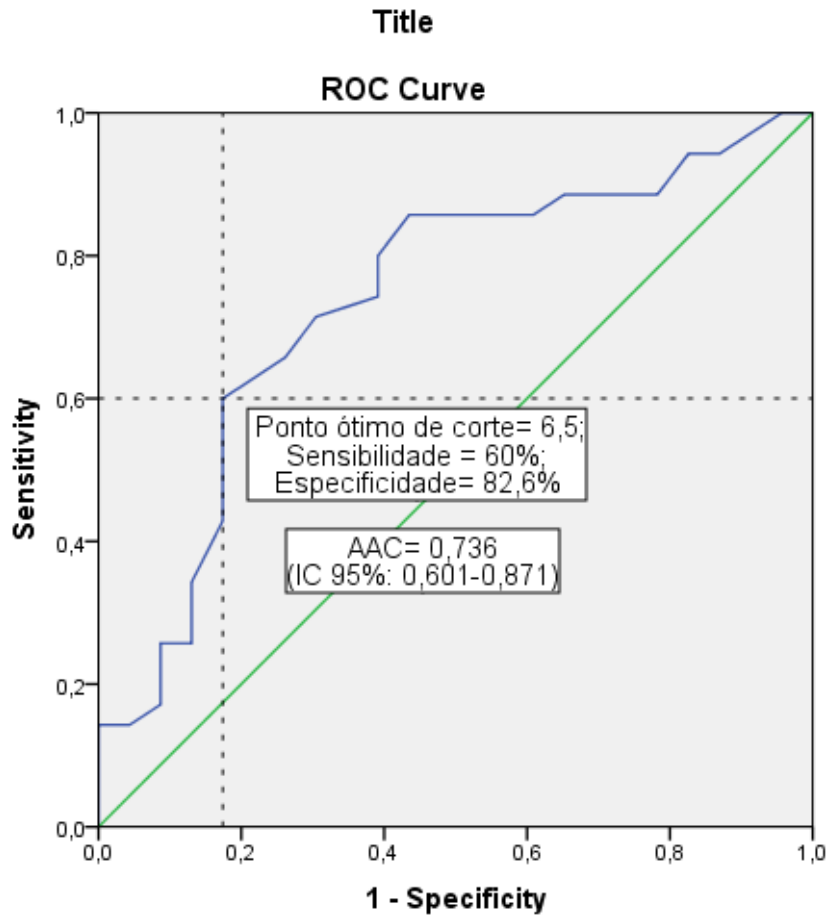
[Curva\_ROC\_menor\_34]

#### Case Processing Summary

Percepção Melhoria (5,6,7)	Valid N (listwise)
Positive <sup>a</sup>	35
Negative	23

Larger values of the test result variable(s) indicate stronger evidence for a positive atual state.

a. The positive atual state is Sucesso (5,6,7).



Diagonal segments are produced by ties.

#### Area Under the Curve

Test Result Variable(s): Diferença QBPDS (T0-T2)

Area	Std. Error <sup>a</sup>	Asymptotic Sig. <sup>b</sup>	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,736	,069	,003	,601	,871

The test result variable(s): Diferença QBPDS (T0-T2) has at least one tie between the positive actual state group and the negative actual state group.

Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

#### Coordinates of the Curve

Test Result Variable(s): Diferença QBPDS (T0-T2)

Positive if Greater Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
-23,0000	1,000	1,000
-17,0000	1,000	,957
-10,0000	,971	,913
-7,5000	,943	,870
-6,5000	,943	,826
-5,5000	,886	,783
-4,5000	,886	,696
-3,0000	,886	,652
-1,5000	,857	,609
-,5000	,857	,435
,5000	,800	,391
2,0000	,743	,391
3,5000	,714	,304
4,5000	,657	,261
5,5000	,629	,217
6,5000	,600	,174
7,5000	,571	,174
8,5000	,543	,174
9,5000	,429	,174
10,5000	,343	,130
12,0000	,257	,130
14,0000	,257	,087
15,5000	,200	,087
16,5000	,171	,087
17,5000	,143	,043
18,5000	,143	,000
20,0000	,114	,000
22,0000	,086	,000
24,0000	,057	,000
26,5000	,029	,000
29,0000	,000	,000

The test result variable(s): Diferença QBPDS (T0-T2) has at least one tie between the positive atual state group and the negative atual state group.

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

## Curva ROC (pontuações iniciais altas)

### ROC Curve

[DataSet1]

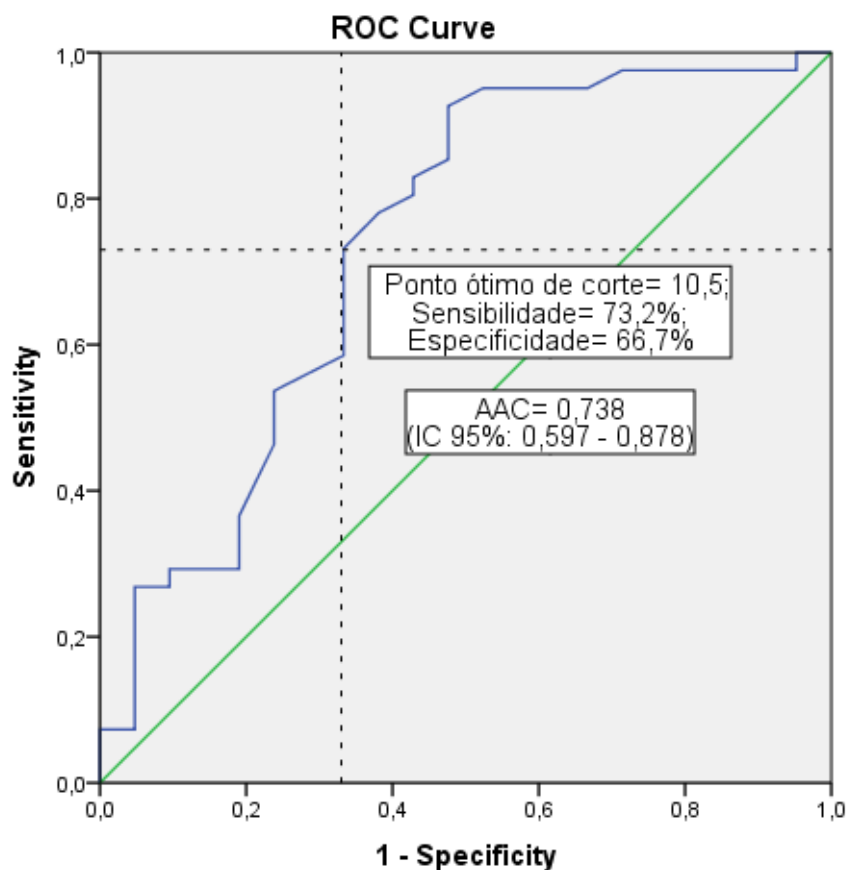
C:\Users\Cristina\Desktop\Cristina\_estudo2\_Modelo\_Predictivo.sav

#### Case Processing Summary

Percepção Melhoria (5,6,7)	Valid N (listwise)
Positive <sup>a</sup>	41
Negative	21

Larger values of the test result variable(s) indicate stronger evidence for a positive atual state.

a. The positive atual state is Sucesso (5,6,7).



Diagonal segments are produced by ties.

#### Area Under the Curve

Test Result Variable(s): Diferença QBPDS (T0-T2)

Area	Std. Error <sup>a</sup>	Asymptotic Sig. <sup>b</sup>	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,738	,072	,002	,597	,878

The test result variable(s): Diferença QBPDS (T0-T2) has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

#### Coordinates of the Curve

Test Result Variable(s): Diferença QBPDS (T0-T2)

Positive if Greater Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
-26,0000	1,000	1,000
-18,0000	1,000	,952

-10,5000	,976	,952
-9,5000	,976	,905
-7,5000	,976	,857
-5,5000	,976	,810
-4,0000	,976	,714
-2,0000	,951	,667
-,5000	,951	,619
,5000	,951	,524
1,5000	,927	,476
2,5000	,878	,476
3,5000	,854	,476
5,5000	,829	,429
8,0000	,805	,429
9,5000	,780	,381
10,5000	,732	,333
11,5000	,659	,333
12,5000	,634	,333
13,5000	,585	,333
14,5000	,537	,238
15,5000	,488	,238
16,5000	,463	,238
17,5000	,366	,190
18,5000	,293	,190
19,5000	,293	,095
20,5000	,268	,095
21,5000	,268	,048
22,5000	,220	,048
24,0000	,195	,048
26,0000	,171	,048
27,5000	,146	,048
28,5000	,122	,048
30,0000	,073	,048
32,0000	,073	,000
39,0000	,049	,000
62,5000	,024	,000
81,0000	,000	,000

The test result variable(s): Diferença QBPDS (T0-T2) has at least one tie between the positive atual state group and the negative atual state group.

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

**ANEXO I:**  
**Autorização da comissão de ética e administração do SESARAM,**  
**E.P.E**



REGIÃO AUTÓNOMA DA MADEIRA  
SERVIÇO DE SAÚDE DA REGIÃO AUTÓNOMA DA MADEIRA, E.P.E.

**Comissão de Ética para a Saúde  
do  
SESARAM,EPE**

**Parecer nº 72/2011**

V. M. A.  
6/12/11

Sobre o estudo " EFEITOS DA INTERVENÇÃO DA FISIOTERAPIA, AO NÍVEL DA DOR, CAPACIDADE FUNCIONAL E DE PERCEÇÃO GLOBAL DE MELHORIA DOS INDIVÍDUOS COM DOR LOMBRA CRÓNICA "

**A - RELATÓRIO**

**A.1.** A Comissão de Ética para a Saúde (CES) do Serviço de Saúde da Região Autónoma da Madeira (SESARAM) iniciou o Processo nº72 . 2011, após solicitação de Parecer enviado pela Direcção Clínica do SESARAM em documento datado de 10/11/2011 , relativa ao estudo " **Efeitos da intervenção da fisioterapia, ao nível da dor, capacidade funcional e de percepção global de melhoria dos indivíduos com dor lombar crónica** " apresentada pela Técnica de Diagnóstico e Terapêutica , no âmbito do mestrado em Fisioterapia, sob orientação do Professor Eduardo Cruz

**A.2.** Fazem parte do processo de avaliação os seguintes documentos: protocolo de investigação, modelo de consentimento informado, 8 questionários de auto preenchimento: Questionário de Caracterização Sócio demográfica e Clínica que inclui Escala Visual Analógica, Quebec Back Pain Disability Scale, Patterns of Activity Measure-Pain, Questionário de Baecke, Roland Morris Questionnaire, Patient Global Impression of Change, Tampa Acale of Kinesiophobia, International Physical Activity Questionnaire.

**A.3.** Trata-se de um estudo observacional multicentrico, prospectivo, por aplicação de questionário estruturado, junto de utentes referenciados para Fisioterapia que cumpram os critérios de inclusão definidos, que aceitem participar e subscrevam um formulário de consentimento informado . O preenchimento do questionário será feito em suporte de papel em 3 momentos distintos.

**A.4.** Os objectivos do estudo são: "Estudar os efeitos da Fisioterapia", assim recolher variáveis demográficas e clinicas dos utentes e registar o tipo de tratamento efectuado pelos Fisioterapeutas.



REGIÃO AUTÓNOMA DA MADEIRA  
SERVIÇO DE SAÚDE DA REGIÃO AUTÓNOMA DA MADEIRA, E.P.E.

**B- IDENTIFICAÇÃO DAS QUESTÕES COM EVENTUAIS IMPLICAÇÕES ÉTICAS**

- B.1. Estão dadas garantias de confidencialidade e anonimato dos participantes.
- B.2. O documento de informação ao sujeito/consentimento informado oferece informação bastante e é claro.
- B.3. Reconhece-se a pertinência do estudo e interesse prático nos resultados esperados, sendo que a metodologia utilizada salvaguarda os direitos dos participantes.
- B.4. Aceita-se como razoável a previsão de que a participação se traduz numa interferência pouco significativa com o regular funcionamento do serviço.

**C - CONCLUSÃO**

Face ao exposto, a CES delibera emitir parecer favorável á autorização deste estudo, nos precisos termos em que o mesmo foi apresentado.

Aprovado em reunião do dia 14 de Novembro de 2011, por unanimidade.

O Presidente da CES/SESARAM, EPE

(Ricardo Jorge Figueira Silva Santos)

**ANEXO II:**  
**Autorização da Santa Casa da Misericórdia de Machico.**



Santa Casa da  
**Misericórdia  
de Machico**

**Exm<sup>a</sup>. Fisioterapeuta,  
Dra. Ana Cristina Pires Vieira**

V/Referência

V/comunicação

N/Referência  
11145/ERS

N/Data  
2011-12-13

**Assunto: Pedido de Autorização para o Desenvolvimento e Projecto de Investigação em Fisioterapia**

De acordo com a solicitação acima referida considera-se autorizada a aplicação do projecto de Investigação em Fisioterapia.

Salientamos que o tratamento dos dados e os seus resultados devem ser apresentados à mesa administrativa da Instituição antes da sua publicação.

O Provedor

(Luís Carlos Araújo Delgado)



Lar Agostinho  
Cupertino da Câmara



Centro  
de Dia



Centro Médico  
e de Reabilitação



Centro Comunitário  
da Bemposta



Residência Sénior  
Quinta Cova do Milho



Quinta Pedagógica  
Cova do Milho

Sede - Rua do Desembarcadouro, nº 50, 9200-144 Machico - Tel.: 291 969 140 - Fax: 291 969 145 - Telm.: 969 570 756 - 919 584 489

Centro Comunitário da Bemposta - Complexo Habitacional da Bemposta - BL A2 Piso 0, 9200-230 Água de Pena - Tel.: 291 523 007

Residência Sénior Quinta Cova do Milho - Caminho do Poiso, Santo da Serra, 9100-262 Santa Cruz - Tel.: 291 552 401 - 291 552 403 - Fax: 291 552 259

E-mail: [scmm@netmadeira.com](mailto:scmm@netmadeira.com) • [www.misericordiademachico.com](http://www.misericordiademachico.com) • NIF: 511 014 244