

UNIVERSIDADE FEDERAL DO MARANHÃO  
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO FÍSICA  
MESTRADO ACADÊMICO

TRADUÇÃO, ADAPTAÇÃO TRANSCULTURAL E VALIDAÇÃO DO *UPPER LIMB  
FUNCTIONAL INDEX (ULFI)* PARA O PORTUGUÊS BRASILEIRO

**Aluno:** Henrique Yuji Takahasi

**Orientador:** Prof. Dr. Almir Vieira Dibai Filho

São Luís, 2020

HENRIQUE YUJI TAKAHASI

TRADUÇÃO, ADAPTAÇÃO TRANSCULTURAL E VALIDAÇÃO DO *UPPER LIMB  
FUNCTIONAL INDEX (ULFI)* PARA O PORTUGUÊS BRASILEIRO

Dissertação apresentada ao Programa de Pós-Graduação em Educação Física da  
Universidade Federal do Maranhão, para defesa para a obtenção do Título de  
Mestre em Educação Física.

Área de Concentração: Biodinâmica do Movimento Humano

Linha de Pesquisa: Atividade Física relacionada à Saúde Humana

Orientador: Prof. Dr. Almir Vieira Dibai Filho

São Luís, 2020

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a)  
autor(a). Núcleo Integrado de Bibliotecas/UFMA

Takahasi, Henrique Yuji.

Tradução, adaptação transcultural e validação do Upper Limb Functional Index ULFI para o português brasileiro / Henrique Yuji Takahasi. - 2020.

68 f.

Orientador(a): Almir Vieira Dibai Filho.

Dissertação (Mestrado) - Programa de Pós-graduação em Educação Física/ccbs, Universidade Federal do Maranhão, São Luís, 2020.

1. Análise fatorial. 2. Estudo de validação. 3. Medidas de resultado relatados pelo paciente. 4. Membro superior. 5. Reprodutibilidade dos testes. I. Dibai Filho, Almir Vieira. II. Título.

HENRIQUE YUJI TAKAHASI

TRADUÇÃO, ADAPTAÇÃO TRANSCULTURAL E VALIDAÇÃO DO *UPPER LIMB FUNCTIONAL INDEX* (ULFI) PARA O PORTUGUÊS BRASILEIRO

Dissertação apresentada ao Programa de Pós-Graduação em Educação Física da Universidade Federal do Maranhão, para defesa para a obtenção do Título de Mestre em Educação Física.

A Banca Examinadora da defesa da Dissertação de Mestrado apresentada em sessão pública, considerou o candidato aprovado em: \_\_\_/\_\_\_/\_\_\_\_.

---

Prof. Dr. Almir Vieira Dibai Filho (Orientador)  
Universidade Federal do Maranhão

---

Profa. Dra. Marisa de Cássia Registro Fonseca (Examinadora-Externo)  
Universidade de São Paulo

---

Prof. Dr. Thiago Teixeira Mendes (Examinador)  
Universidade Federal do Maranhão

---

Prof. Dr. Christian Emmanuel Torres Cabido (Examinador)  
Universidade Federal do Maranhão

São Luís  
2020

## **DEDICATÓRIA**

Dedico este trabalho a todos os pacientes da rede Sarah que buscam reabilitação, alívio de seu sofrimento e melhora de sua qualidade de vida. A eles devemos, como profissionais da saúde, estar preparados para avaliar e prover tratamento humanizado e eficaz baseado na melhor evidência científica disponível.

## **AGRADECIMENTOS**

Agradeço o meu orientador Prof. Dr. Almir Vieira Dibai Filho pelo acolhimento inicial daquele ainda candidato que sonhava em ingressar no mestrado, após três tentativas anteriores frustradas. Agradeço a atenção e sensibilidade para a escolha de um projeto que aliasse o interesse, a experiência e a exequibilidade dentro do ambiente de trabalho. Isso certamente facilitou e me motivou ainda mais. E por fim, agradeço a disponibilidade, a paciência e ajuda durante todas as etapas deste projeto que tornaram esta caminhada muito mais tranquila. Espero que possamos ainda colher muitos frutos esta parceria.

Agradeço minha querida esposa Eliana Harumi Morioka Takahasi e minhas filhas Larissa Akemi Takahasi e Carolina Yumi Takahasi pelo amor, alegrias, apoio incondicional e pela paciência nos momentos ausentes. Vocês me completam e me inspiram a ser uma pessoa melhor.

Agradeço meus pais pelo carinho, pelo exemplo, incentivo e oportunidade de educação e preparação para a vida. Agradeço minhas irmãs pelo apoio mesmo à distância.

Agradeço os meus colegas fisioterapeutas da rede SARAH Bianca Rojas, Camila Pflueger, Ceane Lessa, Felipe Barreto, Fernanda Thomaz, Lucianne Rhoden e Thiago Almeida pelo companheirismo e pelo auxílio inestimável na coleta dos dados.

Agradeço aos membros da banca examinadora Profa. Dra. Marisa de Cássia Registro Fonseca, Prof. Dr. Thiago Teixeira Mendes e Prof. Dr. Christian Emmanuel Torres Cabido pelas contribuições valiosas ao conteúdo e escrita do projeto.

Agradeço ao Prof. Dr. Antônio Augusto Moura da Silva do Programa de Pós-Graduação em Saúde Coletiva, cuja disciplina optativa cursada foi determinante para a definição da metodologia utilizada neste projeto.

Agradeço à coordenação do Programa de Pós-Graduação em Educação Física e a todo o corpo docente do Programa pela dedicação e pela oportunidade de enriquecimento intelectual e pessoal durante todas as disciplinas.

Agradeço aos colegas da turma de mestrado e grupo de pesquisa REMOVI pelo companheirismo e contribuições na construção deste projeto durante as várias disciplinas cursadas e reuniões.

## Resumo

**Objetivo:** O objetivo deste estudo foi traduzir, adaptar culturalmente e validar o *Upper Limb Functional Index* (ULFI) para o português brasileiro. **Métodos:** Foi realizada tradução e adaptação transcultural do ULFI conforme recomendações internacionais e a versão pré-final do ULFI foi aplicada em 30 pacientes com disfunções musculoesqueléticas do membro superior para avaliar a compreensão do questionário e estabelecer a versão final do ULFI. A versão final do ULFI foi aplicada em 200 pacientes para verificar a validade estrutural. Os demais questionários do estudo, isto é, *Quick Disabilities of the Arm, Shoulder and Hand* (QuickDash), *36-item Short Form Health Survey* (SF-36) e Escala Numérica da Dor (END) foram aplicados em 180 pacientes para testar a validade de construto por meio da correlação de Spearman ( $\rho$ ). A estrutura interna do ULFI foi avaliada por análise fatorial exploratória e confirmatória com os índices de ajuste Qui-quadrado ( $X^2$ ), Qui-quadrado dividido pelos graus de liberdade ( $X^2/gl$ ), *Root Mean Square Error of Approximation* (RMSEA), *Comparative Fit Index* (CFI) e *Tucker Lewis Index* (TLI). Uma subamostra de 51 pacientes preencheu o ULFI novamente em um intervalo de tempo médio de 9 dias para avaliar a confiabilidade teste-reteste pelo coeficiente de correlação intraclassa (CCI), erro padrão da medida (EPM) e diferença mínima detectável (DMD). A consistência interna do ULFI foi testada pelo alfa de Cronbach ( $\alpha$ ). Foram avaliados também os efeitos *floor* e *ceiling*. **Resultados:** No processo de tradução e adaptação transcultural, nenhum item da versão pré-final do questionário apresentou problemas de compreensão por mais de 20% dos pacientes resultando na versão final do ULFI. A análise fatorial identificou uma estrutura unidimensional do ULFI com índices de ajuste aceitáveis (RMSEA = 0,063, CFI = 0,918, TLI = 0,910). O ULFI demonstrou excelente confiabilidade (CCI moderada = 0,909) e consistência interna ( $\alpha$  = 0,897). O EPM foi de 6,11 e a DMD foi de 14,26. O ULFI apresentou alta correlação com o QuickDash ( $\rho$  = -0,721), com a END ( $\rho$  = -0,529) e os domínios capacidade funcional ( $\rho$  = 0,529), aspectos físicos ( $\rho$  = 0,572), dor ( $\rho$  = 0,562), estado geral de saúde ( $\rho$  = 0,505) e aspectos sociais ( $\rho$  = 0,557) do SF-36, e baixa com os domínios vitalidade ( $\rho$  = 0,361), aspectos emocionais ( $\rho$  = 0,440) e saúde mental ( $\rho$  = 0,309) do SF-36. Não houve efeitos *floor* e *ceiling*. **Discussão:** Conforme hipotetizado, o ULFI demonstrou alta confiabilidade, consistência interna adequada e alta correlação com o QuickDash, outra medida de resultado relatada pelo paciente que também avalia a função de todo o membro superior. O ULFI é uma medida prática e de fácil compreensão, que permite a avaliação de pacientes com disfunções musculoesqueléticas de todo o membro superior, sobretudo aqueles com mais de uma região acometida. **Conclusão:** O ULFI traduzido e adaptado transculturalmente para o português brasileiro apresenta confiabilidade, consistência interna e validade do construto adequadas, com estrutura fatorial unidimensional.

**Palavras-chave:** membro superior, medidas de resultado relatados pelo paciente, reprodutibilidade dos testes, estudo de validação, análise fatorial



## Abstract

**Objective:** The purpose of this study was to translate, cross-cultural adapt and validate the Upper Limb Functional Index (ULFI) into Brazilian Portuguese.

**Methods:** The translation and cross-cultural adaptation of the ULFI was performed according to international recommendations and the pre-final version of the ULFI was answered by 30 patients with upper limb musculoskeletal disorders to assess questionnaire understanding. The final version of ULFI was applied to 200 patients to verify structural validity. The other questionnaires of the study, i.e., Quick Disabilities of the Arm, Shoulder and Hand (QuickDash), 36-item Short Form Health Survey (SF-36) and Numerical Pain Scale (NPS) were answered by 180 patients to assess its construct validity using the Spearman correlation ( $\rho$ ). The internal structure of the ULFI was evaluated by exploratory and confirmatory factor analysis with fit indices Chi-square ( $X^2$ ), Chi-square/degrees of freedom ( $X^2/df$ ), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) and Tucker Lewis Index (TLI). A subsample of 51 patients answered ULFI after a mean of 9 days to assess test-retest reliability using intraclass correlation coefficient (ICC), standard error of measurement (SEM) and minimum detectable change (MDC). Internal consistency was assessed by Cronbach's alpha ( $\alpha$ ). Floor and ceiling effects were also assessed. **Results:** During translation and cross-cultural adaptation process, no item in the pre-final version of the questionnaire was misunderstood by more than 20% of patients, resulting in the final version of ULFI. Factor analysis identified a unidimensional structure with acceptable fit indices (RMSEA = 0.063, CFI = 0.918, TLI = 0.910). ULFI showed excellent reliability (CCI = 0.909) and adequate internal consistency ( $\alpha$  = 0.897). SEM was 6.11 and the MDC was 14.26. ULFI demonstrated high correlation with QuickDash ( $\rho$  = -0,721), moderate correlation with NPS ( $\rho$  = -0,529), and physical function ( $\rho$  = 0,529), role physical ( $\rho$  = 0,572), pain ( $\rho$  = 0,562), general health status ( $\rho$  = 0,505), social function ( $\rho$  = 0,557) domains of SF-36, and low correlation with vitality ( $\rho$  = 0,361), role emotional ( $\rho$  = 0,440) and mental health ( $\rho$  = 0,309) domains of SF-36. There were no floor and ceiling effects. **Discussion:** As hypothesized, ULFI demonstrated high reliability, adequate internal consistency and high correlation with QuickDash, another region-specific instrument that also assesses function of the entire upper limb. The ULFI is a practical and easy to understand measure, which can be used to assess function of the entire upper limb of patients with musculoskeletal disorders, especially those with more than one segment affected. **Conclusion:** The translated, cross-cultural adapted and validated Brazilian Portuguese version of ULFI has adequate reliability, internal consistency and construct validity with a unidimensional structure.

**Keywords:** upper limb, patient reported outcome measures, reproducibility of results, validation study, factor analysis.

## LISTA DE SIGLAS

AE - Aspectos Emocionais  
AF - Aspectos Físicos  
AFC - Análise Fatorial Confirmatória  
AFE - Análise Fatorial Exploratória  
AP - Análise Paralela  
AS - Aspectos Sociais  
CCI - Coeficiente de Correlação Intraclasse  
CF - Capacidade Funcional Dor  
CFI - *Comparative Fit Index*  
COF - Componente Físico  
COM - Componente Mental  
COSMIN - *COnsensus-based Standards for the selection of health Measurement INstruments*  
DASH - *Disabilities of the Arm, Shoulder and Hand*  
DMCI - Diferença Mínima Clinicamente Importante  
DMD - Diferença Mínima Detectável  
DP - Desvio Padrão  
EGS - Estado Geral de Saúde  
END - Escala Numérica da Dor  
EPM - Erro Padrão da Medida  
ML - *Maximum Likelihood*  
NULI - *Neck and Upper Limb Index*  
PROM - *Patient-Reported Outcome Measure*  
QuickDash - *Quick Disabilities of the Arm, Shoulder and Hand*  
RDWLS - *Robust Diagonally Weighted Least Squares*  
rho - Coeficiente de Correlação de Spearman  
RMSEA - *Root Means Square Error of Approximation*  
SM - Saúde Mental  
SF-36 - *36-Item Short Form Health Survey*  
TLI - *Tucker-Lewis Index*  
UEFI - *Upper Extremity Functional Index*

UEFS - *Upper Extremity Functional Scale*

ULFI – *Upper Limb Functional Index*

## SUMÁRIO

1. INTRODUÇÃO .....	12
2. OBJETIVO.....	15
2.1 OBJETIVO GERAL .....	15
2.2 OBJETIVOS ESPECÍFICOS.....	15
3. HIPÓTESE .....	16
HIPÓTESE AFIRMATIVA .....	16
4. REVISÃO DA LITERATURA.....	16
4.2 ADAPTAÇÃO TRANSCULTURAL.....	22
4.3 VALIDADE DE CONTEÚDO .....	24
4.4 ESTRUTURA INTERNA .....	24
4.4.1 Análise Fatorial .....	25
4.4.2 Consistência Interna .....	28
4.4.3 Validade Transcultural (invariância de medida).....	28
4.5 Demais propriedades de medida .....	29
4.5.1 Confiabilidade .....	29
4.5.2 Validade de critério .....	31
4.5.3 Validade de construto .....	31
4.5.4 Responsividade .....	32
5. MÉTODOS .....	33
5.1 DESENHO DO ESTUDO .....	33
5.2 LOCAL E PARTICIPANTES DA PESQUISA .....	36
5.3 TRADUÇÃO E ADAPTAÇÃO TRANSCULTURAL .....	36
5.4 APLICAÇÃO DOS QUESTIONÁRIOS .....	38
5.5 OUTROS QUESTIONÁRIOS .....	38
5.6 ANÁLISE ESTATÍSTICA.....	40
6. RESULTADOS.....	42
7. DISCUSSÃO .....	48
7.1 LIMITAÇÕES.....	53
8. CONCLUSÃO.....	54
9. REFERÊNCIAS.....	54

ANEXO 1 - PARECER CONSUBSTANCIADO DO CEP .....	60
ANEXO 2 – ULFI ORIGINAL .....	61
ANEXO 3 – QUICKDASH.....	62
ANEXO 4 – SF-36 .....	64
ANEXO 5 – ESCALA NUMÉRICA DA DOR .....	68
APÊNDICE 1 – VERSÃO FINAL DO ULFI-BR .....	69

## 1. INTRODUÇÃO

Ao longo da evolução, o membro superior do ser humano passou por mudanças nas suas funções de locomoção e preensão, para um papel predominantemente de preensão e manipulação (CHURCHILL et al., 2013). Esta função manipulativa do membro superior o coloca em risco de desenvolver sintomas e disfunções nas diversas situações da vida moderna, sobretudo aquelas com movimentos repetitivos, posturas desfavoráveis e altas demandas físicas e psicológicas (APOSTOLI et al., 2012; BODIN; HA; DESCATHA, 2012).

Disfunções no membro superior de origem musculoesquelética causam maior incapacidade do que aquelas que acometem outras partes do corpo. A incapacidade de se alimentar, se vestir e se limpar podem ser debilitantes e estão frequentemente associadas a depressão, ansiedade e pensamentos catastróficos (JAYAKUMAR et al., 2018). As disfunções dos membros superiores também são comuns na população adulta inserida no mercado de trabalho. No Brasil, estudos mostram que a prevalência de sintomas no membro superior entre trabalhadores de diferentes segmentos varia de 19 a 69% (BARBOSA; ASSUNÇÃO; ARAÚJO, 2012; BARRO et al., 2015; CARDOSO et al., 2011; FERNANDES; CARVALHO; ASSUNÇÃO, 2011; FERRAZ et al., 1995; HEMBECKER et al., 2017; SANTOS FILHO; BARRETO, 2001). Os distúrbios de membros superiores podem causar prejuízos financeiros para a sociedade decorrentes tanto de custos relacionados a tratamento quanto a afastamentos do trabalho (HAMILTON; CHESWORTH, 2013). O Instituto Nacional de Seguridade Social (INSS) registrou 12.502 casos de doenças ocupacionais em 2016, sendo que 51,26% correspondiam a acometimento dos membros superiores (BRASIL, 2016).

A avaliação das disfunções do membro superior envolve o exame clínico (anamnese, inspeção, palpação, testes diagnósticos específicos, avaliações da amplitude de movimento, força e sensibilidade) e exames de imagem que auxiliam no diagnóstico diferencial (LUKE et al., 2013). Entretanto, estes achados não conseguem mensurar o quanto aquela disfunção interfere na capacidade funcional e qualidade de vida do paciente. Para isso, as medidas de resultados relatados pelo paciente ou *patient-reported outcome measures* (PROMs) podem ser utilizadas para complementar a avaliação tradicional (TONGA et al., 2015).

As PROMs são instrumentos de avaliação centrados na percepção que o próprio paciente tem de seus sintomas e podem ser divididas em genéricas, articulação-específica, doença-específica ou regional-específica. As genéricas podem ser utilizadas em qualquer paciente, independentemente de sua doença e são particularmente úteis para comparação de resultados de grupos (EMERY; PERRIER; ACQUADRO, 2005; KYTE et al., 2015). Um exemplo de PROM genérica é o *36-item Short Form Health Survey* (SF-36) que avalia qualidade de vida (CICONELLI et al., 1999). As PROMs doença-específicas são direcionadas a populações com determinadas doenças ou distúrbios, tendem a ser mais responsivas a pequenas mudanças e são mais apropriadas para a avaliação de indivíduos isolados (KYTE et al., 2015). Exemplos de PROMs doença-específica e articulação-específica são o *Boston Carpal Tunnel Questionnaire* (LEVINE et al., 2011) para síndrome do túnel do carpo e o *Oswestry Disability Index* (FAIRBANK et al., 1980) para avaliação da coluna respectivamente.

Mais recentemente, foram desenvolvidas PROMs regionais-específicas que permitem a avaliação do membro superior, por exemplo, como uma cadeia cinética única. As PROMs regionais-específicas têm sido bastante utilizadas em uma

variedade de condições musculoesqueléticas, em virtude de sua praticidade, facilidade de administração e versatilidade por permitir a comparação entre diferentes patologias que afetam o membro superior (BADALAMENTE et al., 2013a; BEATON et al., 2005; TONGA et al., 2015).

Existem várias PROMs regionais-específicas descritas na literatura. Destas, apenas o *Disabilities of the Arm, Shoulder and Hand* (DASH) e sua versão reduzida (QuickDash) foram validados para o português brasileiro (ORFALE et al., 2005; PUGA et al., 2013).

O DASH é um questionário regional-específico amplamente utilizado na literatura. Foi traduzido para diversas línguas e suas propriedades psicométricas já foram extensamente investigadas. Porém, a qualidade dos processos de tradução e validação do DASH em países em desenvolvimento foi recentemente questionada em uma revisão sistemática (DE KLERK; BUCHANAN; JEROSCH-HEROLD, 2018). No estudo de tradução e validação do DASH para o português, não foi descrito o teste de hipótese para a validação do construto, nem realizada análise da estrutura fatorial (ORFALE et al., 2005). Em outro estudo, foi realizada análise fatorial exploratória por análise de componentes principais, a qual não é a mais indicada para a determinação do número de fatores (CHENG et al., 2008)

O *Upper Limb Functional Index* (Índice Funcional do Membro Superior [ULFI]) é um questionário regional-específico para o membro superior com 25 itens em uma página que apresenta boas propriedades psicométricas e alta correlação com o DASH, mas com maior praticidade pelo tempo de preenchimento e pontuação reduzidos (GABEL et al., 2006).

O ULFI já foi traduzido e validado para o turco (TONGA et al., 2015), espanhol (CUESTA-VARGAS; GABEL, 2013), francês-canadense (HAMASAKI et al., 2014),



italiano (SARTORIO et al., 2015) e coreano (IN et al., 2017). Estas versões demonstraram propriedades psicométricas equivalentes ao original em inglês. Entretanto, o ULFI não possui nenhuma versão em português brasileiro o que impede seu uso nesta população e a comparação de resultados com estudos de outros países.

## **2. OBJETIVO**

### **2.1 OBJETIVO GERAL**

O objetivo deste estudo foi realizar a tradução, adaptação transcultural e a validação do ULFI para o português brasileiro em indivíduos com disfunções musculoesqueléticas no membro superior.

### **2.2 OBJETIVOS ESPECÍFICOS**

- Realizar o teste da versão pré-final do instrumento ULFI;
- Verificar a confiabilidade teste-reteste e a concordância da versão traduzida do ULFI;
- Verificar a consistência interna da versão traduzida do ULFI.
- Verificar a validade de construto do ULFI por meio da avaliação de correlação do ULFI com outros instrumentos;
- Avaliar os efeitos *floor* e *ceiling*.

### **3. HIPÓTESE**

#### **HIPÓTESE AFIRMATIVA**

A versão brasileira do ULFI apresenta valores adequados de consistência interna e confiabilidade e correlação de moderada a alta com o QuickDash, SF-36 e END.

### **4. REVISÃO DA LITERATURA**

#### **4.1 MEDIDAS DE RESULTADO REPORTADAS PELO PACIENTE**

As PROMs são quaisquer relatórios da condição de saúde de um paciente que provém diretamente dele, sem interpretação da resposta por parte do clínico ou qualquer outra pessoa. Elas têm sido cada vez mais usadas na prática clínica e já são bem consolidadas na pesquisa. Estes instrumentos podem ajudar os fisioterapeutas e demais profissionais da saúde a monitorar de forma mais efetiva o impacto de seu tratamento e desenvolver melhores estratégias de intervenção (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; KYTE et al., 2015; METCALF et al., 2007).

Muitas vezes, os parâmetros objetivos de avaliação (amplitude de movimento, sensibilidade, força, etc.) ou mesmo a percepção subjetiva das limitações funcionais do paciente pelo profissional da saúde não refletem o grau de incapacidade percebido pelo próprio paciente (BRAUN et al., 2017; METCALF et al., 2007).

As PROMs medem aspectos específicos da saúde percebidos pelo paciente que não podem ser diretamente mensurados (por exemplo, a dor) ou que são difíceis de serem observados (por exemplo, o desempenho nas atividades de vida diária) (DAVIDSON; KEATING, 2014; METCALF et al., 2007). Os questionários são normalmente desenvolvidos para focar em um ou mais elementos do bem-estar do paciente. Algumas PROMs medem uma combinação de aspectos físico, mental e social que juntos formam a chamada qualidade de vida relacionada à saúde, enquanto outras avaliam apenas um aspecto como atividade física, por exemplo (BADALAMENTE et al., 2013a; KYTE et al., 2015).

Para se usar uma PROM, é necessário que os respondentes tenham cognição preservada e capacidade de ler. Por outro lado, algumas PROMs pode ser aplicadas em forma de entrevista. Nestes casos, deve-se levar em consideração que o modo de preenchimento pode gerar resultados diferentes, por exemplo, quando auto-preenchidos ou aplicados por meio de entrevista pessoalmente ou telefone (DAVIDSON; KEATING, 2014; HAYS et al., 2009; VARNI; LIMBERS; NEWMAN, 2009).

Questões específicas de uma PROM (conhecidas como “itens”) normalmente são agrupadas para formar uma subcategoria ou domínio. Por exemplo, várias questões sobre deambulação podem ser agrupadas no domínio “mobilidade”. Uma PROM pode conter vários domínios formados por diferentes itens (KYTE et al., 2015). A escala genérica de qualidade de vida SF-36, por exemplo, possui 36 itens que avaliam 8 domínios (capacidade funcional, limitação por aspectos físicos, dor, estado geral de saúde, vitalidade, aspectos sociais, limitação por aspectos emocionais e saúde mental) (CICONELLI et al., 1999).

As PROMs podem ser categorizadas em genéricas, articulação-específica, doença-específica e regional-específica. As escalas genéricas medem a saúde em geral e o bem-estar em qualquer tipo de pessoa, independentemente de sua doença (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; KYTE et al., 2015). Muitas destas escalas como a SF-36 e a EQ-5D têm sido usadas extensivamente em diversas áreas clínicas e sua confiabilidade e validade já estão bem estabelecidas (CICONELLI et al., 1999; RABIN; CHARRO, 2001). São menos sensíveis tendo, assim, menor capacidade para medir mudanças no estado clínico de um único paciente ao longo do tempo. Assim, elas são mais úteis para análise e comparação de grupos ou como uma ferramenta de medida de desfecho secundária. Normalmente, escalas genéricas não são utilizadas como medida de desfecho primária nas condições que afetam o membro superior, pois estas tendem a impactar pouco a saúde em geral, exceto por algumas condições que também apresentam repercussões sistêmicas, como artrite reumatoide. Porém, em combinação com escalas doença-específicas, elas podem fornecer um panorama mais completo da saúde do indivíduo (BADALAMENTE et al., 2013a; KYTE et al., 2015).

As escalas articulação-específica e doença-específica tendem a ser mais responsivas a mudanças sutis na condição do paciente e, portanto, são mais indicadas para avaliação individual. Seus itens focam em aspectos importantes daquela condição em particular e por isso conseguem medir melhor o estado de saúde da população alvo. Por outro lado, elas não necessariamente funcionarão bem para avaliar condições diferentes daquelas para as quais foi desenvolvida (BADALAMENTE et al., 2013a; KYTE et al., 2015).

As escalas regionais-específicas, como o ULFI e o DASH avaliam o membro superior como uma cadeia cinética única. Assim, o resultado da escala reflete a influência que uma parte do membro tem sobre toda a extremidade. Este tipo de escala também possibilita a comparação de diferentes condições que afetam o membro superior como um todo ou parte dele em um mesmo instrumento. Com isso, elimina-se a necessidade de vários instrumentos para avaliar um paciente que tenha mais de uma articulação afetada. Esta versatilidade torna este tipo de questionário mais prático pelo seus tempos de preenchimento e de cálculo do escore mais rápidos, facilitando seu uso no ambiente clínico onde muitas vezes o tempo disponível é reduzido (BADALAMENTE et al., 2013b; DAVIS et al., 1999; KYTE et al., 2015).

Escalas regionais-específicas podem ser tão responsivas quanto escalas doença-específica. Beaton et al (2001) verificaram que a responsividade do DASH era equivalente ou maior do que instrumentos específicos da articulação do ombro (*Shoulder Pain and Disability Index*) e para síndrome do túnel do carpo (Questionário *Brigham*). Além disso, também verificaram que itens do questionário SPADI eram sensíveis a disfunções no punho e itens do Questionário *Brigham* eram sensíveis a disfunções no ombro.

Estes achados corroboram os resultados de outro trabalho que identificou que itens existentes em questionários específicos de membro superior também são relevantes para outras regiões e condições. Neste estudo, quinze clínicos avaliaram uma amostra de 132 itens de questionários específicos de articulação do membro superior e consideraram que 72% dos itens também poderiam ser influenciados por disfunções em outras partes do membro superior e não exclusivamente de uma só articulação. Por exemplo, “usar o telefone” pode ser um item relevante para

pacientes que possuem problemas tanto no ombro, como no cotovelo ou na mão. (DAVIS et al., 1999).

A efetividade de um tratamento deve ser avaliada por meio de escalas que contenham itens que sejam relevantes para os pacientes. Como o alívio dos sintomas e a redução da incapacidade são os principais motivos que os pacientes com distúrbios musculoesqueléticos do membro superior procuram ajuda, clínicos e pesquisadores normalmente utilizam instrumentos que refletem estes aspectos de saúde (KYTE et al., 2015). A escolha de qual escala utilizar depende da população alvo, do que se pretende medir, de suas características práticas (tempo de preenchimento e facilidade de uso) e de propriedades psicométricas adequadas (BADALAMENTE et al., 2013a).

Oito PROMs regionais-específicas de membro superior para uso na população em geral foram encontradas na literatura: o *Neck and Upper Limb Index* (NULI) (STOCK; STREINER; REARDON, 1995), o *Upper Extremity Functional Index* (UEFI) (STRATFORD; BINKLEY; STRATFORD, 2001) *Upper Extremity Functional Scale* (UEFS) (PRANSKY et al., 1997), o WORQ-UP (AERTS et al., 2017), o DASH (HUDAK; AMADIO; BOMBARDIER, 1996), QuickDASH com 11 itens (BEATON et al., 2005), o QuickDASH com 9 itens (GABEL et al., 2009) e o ULFI (GABEL et al., 2006).

A UEFS foi desenvolvida e validada em uma população apenas de trabalhadores e sua confiabilidade não foi testada (MICHENER; LEGGIN, 2001; PRANSKY et al., 1997). O NULI também foi desenvolvido para avaliação apenas de trabalhadores e não há evidência de sua validade de critério, validade de conteúdo, erro de medida e responsividade. A análise fatorial identificou 4 dimensões de seus 20 itens (ALRENI et al., 2017; SALERNO et al., 2002; STOCK; STREINER;

REARDON, 1995). Quanto ao UEFI, sua versão original com 20 itens se revelou multidimensional por análise Rasch, sendo proposta uma versão de 15 itens que teria propriedades de medida mais adequadas (HAMILTON; CHESWORTH, 2013; STRATFORD; BINKLEY; STRATFORD, 2001). Porém, recomendou-se mais estudos para esclarecer os melhores itens da escala e sua performance em contextos culturais e clínicos diferentes (ARUMUGAM; MACDERMID, 2018). O WORQ-UP avalia limitações especificamente relacionadas ao trabalho, sua estrutura possui 4 fatores, mas sua confiabilidade e validade não foram testadas (AERTS et al., 2018).

O DASH é o questionário regional-específico de membro superior mais utilizado na literatura. Porém, estudos recentes apontaram performance reduzida em análise fatorial e teoria de resposta ao item (WORMALD et al., 2019), além de multidimensionalidade o que inviabiliza o uso de um escore único da escala (CHENG et al., 2008; FRANCHIGNONI et al., 2010). Outros pontos a serem considerados do DASH são seu tamanho, complexidade, erros de cálculo do escore e consistência interna excessiva indicando redundância de itens. Na tentativa de melhorar sua praticidade, o DASH foi modificado para uma versão mais curta de 11 itens, o QuickDASH, que demonstrou propriedades psicométricas equivalentes ao DASH original (BEATON et al., 2005). Entretanto, sua validade foi questionada em virtude de resultados conflitantes sobre sua estrutura fatorial. Uma outra versão de 9 itens, o QuickDASH-9, foi testada e demonstrou uma estrutura unidimensional válida (GABEL et al., 2009). O QuickDASH-9 não foi validado para o português brasileiro.

O *Upper Limb Function Index* (ULFI) foi desenvolvido na Austrália em 2006 para suprir as deficiências dos poucos instrumentos de avaliação regional do membro superior existentes na época. A proposta foi de desenvolver um instrumento abrangente (que pudesse ser utilizado em diferentes populações e faixas etárias),

prático (número de itens reduzidos para diminuir redundâncias, erros e tempos de preenchimento e de pontuação) e com boas propriedades psicométricas. O questionário consiste em uma lista com 25 sentenças que descrevem sintomas e dificuldades normalmente vivenciadas por pessoas com problemas nos membros superiores. O questionário monitora sintomas e função nas regiões proximal, central e distal do membro superior permitindo seu uso em uma variedade de distúrbios musculoesqueléticos (GABEL et al., 2006).

Posteriormente, o autor testou uma versão modificada do ULFI com opção de 3 respostas: sim (1 ponto), parcialmente (0,5 ponto) e não (0). Esta versão demonstrou confiabilidade, consistência interna e responsividade ainda maiores que sua versão original e diminuiu erros relacionadas à mensuração da funcionalidade do membro superior, sem perder sua praticidade (GABEL et al., 2010). Foi esta versão modificada de 3 respostas a utilizada no presente estudo. Somam-se os pontos assinalados e multiplica-se por 4. A partir de então, subtrai-se 100 desse valor encontrado, obtendo-se um escore final que varia de 0 a 100, de forma que quanto maior o escore, maior a função.

Charles Gabel, um dos autores do ULFI, desenvolveu em 2012 o *Lower Limb Functional Index* e em 2013 o *Spine Functional Index*. São questionários regionais do membro inferior e coluna respectivamente e têm um formato parecido com o ULFI. Apresentam-se em uma única página e os itens são dispostos em caixas de texto sombreadas alternadamente para facilitar a leitura e torná-la mais agradável aos respondentes (GABEL et al., 2012, 2013).

## **4.2 ADAPTAÇÃO TRANSCULTURAL**



Normalmente, os questionários autoaplicáveis são desenvolvidos em países de língua inglesa e o seu uso em populações de outros países requer não só tradução, mas também uma adaptação transcultural para manter a validade de conteúdo do instrumento. O cuidado a esses detalhes assegura que o impacto da doença ou de seu tratamento seja descrito de maneiras similares em ensaios clínicos em diferentes locais ou em diferentes avaliações (BEATON et al., 2000).

A necessidade de adaptações transculturais de questionários tem aumentado com o número crescentes de estudos multicêntricos. A simples tradução de um questionário pode não ser suficiente em função das diferenças linguísticas e culturais. A percepção de qualidade de vida e a maneira como os problemas de saúde se manifestam variam de cultura para cultura (GUILLEMIN; BOMBARDIER; BEATON, 1993).

A adaptação transcultural tem como objetivo produzir equivalência de conteúdo entre o instrumento original e o novo. Supõe-se que esse processo garantirá a retenção das propriedades psicométricas como validade e confiabilidade da escala e/ou dos itens (BEATON et al., 2000; GUILLEMIN; BOMBARDIER; BEATON, 1993).

Beaton et al. (2000) descreveram orientações para a realização do processo de tradução e adaptação transcultural de questionários. Após esta etapa, é importante a análise das propriedades de medida do instrumento como validade de conteúdo, estrutura interna (validade estrutural, consistência interna, validade transcultural) e outras propriedades como confiabilidade, erro de medida, validade de critério, teste de hipótese para validade de construto e responsividade (PRINSEN et al., 2018).

### **4.3 VALIDADE DE CONTEÚDO**

Validade de conteúdo se refere ao grau em que o conteúdo de uma PROM é um reflexo adequado do construto a ser medido. É a propriedade de medida mais importante e também a mais difícil de ser avaliada. Existem três aspectos que são destacados: (1) relevância (todos os itens da PROM devem ser relevantes para o construto de interesse em uma determinada população e contexto de uso), (2) abrangência (nenhum aspecto chave do construto podem estar ausente) e (3) compreensibilidade (os itens devem ser compreendidos pelos pacientes/usuários da escala) (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; MOKKINK et al., 2019; SOUZA; ALEXANDRE; GUIRARDELLO, 2017; TERWEE et al., 2018).

A avaliação da validade de conteúdo deve ser feita perguntando-se aos pacientes e profissionais (pesquisadores, clínicos) sobre a relevância, abrangência e compreensibilidade dos itens (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; MOKKINK et al., 2019; TERWEE et al., 2018).

A falta de validade de conteúdo pode afetar todas as outras propriedades de medida. Itens irrelevantes podem diminuir a consistência interna, a validade estrutural e a interpretabilidade de uma PROM. A ausência de conceitos pode diminuir a validade e a responsividade (PRINSEN et al., 2018; TERWEE et al., 2018).

### **4.4 ESTRUTURA INTERNA**

A estrutura interna de um instrumento se refere a como os diferentes itens de uma PROM se relacionam e estabelecer se podem ser combinados em uma única

escala ou subescalas. É composta da validade estrutural (medida por análise fatorial ou teoria de resposta ao item ou análise Rasch), a consistência interna (medida pelo alfa de Cronbach) e a validade transcultural. Estas três propriedades de medida focam na qualidade dos itens e a relação entre eles, ao contrário das demais propriedades de medida que se concentram na qualidade da escala ou subescalas (MOKKINK et al., 2018a, 2019; PRINSEN et al., 2018).

A estrutura interna é relevante para instrumentos baseados em um modelo reflexivo, no qual os itens são um reflexo do construto a ser medido (também conhecido como traço latente, fator ou dimensão). Isto é, o traço latente (algo que não pode ser medido diretamente) é estimado por outras variáveis observáveis. Como a unidimensionalidade (um fator) ou validade estrutural é pré-requisito para a interpretação da consistência interna, é recomendado primeiro avaliar a validade estrutural para depois avaliar a consistência interna (MOKKINK et al., 2018b, 2018a; PRINSEN et al., 2018).

Um exemplo de modelo reflexivo é a medida de ansiedade, que se manifesta por meio de sinais específicos como pensamentos preocupantes, pânico, inquietação. Perguntando-se os pacientes sobre estes aspectos, podemos avaliar o grau de ansiedade (todos os itens são reflexo do mesmo construto) (PRINSEN et al., 2018).

Se os itens da escala ou subescala não estão correlacionados (modelo formativo), estas análises não são relevantes e a análise fatorial ou análise Rasch podem ser ignoradas pois não são interpretáveis (PRINSEN et al., 2018).

#### **4.4.1 Análise Fatorial**

A análise fatorial é uma das técnicas estatísticas que podem ser usadas para se avaliar a estrutura interna de um questionário. A Análise Fatorial pode ser do tipo exploratória (AFE) ou confirmatória (AFC) e ambas têm como objetivo agrupar uma grande quantidade de itens (variáveis observáveis) em um número reduzido de fatores (também conhecidos como construtos, dimensões ou traços latentes) que não podem ser medidos diretamente. Elas se diferem fundamentalmente pelo número e natureza das especificações e restrições feitas *a priori* no modelo fatorial (BROWN, 2006; DAMASIO, 2012; SCHREIBER et al., 2006).

Quando não há informações prévias sobre a estrutura do questionário, utiliza-se a AFE para determinar o menor número de fatores que explicam as variáveis observáveis (por exemplo, por meio do valor das cargas fatoriais de cada variável observável). É uma análise “descritiva” porque nenhuma restrição *a priori* é imposta quanto ao padrão de relações entre as variáveis observadas e latentes (BROWN, 2006; DAMASIO, 2012; SOUZA; ALEXANDRE; GUIRARDELLO, 2017).

Há evidências de que o método mais preciso para se realizar a AFE é a Análise Paralela (AP) desenvolvido por Horn em 1965, embora seja um dos métodos mais subutilizados. A AP supera as limitações de outros métodos que tendem a superestimar o número de fatores como o K1 (Kaiser, 1960), que retém fatores com *eigenvalues*  $> 1$ , e o *scree plot* proposto por Cattell (1966), que retém fatores acima do ponto de inflexão do gráfico (DAMASIO, 2012; HAYTON; ALLEN; SCARPELLO, 2004).

A AP simula matrizes de correlação de variáveis aleatórias baseadas no tamanho da amostra e no número de variáveis do banco de dados real. Os *eigenvalues* das matrizes de correlação aleatórias são comparados aos *eigenvalues* correspondentes da matriz de correlação da matriz do banco de dados real. Fatores

correspondentes dos *eigenvalues* reais maiores do que a média dos *eigenvalues* paralelos aleatórios devem ser retidos. *Eigenvalues* reais menores do que seus correspondentes *eigenvalues* paralelos aleatórios são considerados erros de amostragem e, portanto, não retidos (HAYTON; ALLEN; SCARPELLO, 2004).

A AFC é um tipo de modelagem por equações estruturais que lida especificamente com modelos de medição, isto é, relações entre medidas observáveis ou indicadores e variáveis latentes ou fatores. Ao contrário da AFE, na AFC o pesquisador precisa pré-estabelecer todos os aspectos do modelo como o número de fatores, quais variáveis indicadoras são relacionadas com quais fatores e assim por diante (BROWN, 2006). A AFC é uma técnica mais restritiva e rigorosa sendo mais indicada na validação de questionários (SCHREIBER et al., 2006; SOUZA; ALEXANDRE; GUIRARDELLO, 2017).

Na modelagem por equações estruturais, a avaliação de ajuste do modelo não é tão direta como em abordagens estatísticas baseadas em variáveis medidas sem erro. Como não há um teste único de significância estatística que identifique o modelo correto para aquela amostra, é necessário considerar múltiplos critérios simultaneamente chamados de índices de ajuste (SCHREIBER et al., 2006).

O teste do qui-quadrado ( $X^2$ ) avalia a hipótese nula que as matrizes de covariância da população e do modelo proposto são iguais. Ou seja, espera-se que o p valor seja maior que 0,05. Porém, o  $X^2$  é muito sensível ao tamanho da amostra e pequenas diferenças entre os dados observados e o modelo tendem a ser estatisticamente significantes, podendo levar a rejeição de modelos plausíveis. No intuito de compensar a imprecisão do  $X^2$ , propõe-se a comparação do  $X^2$  com o valor esperado da distribuição da amostra (número de graus de liberdade [gl]). Para um bom ajuste, a razão  $X^2/gl$  deve ser a menor possível. Razões menores que 2 e 3 são

consideradas boas e aceitáveis, respectivamente (SCHERMELLEH-ENGEL; MOOSBRUGGER; MÜLLER, 2003).

O *Root Means Square Error of Approximation* (RMSEA) é um dos índices de ajuste absolutos mais utilizados e avalia em que medida o modelo se ajusta razoavelmente na população. Quanto mais o valor do índice se aproxima de 0 melhor é o ajuste do modelo. Diversos pontos de corte são descritos na literatura. De forma geral, valores menores que 0,05 são considerados bons, entre 0,05 e 0,08 são adequados, entre 0,08 e 0,10 medíocres e acima de 0,10 inaceitáveis (SCHERMELLEH-ENGEL; MOOSBRUGGER; MÜLLER, 2003).

O *Comparative Fit Index* (CFI) e o *Tucker-Lewis Index* (TLI) são índices comparativos ou incrementais que avaliam quanto o modelo se distancia de um modelo de base nulo ou de independência cujos ajustes são ruins. Portanto, quanto mais esses índices se aproximarem de 1 melhor. Recomenda-se valores próximos de 0,95 para bom ajuste (BROWN, 2006; HU; BENTLER, 1999).

#### **4.4.2 Consistência Interna**

Consistência interna se refere ao grau de interrelação entre os itens e normalmente é avaliada pelo alfa de Cronbach. A unidimensionalidade é um pré-requisito para sua interpretação. Portanto, se o resultado da análise fatorial demonstrar que a escala possui mais de um fator, deve-se avaliar a consistência interna de cada uma das subescalas (PRINSEN et al., 2018).

#### **4.4.3 Validade Transcultural (invariância de medida)**

A validade transcultural ou invariância de medida se refere ao grau em que a performance dos itens de uma PROM adaptada culturalmente é um reflexo adequado da performance dos itens da escala original. Para ser avaliada, são necessários pelo menos dois grupos que se difiram em termos de etnias, língua, sexo ou idade. A avaliação é feita verificando-se a ocorrência do *differential item functioning* por meio de análises de regressão logística ou se a estrutura fatorial e cargas fatoriais são equivalentes entre os grupos utilizando-se a análise fatorial confirmatória multigrupo (DAMASIO, 2012; PRINSEN et al., 2018).

#### **4.5 Demais propriedades de medida**

Diferentemente da validade de conteúdo e estrutura interna, as demais propriedades de medida (confiabilidade, erro de medida, validade de critério, teste de hipótese para validade de construto e responsividade) avaliam a qualidade da escala ou subescala como um todo ao invés dos itens.

##### **4.5.1 Confiabilidade**

Confiabilidade se refere a proporção de variância total da medida que é devida a diferenças “verdadeiras” entre os pacientes e não ao erro de observação. É um escore médio que seria obtido se a escala fosse administrada infinitas vezes na mesma pessoa. Portanto, ela está associada a consistência do escore e não a sua acurácia (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; MOKKINK et al., 2019; PRINSEN et al., 2018).

Os testes de confiabilidade podem ser: (1) interobservadores, usado para avaliar o grau de consistência entre diferentes avaliadores nos mesmos pacientes em um determinado momento ou (2) intraobservador (ou teste-reteste), em que o mesmo avaliador avalia os mesmos pacientes em ocasiões diferentes (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014).

Um requisito importante para se medir a confiabilidade intraobservador e o erro de medida do instrumento é que os pacientes devem ser estáveis em relação ao construto a ser medido entre as avaliações. Isto é, idealmente não devem sofrer nenhuma intervenção/tratamento ou mudanças em seu quadro clínico de uma medida para a outra. Para isso, o intervalo de tempo entre as medições deve ser longo o bastante para prevenir que o paciente lembre do preenchimento anterior da escala e curto o suficiente para não haver mudança clínica ou do construto de interesse (DAVIDSON; KEATING, 2014; PRINSEN et al., 2018). Não há um consenso sobre qual o intervalo de tempo adequado entre as avaliações. Marx et al. (2003) verificaram não haver diferenças da confiabilidade teste-reteste entre intervalos de 2 dias e 2 semanas em instrumentos de medida em saúde (MARX et al., 2003).

As condições de avaliação também devem ser similares. Isso incluiu o tipo de administração (autoaplicado, entrevista), o local (hospital, em casa) e as instruções dadas para preenchimento (HAYS et al., 2009; PRINSEN et al., 2018).

O método utilizado para se avaliar a confiabilidade depende do tipo de escore do instrumento. Para escores contínuos, utiliza-se o coeficiente de correlação intraclassa (CCI) que varia de 0 a 1, em que 0 indica nenhuma confiabilidade e 1 confiabilidade perfeita. Sugere-se valores de CCI maiores que 0,75 para instrumentos de medida em saúde (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014).



Para escores dicotômicos, nominais ou ordinais utiliza-se o kappa (MOKKINK et al., 2019; PRINSEN et al., 2018).

Na prática clínica, mais do que o CCI, é importante conhecer a diferença mínima detectável (DMD) que é a variação no escore do instrumento que é devida ao erro de medida. Portanto, uma variação no escore só pode ser considerado real se ela for maior que a DMD (VAN KAMPEN et al., 2013).

#### **4.5.2 Validade de critério**

A validade de critério avalia a correlação do novo instrumento com uma medida padrão ouro. Quando a correlação é investigada ao mesmo tempo é descrita como “validade concorrente”. Isso normalmente ocorre quando um instrumento tem potencial de ser substituído por uma versão mais curta de um questionário já validado (considerado o padrão-ouro) ou uma medida mais barata ou menos invasiva. Neste caso, espera-se uma alta correlação ( $\geq 0,8$ ). Quando um novo instrumento é comparado a um critério medido depois, tem-se a “validade preditiva”. Este tipo de validação é usada com medidas que preveem eventos futuros como resposta ao tratamento ou mortalidade (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; MOKKINK et al., 2019; TERWEE et al., 2007).

#### **4.5.3 Validade de construto**

A princípio, não há medida padrão ouro para PROMs já que estas medem construtos que não podem ser medidos com equipamentos. Apenas a versão

original de uma versão curta de uma PROM pode ser considerada um padrão ouro (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; PRINSEN et al., 2018).

Muitos testes de hipóteses podem ser feitos para se avaliar a validade de construto de uma PROM. Em geral, compara-se a nova PROM a outros instrumentos que medem construtos similares ou avalia-se diferenças de escores entre grupos conhecidos (por exemplo, indivíduos com disfunção de ombro devem apresentar escores diferentes de indivíduos saudáveis) (MOKKINK et al., 2019; PRINSEN et al., 2018).

Sugere-se a formulação de testes de hipóteses sobre a direção esperada (positivo ou negativo) e magnitude (absoluta ou relativa) das correlações. As hipóteses podem ser baseadas na literatura ou na experiência dos autores (MOKKINK et al., 2019; PRINSEN et al., 2018; TERWEE et al., 2007).

A validade de construto é calculada por meio de algum coeficiente de correlação que mede o grau de associação entre dois escores. Existem vários coeficientes de correlação, mas os mais utilizados em estudos de validação de construto são os coeficientes de correlação de Pearson e de Spearman. O coeficiente de Pearson é utilizado para variáveis contínuas e com distribuição normal. O coeficiente de Spearman é utilizado quando a escala é ordinal ou quando os dados não têm distribuição normal (DAVIDSON; KEATING, 2014).

#### **4.5.4 Responsividade**

Responsividade se refere a habilidade da PROM em detectar mudanças ao longo do tempo no construto medido. Para isso, aplica-se o instrumento em um grupo de pacientes cuja condição clínica se alterou (por exemplo, após o tratamento)

e compara-se o escore da escala a outro parâmetro que identifique a melhora ou piora do paciente. Na ausência de um padrão ouro do construto medido, pesquisadores normalmente utilizam uma escala de percepção global de mudança para comparação. A escala de percepção global de mudança é preenchida pelos pacientes ao mesmo tempo em que a escala a ser validada é reaplicada (ALRUBAIY; HUTCHINGS; WILLIAMS, 2014; DAVIDSON; KEATING, 2014; MOKKINK et al., 2019).

Existem diversas formas de se calcular a responsividade que pode ser subdividida em responsividade interna e responsividade externa. Não há um consenso sobre o que responsividade realmente significa o que se reflete na abundância de métodos estatísticos sobre o assunto. O *COnsensus-based Standards for the selection of health Measurement INstruments* (COSMIN) sugere o uso de teste de correlação ou o cálculo da área sob a curva ROC para variáveis contínuas e cálculo da sensibilidade e especificidade para escores dicotômicos (MOKKINK et al., 2019; PRINSEN et al., 2018).

Os estudos de responsividade permitem a definição da diferença mínima clinicamente importante (DMCI) que é a menor variação no escore de uma PROM que é percebida pelo paciente e que representa uma melhora significativa no seu estado clínico (DAVIDSON, 2014).

## **5. MÉTODOS**

### **5.1 DESENHO DO ESTUDO**

Trata-se de um estudo de validação de questionário realizado conforme as *Guidelines for the Process of Cross-cultural Adaptation of Self-Report Measures* (BEATON et al., 2000) e *Consensus-based Standards for the Selection of Health Measurement Instruments* (COSMIN) (MOKKINK et al., 2019). A autorização para realização da tradução do ULFI para o português brasileiro foi concedida via e-mail por um dos autores do questionário (Dr. Charles Philip Gabel). O estudo foi realizado em três fases: (1) tradução e adaptação transcultural do questionário, (2) teste da versão pré-final traduzida do ULFI para o português brasileiro, e (3) validação da versão final do ULFI transculturalmente adaptada para o português brasileiro (Figura 1).

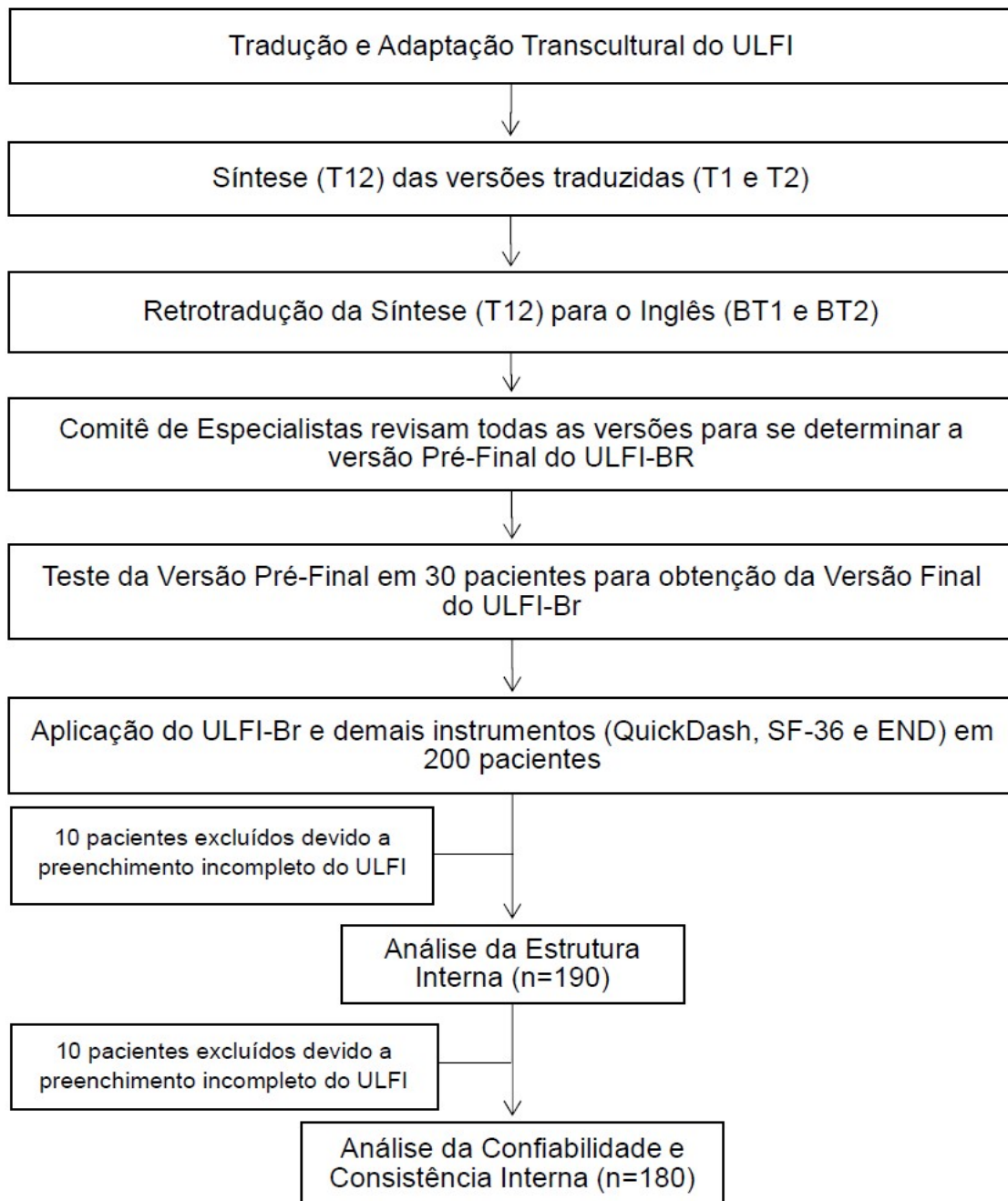


Figura 1: Desenho Experimental do estudo. Fonte: autoria própria.

## **5.2 LOCAL E PARTICIPANTES DA PESQUISA**

A pesquisa foi realizada no ginásio de fisioterapia adulto do Hospital Sarah (São Luís, MA, Brasil). O projeto foi aprovado pelo Comitê de Ética em Pesquisa da referida instituição, sob o CAAE 94496318.4.0000.0022 em 30/10/2018 (ANEXO 1).

Foram incluídos no estudo pacientes admitidos com queixa de dor e/ou alteração de função no(s) membro(s) superior(es) de origem musculoesquelética com duração maior ou igual a 12 semanas, com diagnóstico ortopédico definido. Além desses critérios, os pacientes deveriam ser capazes de ler e escrever em português brasileiro, não possuir alterações cognitivas e ter mais de 18 anos. Não foram incluídos pacientes com histórico de cirurgia há menos de 6 meses, doenças infecciosas, neurológicas (sistema nervoso central), câncer e alterações psiquiátricas graves que impossibilitassem o preenchimento dos questionários.

## **5.3 TRADUÇÃO E ADAPTAÇÃO TRANSCULTURAL**

O processo de tradução e adaptação transcultural do ULFI para o português brasileiro seguiu os critérios de Beaton et al. (2000) e foi realizado em 5 estágios, conforme descrito abaixo.

1) Tradução: dois tradutores independentes (um fisioterapeuta com experiência há 7 anos na área e um professor de inglês com experiência em traduções há 20 anos sem conhecimentos técnicos de assuntos na área da saúde), ambos com português brasileiro como a língua mãe e com fluência na língua inglesa, realizaram a tradução da versão original do ULFI para o português brasileiro produzindo duas versões (T1 e T2).

2) Síntese das traduções: após discussões e revisões, os dois tradutores, juntamente com os pesquisadores, sintetizaram as duas versões do questionário traduzidas de forma independente (T1 e T2) e produziram uma única versão do ULFI de forma consensual (T12).

3) Retrotradução: dois tradutores independentes (sem conhecimentos técnicos de assuntos na área da saúde e sem conhecimento da versão original do questionário), ambos com o inglês como a língua mãe e com fluência na língua portuguesa do Brasil, realizaram a retrotradução da versão em português do ULFI de volta para o inglês produzindo duas versões (BT1 e BT2).

4) Análise de um comitê de especialistas: 2 especialistas da área de reabilitação, em conjunto com os 4 tradutores envolvidos no projeto, revisaram todas as versões traduzidas e retrotraduzidas, para correções de possíveis discrepâncias, alcançando-se, assim, uma versão pré-final do ULFI em português (ULFI-Br) de forma acordada entre todos os membros do comitê.

5) Teste da versão pré-final: a versão pré-final do ULFI-Br (APÊNDICE 1) foi aplicada em 30 indivíduos com disfunções nos membros superiores e com o português brasileiro como língua mãe. Os participantes leram e preencheram o questionário e, ao final do preenchimento, foram interrogados se houve alguma questão que gerou dúvidas de compreensão ou interpretação. As questões que não compreendidas por mais de 20% dos participantes, seriam reformuladas e testadas novamente em uma nova amostra de 30 participantes (RODRIGUES; FONSECA; MACDERMID, 2015), até que se alcançasse o nível de compreensão desejado, estabelecendo-se, assim, a versão final do ULFI no português brasileiro.

## **5.4 APLICAÇÃO DOS QUESTIONÁRIOS**

Após a definição da versão final do ULFI-Br, este foi aplicado juntamente com os demais instrumentos (QuickDash, SF-36 e a END) por um dos fisioterapeutas da equipe de reabilitação em pacientes consecutivos na consulta de admissão da fisioterapia entre dezembro de 2018 e abril de 2020. A amostra foi calculada conforme a recomendação do COSMIN de 7 pacientes para cada item do questionário (7 x 25 itens = 175 pacientes) (MOKKINK et al., 2019).

Cada paciente foi instruído sobre os objetivos da pesquisa e, após concordar em participar, assinou o termo de consentimento livre e esclarecido e preencheu os questionários sozinho.

Para a análise da confiabilidade teste-reteste, o ULFI-Br (ANEXO 2) foi preenchido em dois momentos por uma amostra de conveniência dos 51 primeiros pacientes avaliados. O primeiro preenchimento ocorreu no dia da avaliação fisioterapêutica e o segundo preenchimento foi realizado no dia de início da fisioterapia. Este intervalo entre o primeiro e segundo preenchimentos do ULFI foi de no mínimo 48h com o objetivo de minimizar o efeito de memória da primeira aplicação (TERWEE et al., 2012).

## **5.5 OUTROS QUESTIONÁRIOS**

No intuito de determinar a validade do construto, outros três questionários já validados para o português brasileiro e comumente utilizados em pacientes com disfunções do membro superior foram empregados no estudo.



O QuickDASH (ANEXO 3), assim como o ULFI, é um questionário preenchido pelo próprio paciente que avalia função e sintomas de pessoas com qualquer distúrbio musculoesquelético de todo o membro superior. Ele possui 11 itens e deriva do questionário DASH de 30 itens e sua confiabilidade, validade e responsividade já foram demonstradas em pacientes com distúrbios proximais e distais do membro superior. Cada item possui uma escala do tipo *Likert* que varia de 1 a 5 em que quanto maior o valor pior a função ou o sintoma. O score é obtido somando-se os itens assinalados, dividindo pelo número de itens completados. Desta razão, subtrai-se um e multiplica-se por 25 para gerar um valor numa escala de 0 a 100, na qual 100 representa maior comprometimento da função. Apenas um item em branco (10%) é permitido (BEATON et al., 2005). Optou-se pelo uso do QuickDash para diminuir o desgaste dos pacientes relacionado a quantidade e tempo de preenchimento dos questionários. O QuickDash demonstra forte correlação com o DASH original e sua versão brasileira demonstrou bons índices de confiabilidade, validade e responsividade (DA SILVA et al., 2020; PUGA et al., 2013).

O SF-36 (ANEXO 4) é um instrumento genérico de avaliação de qualidade de vida que avalia 8 domínios: capacidade funcional, aspectos físicos, dor, estado geral de saúde, vitalidade, aspectos sociais, aspectos emocionais e saúde mental. A pontuação de cada domínio é transformada em um score de 0 a 100 no qual zero corresponde a pior estado de saúde e 100 melhor estado de saúde. O SF-36 já foi adaptado culturalmente e validado para o português brasileiro (CICONELLI et al., 1999). Optou-se pelo uso de uma escala de qualidade de vida com o objetivo de avaliar o quanto o ULFI poderia se correlacionar com um construto relacionado, mas diferente daquele para o qual ele foi desenvolvido que é função de membro superior.

A END (ANEXO 5) é uma escala simples e de fácil mensuração que consiste em uma sequência de números, de 0 a 10, no qual o valor 0 representa “sem dor” e o número 10 representa “pior dor que se pode imaginar”. Desta forma, os indivíduos graduam a intensidade da sua dor nos últimos dias tendo como base esses parâmetros (FERREIRA-VALENTE; PAIS-RIBEIRO; JENSEN, 2011).

## 5.6 ANÁLISE ESTATÍSTICA

A estrutura interna do ULFI foi avaliada por meio da análise fatorial exploratória (AFE) e análise fatorial confirmatória (AFC). A AFE e AFC foram realizadas utilizando-se correlações policóricas e método de extração *Robust Diagonally Weighted Least Squares* (RDWLS) (ASPAROUHOV; MUTH; MUTHÉN, 2010). A análise paralela (AP) com permutação dos dados aleatórios foi utilizada para se determinar o número de fatores da escala (TIMMERMAN; LORENZO-SEVA, 2011).

A adequação do modelo foi avaliada pelos índices de ajuste Qui-quadrado ( $X^2$ ), Qui-quadrado/graus de liberdade ( $X^2/gl$ ), *Root Means Square Error of Approximation* (RMSEA), *Comparative Fit Index* (CFI), *Tucker-Lewis Index* (TLI) sendo considerados bons índices de ajuste o  $X^2$  não significativo,  $X^2/gl < 3$ , RMSEA menor que 0,08, CFI e TLI maior que 0,90 (SCHERMELLEH-ENGEL; MOOSBRUGGER; MÜLLER, 2003).

A confiabilidade foi avaliada com base em um modelo teste-reteste e por meio de duas propriedades: confiabilidade teste-reteste (erro relativo da medida) e concordância (erro absoluto da medida). Confiabilidade foi avaliada por meio do coeficiente de correlação intraclassa ( $CCI_{2,1}$ ). A interpretação do valor do CCI foi

realizada da seguinte forma: para valores abaixo de 0,40, a confiabilidade foi considerada baixa; entre 0,40 e 0,75, moderada; entre 0,75 e 0,90, substancial, e, finalmente, valores maiores que 0,90, a confiabilidade foi considerada excelente (FLEISS, 1986). Duas medidas foram usadas para avaliar a concordância: erro padrão da medida (EPM) e diferença mínima detectável (DMD) (WEIR, 2005).

O erro padrão da medida foi calculado pela fórmula:  $EPM = DP\sqrt{1-CCI}$ , em que DP é a média dos desvios padrão do teste e reteste. A DMD foi medida pela fórmula  $DMD_{90} = EPM \times 1,65 \times \sqrt{2}$  (GABEL et al., 2010).

A consistência interna foi calculada por meio do alfa de Cronbach para identificar se existem itens redundantes ou heterogêneos do questionário. Valores do alfa de Cronbach iguais ou maiores que 0,70 foram considerados adequados (PRINSEN et al., 2018).

Para se analisar a validade do construto, foi utilizado o coeficiente de correlação de Spearman ( $\rho$ ) para se determinar a magnitude de correlação entre o ULFI-Br e outros instrumentos. A interpretação da magnitude seguiu os seguintes critérios: baixa, de 0,26 a 0,49; moderada, de 0,50 a 0,69; alta, de 0,70 a 0,89; e muito alta, de 0,90 a 1,00 (VAN ARK et al., 2014).

Efeitos *ceiling* e *floor* foram avaliados no presente estudo. Por definição, esses efeitos ocorrem quando uma quantidade dos participantes do estudo (mais de 15%) atinge como escore total do questionário os valores mínimo ou máximo, o que indica um problema quando se avalia a responsividade do instrumento (TERWEE et al., 2007).

Dados sociodemográficos e valores dos escores dos questionários foram descritos por meio de média, mediana, desvio padrão (DP) e intervalo interquartil

para os dados quantitativos ou em número e porcentagem para os dados qualitativos.

Foram utilizados os softwares Factor 10.10.03 (Universitat Rovira i Virgili, Tarragona, Catalunha, Espanha) para a AFE, RStudio versão 1.1.453 (Boston, MA, EUA) e seus pacotes Lavaan versão 0.6-5 e SemPlot versão 1.1.2 para a AFC, e o SPSS versão 20 (Chicago, IL, EUA) para a análise das demais propriedades psicométricas e análise descritiva.

## **6. RESULTADOS**

Na etapa de tradução e adaptação transcultural, assim como na versão turca, optou-se por excluir a tradução do termo “*chop sticks*” (hashis) do item 20 por ser um utensílio da culinária oriental, pouco utilizado pela maioria da população brasileira e que poderia gerar dúvidas no preenchimento do questionário.

A versão pré-final do ULFI-Br foi então aplicada em 30 pacientes com disfunções musculoesqueléticas de membro superior. Nenhum dos itens do questionário foi incompreendido por mais de 20% dos pacientes. Assim, a versão pré-final do ULFI foi estabelecida como a versão final do ULFI no português brasileiro. Com relação a amostra do estudo, foram utilizados tamanhos amostrais diferentes a depender da análise estatística utilizada. Dos 200 pacientes inicialmente incluídos, para as análises fatoriais e descrição da amostra, foram utilizados dados de 190 pacientes, já que 10 foram excluídos por preenchimento incompleto do ULFI-Br; para validade de construto, foram utilizados dados de 180 pacientes, já que 10 pacientes foram excluídos dessa análise por não preencherem o QuickDash, SF-36

e/ou END; para confiabilidade, foi utilizada uma subamostra composta por 51 pacientes.

Assim sendo, a amostra composta por 190 pacientes apresentou média de idade de 52 anos (DP = 12), com disfunções musculoesqueléticas no membro superior em média há 42 meses (DP = 51). As características demográficas e clínicas estão descritas na Tabela 1.

**Tabela 1.** Características demográficas e clínicas da amostra total de pacientes incluídos no estudo (n = 190).

Características	Número absoluto (%)
Sexo feminino	155 (81,6%)
Estado civil	
Casado	101 (53,2%)
Solteiro	56 (29,5%)
Divorciado	19 (10%)
Viúvo	14 (7,3%)
Lado acometido	
Direito	61 (32,1%)
Esquerdo	33 (17,4%)
Bilateral	96 (50,5%)
Escolaridade	
Ensino básico (incompleto ou completo)	24 (12,6%)
Ensino médio (incompleto ou completo)	83 (43,7%)
Ensino superior (incompleto ou completo)	52 (27,4%)
Pós-graduação (incompleto ou completo)	31 (16,3%)
Diagnósticos	
Tendinopatia/bursite do ombro	103 (54,2%)
Síndrome do túnel do carpo	60 (31,6%)
Epicondilite do cotovelo	29 (15,3%)
Tendinopatia do punho/mão	15 (7,9%)
Dedo em gatilho	14 (7,4%)
Artrose punho/mão	9 (4,7%)
Tenossinovite de De Quervain	8 (4,2%)
Capsulite adesiva do ombro	3 (1,5%)
Outros	24 (12,6%)
Região acometida	
Proximal	75 (39%)
Distal	64 (34%)
Proximal e distal	22 (12%)
Intermediária	10 (5%)

Proximal e intermediária	9 (5%)
Proximal, intermediária e distal	6 (3%)
Intermediária e distal	4 (2%)

Fonte: autoria própria.

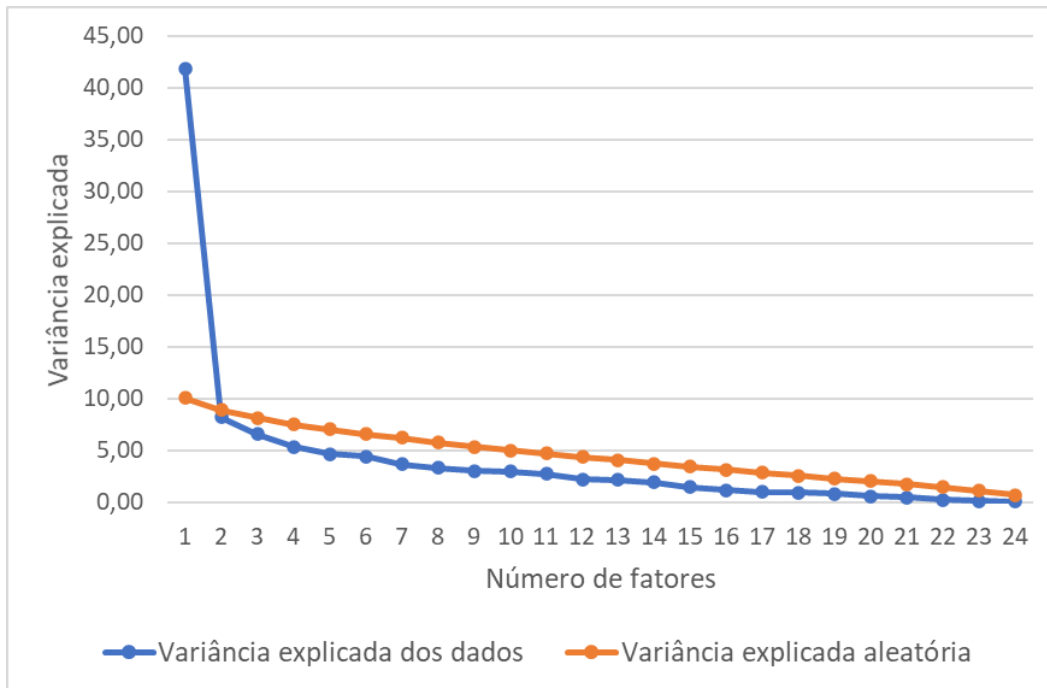
Com relação aos escores dos questionários utilizados neste estudo, a Tabela 2 descreve os dados de tendência central e dispersão. Para todos os questionários relacionados à função e dor, o escore foi, em média, próximo a 50% de acometimento.

**Tabela 2.** Escores dos questionários utilizados no presente estudo (n = 180).

Questionários	Média	Desvio padrão	Mediana	Intervalo interquartil
ULFI	53,5	21,5	52	30
QuickDash	49,38	20,82	47,73	31,25
SF-36				
Capacidade Funcional	48,39	24,04	45	40
Aspectos Físicos	29,17	37,04	0	50
Dor	39,9	21,27	41	31,75
Estado Geral de Saúde	58,63	21,06	57	30
Vitalidade	51,92	21,66	55	35
Aspectos Sociais	61,64	25,21	62,5	37
Aspectos Emocionais	39,62	41,97	33	100
Saúde Mental	65,07	20,16	64	31
Escala Numérica de Dor	6,66	2,31	7	3

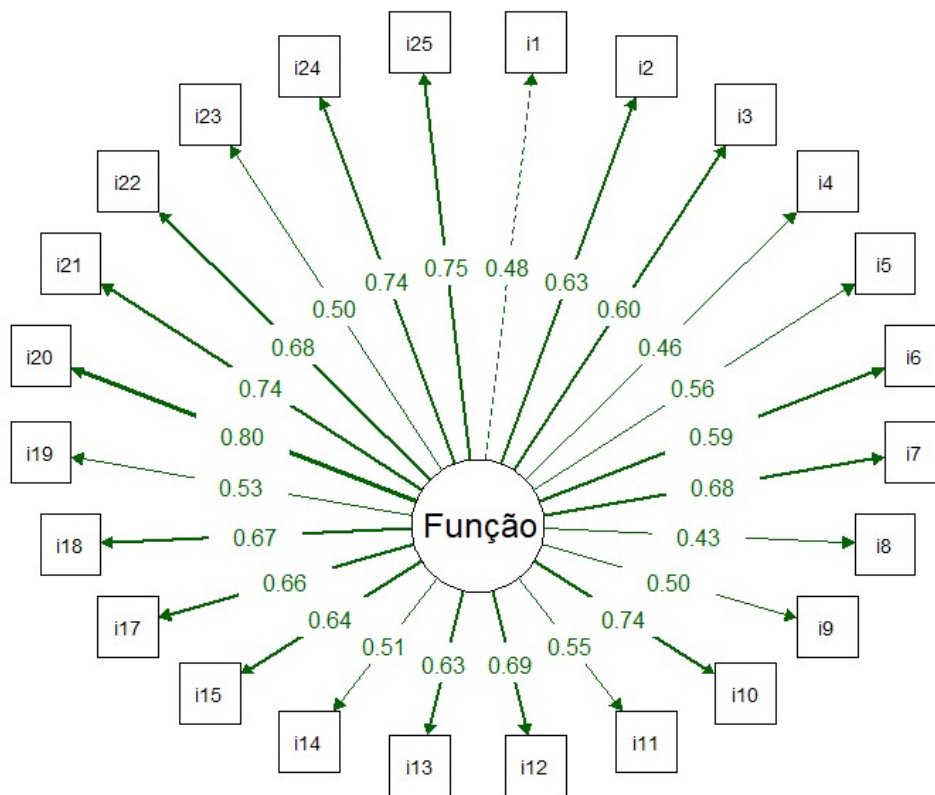
Fonte: autoria própria. ULFI: *Upper Limb Functional Index*; QuickDash: *Quick Disabilities of Arm, Shoulder and Hand*; SF-36: *36-item Short Form Health Survey*.

A AFE com implementação da análise paralela sugeriu uma estrutura unidimensional do ULFI, pois apenas um fator dos dados reais apresentou percentual de variância explicada maior do que os dados aleatórios (Figura 2). Os índices de ajuste da AFE foram adequados ( $X^2 = 424,427$ ,  $gl = 275$ ,  $p < 0,001$ ;  $X^2/gl = 1,54$ ;  $RMSEA = 0,054$ ;  $CFI = 0,973$ ;  $TLI = 0,975$ ).



**Figura 2.** Scree plot da análise paralela do *Upper Limb Functional Index* (n = 190). Apenas um fator da variância explicada dos dados é maior do que os fatores da variância explicada aleatória, indicando a retenção de apenas um fator. Fonte: autoria própria.

O modelo unifatorial do ULFI-Br foi testado pela AFC. Todos os 25 itens apresentaram cargas fatoriais maiores que 0,4 (Figura 3). Os índices de ajustes se mostraram aceitáveis para essa estrutura ( $X^2 = 441,860$ ,  $gl = 252$ ,  $p < 0,001$ ;  $X^2/gl = 1,75$ ;  $RMSEA = 0,063$ ;  $CFI = 0,918$ ;  $TLI = 0,910$ ).



**Figura 3.** Gráfico da Análise Fatorial Confirmatória do *Upper Limb Functional Index* com as cargas fatoriais (n = 190). Fonte: autoria própria.

Conforme mostra a Tabela 3, observou-se valores de alfa de Cronbach  $\geq 0,890$ , o que demonstra consistência interna adequada do ULFI-Br. Além disso, foi observado valor de CCI de 0,909, indicando excelente confiabilidade teste-reteste, com intervalo de tempo médio entre as avaliações de 9 dias. O erro padrão da medida foi de 6,11 e a DMD foi de 14,26.



**Tabela 3.** Consistência interna do *Upper Limb Functional Index* com apresentação do alfa de Cronbach total e se cada item for deletado (n = 51).

Item	Alfa de Cronbach se o item for deletado
Item 1	0,895
Item 2	0,893
Item 3	0,893
Item 4	0,895
Item 5	0,893
Item 6	0,893
Item 7	0,892
Item 8	0,897
Item 9	0,894
Item 10	0,890
Item11	0,894
Item 12	0,894
Item 13	0,892
Item 14	0,894
Item 15	0,891
Item 16	0,890
Item 17	0,892
Item 18	0,892
Item 19	0,894
Item 20	0,893
Item 21	0,892
Item 22	0,892
Item 23	0,895
Item 24	0,891
Item 25	0,890
<b>Total</b>	<b>0,897</b>

Fonte: autoria própria.

O ULFI demonstrou coeficientes de correlação ( $\rho$ ) com alta magnitude para o QuickDash, moderada magnitude para os domínios capacidade funcional, aspectos físicos, dor, estado geral de saúde, aspectos sociais do SF-36 e END, e baixa magnitude para os domínios vitalidade, aspectos emocionais e saúde mental do SF-36 (Tabela 4).

**Tabela 4.** Correlação entre os escores do *Upper Limb Functional Index* e os demais questionários do presente estudo (n = 180).

Questionários	ULFI	
	rho	p
QuickDash	-0,721	< 0,001
SF-36		
Capacidade Funcional	0,528	< 0,001
Aspectos Físicos	0,572	< 0,001
Dor	0,562	< 0,001
Estado Geral de Saúde	0,505	< 0,001
Vitalidade	0,361	< 0,001
Aspectos Sociais	0,557	< 0,001
Aspectos Emocionais	0,440	< 0,001
Saúde Mental	0,309	< 0,001
Escala Numérica de Dor	-0,529	< 0,001

Fonte: autoria própria. QuickDash: *Quick Disabilities of Arm, Shoulder and Hand*. SF-36: *36-item Short Form Health Survey*.

Quanto aos efeitos *floor* e *ceiling*, apenas um paciente atingiu o escore mínimo de 0 do ULFI-Br. Nenhum paciente atingiu o escore máximo de 100. Assim, não foram observados tais efeitos para o ULFI na versão em português brasileiro.

## 7. DISCUSSÃO

O presente estudo realizou a tradução, adaptação transcultural e validação do ULFI para o português brasileiro. O ULFI-Br demonstrou ser de fácil compreensão e possuir confiabilidade, consistência interna e validade de construtos adequados. Sua estrutura unidimensional permite o cálculo de um escore único, equivante ao fator função construto o qual o instrumento se propõe a medir.

A versão pré-final do ULFI-Br foi respondida por 30 pacientes que não apresentaram dificuldades de compreensão de nenhum dos itens, não sendo necessária a realização de mudanças. Assim, a versão pré-final pôde ser estabelecida como a versão final do ULFI-Br. Beaton et al. (2000) sugere que o

questionário possa ser entendido por uma criança de 12 anos (equivalente ao 6º ano do ensino fundamental). Gabel et al. (2010) verificaram que o ULFI foi mais fácil de ser lido e compreendido em relação ao QuickDash (nível de compreensão equivalentes ao 7º e 12º anos respectivamente).

A versão final do ULFI-Br foi respondida por uma amostra de 190 pacientes com uma grande variedade de disfunções musculoesqueléticas acometendo as diferentes regiões do membro superior. A proporção das regiões acometidas foi bem distribuída entre proximal (39%) e distal (34%), diferentemente dos outros estudos de validação do ULFI em que a região proximal foi predominantemente afetada (variando de 49% a 76%) (CUESTA-VARGAS; GABEL, 2013; IN et al., 2017; TONGA et al., 2015). Vinte e dois por cento dos pacientes apresentaram disfunção em mais de uma região, o que destaca a versatilidade de um questionário regional-específico como o ULFI para avaliação de pacientes com múltiplas queixas do membro superior.

A AFE revelou uma estrutura unidimensional do ULFI-Br assim como as versões original (GABEL et al., 2010) e espanhola (CUESTA-VARGAS; GABEL, 2013) e a AFC demonstrou índices de ajuste aceitáveis deste modelo. Por outro lado, as versões turca (TONGA et al., 2015) e a versão italiana (BRAVINI et al., 2015) encontraram uma estrutura de dois fatores, sendo que a italiana propôs a exclusão de 6 itens para se obter uma estrutura unidimensional de 19 itens. O método utilizado para determinação do número de fatores no presente estudo foi a análise paralela que tem se mostrado mais precisa do que os métodos empregados na maioria dos estudos de validação de questionários (K1, *scree plot* ou ambos). O K1 tende a superestimar o número de fatores o que pode desviar o foco para fatores menores em detrimento dos mais importantes, dificultar a interpretação e criar

fatores que provavelmente não sejam replicados em outros estudos. O *scree plot*, por ser um método visual, sofre de subjetividade e ambiguidade, especialmente quando gráfico não apresenta pontos claros de inflexão ou há dois ou mais pontos de inflexão (HAYTON; ALLEN; SCARPELLO, 2004).

O método de extração dos fatores utilizado neste estudo foi o RDWLS, diferente do estimador *maximum likelihood* (ML), comumente usando nas demais validações do ULFI. O ML é um estimador utilizado para variáveis contínuas e distribuição normal. Já o RDWLS é um estimador robusto para variáveis categóricas e distribuição não normal, que é o mais indicado para questionários com opções de resposta do tipo *Likert* (XIA; YANG, 2019).

Três itens apresentaram carga fatorial menor que 0,5 (item 1, “eu fico em casa a maior parte do tempo”; item 4, “eu descanso com maior frequência”; e item 8, “meu apetite agora está diferente”). Embora o ULFI seja denominado um questionário que avalia função, tais itens não avaliam este construto diretamente e se correlacionam mais com domínios de naturezas diferentes como participação, energia e apetite. Entretanto, esses 3 itens apresentaram carga fatorial aceitável ( $> 0,40$ ).

Na versão turca, 15 itens apresentaram carga fatorial menor que 0,5. Na versão original, 14 itens foram menores que 0,5 e na versão espanhola, 5 itens foram menores que 0,5. Não há um consenso de ponto de corte para se excluir ou se manter um item. As cargas fatoriais indicam em que medida o fator explica aquele item (variável observada) e quanto maior forem elas melhor. Cargas fatoriais muito baixas ( $< 0,4$ ) podem indicar potencial para exclusão do item do questionário.

Bravini et al. (2015) realizaram análise Rasch do ULFI e sugeriram uma redução para 19 itens para se obter uma estrutura unidimensional. Foram excluídos o item 4, item 8, item 9 (“caminhada, atividade esportiva ou recreativa está

prejudicada”), item 11 (“eu durmo mal”), item 14 (eu estou mais irritado e/ou mal-humorado”) e item 17 (“eu tenho dificuldade para colocar o braço na manga da camisa ou blusa ou preciso de ajuda para me vestir”). No presente estudo, os itens 4, 8, 9, 11 e 14 excluídos por Bravini et al. (2015) estão entre os itens que apresentaram menor carga fatorial do modelo (0,46, 0,43, 0,50, 0,55 e 0,51). Embora aparentemente sejam itens que não avaliam função do membro superior diretamente, eles apresentaram carga fatorial aceitável ( $>0,4$ ) e abordam aspectos contemplados nos códigos da Classificação Internacional de Funcionalidades (CIF) como energia (b1300), apetite (b1302), recreação/lazer (d920), sono (b134) e estabilidade psíquica (b1263). A CIF classifica a funcionalidade humana em três níveis: funcionalidade ao nível do copo ou parte dele, a pessoa como um todo e a pessoa como um todo em um contexto social. Portanto, consideramos que o ULFI-Br com todos os seus 25 itens avalia a funcionalidade do indivíduo de uma forma mais ampla.

A consistência interna do ULFI foi adequada ( $\alpha = 0,897$ ), o que demonstra que os itens da escala são correlacionados (homogêneos), ou seja, medem o mesmo construto.

A confiabilidade do ULFI foi alta no teste-reteste (CCI = 0,909), indicando um instrumento consistente ao longo do tempo. O CCI do presente estudo foi superior ao da versão turca (TONGA) (CCI = 0,87), mas inferior ao das versões original, espanhola, coreana e francesa-canadense (CCI = 0,93 a 0,94)(CUESTA-VARGAS; GABEL, 2013; GABEL et al., 2006; HAMASAKI; DEMERS; FILIATRAULT, 2015; IN et al., 2017). Tal diferença pode ser atribuída ao intervalo médio do teste-reteste de nosso estudo que foi de 9 dias comparado a 3 dias das demais versões. A avaliação da confiabilidade pelo teste-reteste pressupõe que os pacientes sejam estáveis entre

as avaliações. Os pacientes preencheram o ULFI-Br entre a primeira avaliação da fisioterapia e a primeira sessão de tratamento. Portanto, não receberam nenhum tratamento neste período. Além disso, eram pacientes crônicos, com menor possibilidade de alterar o quadro clínico no período. Mesmo assim, este intervalo maior pode ter contribuído para uma maior variação dos escores do ULFI entre as aplicações.

No contexto clínico, é importante se conhecer o valor da DMD, pois ela permite avaliar se a mudança ocorrida na função entre duas avaliações foi real ou se foi devida apenas a erro do instrumento. A DMD do ULFI no presente estudo foi de 14,26. Assim, apenas diferenças maiores que este valor entre escores em diferentes momentos podem ser consideradas mudanças reais da função do paciente. A DMD do ULFI-Br foi superior a DMD encontrada nas demais versões do ULFI. Mas ainda assim, podemos considerá-la aceitável, pois DMDs superiores a 10% são comumente encontradas em outros questionários (FROUD et al., 2018).

Foi analisada a validade de construto do ULFI-Br, pois não há um critério padrão ouro para se avaliar função de membro superior. O ULFI-Br, conforme hipotetizado, apresentou correlação inversa alta com o QuickDash ( $\rho = -0,721$ ), outro questionário regional-específico que também avalia função do membro superior. O ULFI-Br, embora contenha mais itens que o QuickDash, também apresenta características práticas (fácil compreensão e rápido preenchimento e cálculo do escore). Suas 3 opções de resposta (não, parcialmente e sim) são mais fáceis de compreender do que a escala *Likert* com 5 opções de resposta do QuickDash.

O ULFI apresentou correlação moderada com a END e os domínios capacidade funcional (CF), aspectos físicos (AF), dor, estado geral de saúde (EGS)

e aspectos sociais (AS) do SF-36 conforme nossa hipótese, mas correlação baixa com os domínios vitalidade (VIT), aspectos emocionais (AE) e saúde mental (SM) do mesmo questionário. O SF-36 é uma escala genérica de qualidade de vida que pode ser dividida em componentes físico (COF) e mental (COM). O COF é calculado utilizando-se os domínios CF, AF, dor e EGS e o COM, por sua vez, é gerado com os domínios VIT, AS, AE e SM. O COF, por ser composto por domínios de natureza mais física, provavelmente está mais correlacionado à função do que o COM. Silva et al. (2020) verificaram que o QuickDash, uma escala regional-específica que avalia função assim como ULFI, também apresentou correlação moderada com o COF e baixa com o COM do SF-12, que deriva do SF-36.

Não foram identificados efeitos *floor* e *ceiling*, assim como nos demais estudos de validação do ULFI-Br em outros idiomas. A maior parte dos pacientes apresentaram escores na faixa intermediária de 0 a 100 e apenas um apresentou escore mínimo de 0. Caso houvesse muitos pacientes em tratamento com escore mínimo de 0 (que indica pouca função) e eles pioram após o tratamento, o questionário não teria capacidade de detectar esta piora pois eles já se encontrariam no escore mínimo possível. Da mesma forma, caso pacientes com escore máximo de 100 melhorassem após um período, não seria possível detectar esta melhora por meio do questionário.

## **7.1 LIMITAÇÕES**

Este estudo apresenta algumas limitações a serem destacadas. Não foi avaliada a responsividade, uma propriedade psicométrica importante. Ela requer um delineamento longitudinal para verificar a capacidade de o questionário detectar

melhora clínica após um tratamento e estabelecer a diferença mínima clinicamente importante. A amostra do estudo foi composta por pacientes com disfunções musculoesqueléticas de caráter crônico (mais de 3 meses). O uso em pacientes agudos ou com disfunções do membro superior com causas diferentes (neurológicos, por exemplo) requer a validação para a população em questão.

## 8. CONCLUSÃO

O ULFI-Br traduzido, adaptado transculturalmente e validado para o português brasileiro apresenta confiabilidade, consistência interna e validade do construto adequadas, com estrutura fatorial unidimensional. É um instrumento prático que pode ser usado na avaliação de função de membro superior de indivíduos com diferentes condições musculoesqueléticas crônicas no Brasil.

## 9. REFERÊNCIAS

- AERTS, B. R. et al. Work-Related Questionnaire for UPper extremity disorders (WORQ-UP): Factor Analysis and Internal Consistency. **Archives of Physical Medicine and Rehabilitation**, v. 99, n. 9, p. 1818–1826, 2018.
- AERTS, B. R. J. et al. Development of a novel Work-Related Questionnaire for UPper extremity disorders (WORQ-UP). **International Archives of Occupational and Environmental Health**, v. 90, n. 8, p. 823–833, 2017.
- ALRENI, A. S. E. et al. Measures of upper limb function for people with neck pain. A systematic review of measurement and practical properties. **Musculoskeletal Science and Practice**, v. 29, p. 155–163, 2017.
- ALRUBAIY, L.; HUTCHINGS, H. A.; WILLIAMS, J. G. Assessing patient reported outcome measures: A practical guide for gastroenterologists. **United European Gastroenterology Journal**, v. 2, n. 6, p. 463–470, 2014.
- APOSTOLI, P. et al. Loads of housework? Biomechanical assessments of the upper limbs in women performing common household tasks. **International Archives of Occupational and Environmental Health**, v. 85, n. 4, p. 421–425, 2012.



- ARUMUGAM, V.; MACDERMID, J. C. Clinimetrics: Upper Extremity Functional Index. **Journal of Physiotherapy**, v. 64, n. 2, p. 125, 2018.
- ASPAROUHOV, T.; MUTH, B.; MUTHÉN, B. Simple Second Order Chi-Square Correction. Technical appendix. Retrieved from [https://www.statmodel.com/download/WLSMV\\_new\\_chi21.pdf](https://www.statmodel.com/download/WLSMV_new_chi21.pdf) [01.09.2016]. p. 1–8, 2010.
- BADALAMENTE, M. et al. Measurement scales in clinical research of the upper extremity, part 1: General principles, measures of general health, pain, and patient satisfaction. **Journal of Hand Surgery**, v. 38, n. 2, p. 401–406, 2013a.
- BADALAMENTE, M. et al. Measurement scales in clinical research of the upper extremity, part 2: Outcome measures in studies of the hand/wrist and shoulder/elbow. **Journal of Hand Surgery**, v. 38, n. 2, p. 407–412, 2013b.
- BARBOSA, R. E. C.; ASSUNÇÃO, A. Á.; ARAÚJO, T. M. DE. Distúrbios musculoesqueléticos em trabalhadores do setor saúde de Belo Horizonte, Minas Gerais, Brasil TT - Musculoskeletal disorders among healthcare workers in Belo Horizonte, Minas Gerais State, Brazil. **Cad Saude Publica**, v. 28, n. 8, p. 1569–1580, 2012.
- BARRO, D. et al. Job characteristics and musculoskeletal pain among shift workers of a poultry processing plant in Southern Brazil. **Journal of Occupational Health**, v. 57, n. 5, p. 448–456, 2015.
- BEATON, D. E. et al. Guidelines for the process of cross-cultural adaptation of self-report measures. **Spine**, v. 25, n. 24, p. 3186–3191, 2000.
- BEATON, D. E. et al. Development of the QuickDASH: COmparison of three item-reduction approaches. **Journal of Bone and Joint Surgery - Series A**, v. 87, n. 5, p. 1038–1046, 2005.
- BODIN, J.; HA, A. C.; DESCATHA, A. Comparison of Risk Factors for Shoulder Pain and Rotator Cuff Syndrome in the Working Population. **Am J Ind Med**, v. 615, n. December, p. 605–615, 2012.
- BRASIL. **Anuário Estatístico da Previdência Social AEPS 2016** Secretaria de Previdência, Empresa de Tecnologia e Informações da Previdência Brasília, 2016.
- BRAUN, Y. et al. The relationship between therapist-rated function and patient-reported outcome measures. **Journal of Hand Therapy**, v. 30, n. 4, p. 516–521, 2017.
- BRAVINI, E. et al. Classical test theory and Rasch analysis validation of the upper limb functional index in subjects with upper limb musculoskeletal disorders. **Archives of Physical Medicine and Rehabilitation**, v. 96, n. 1, p. 98–104, 2015.
- BROWN, T. A. **Methodology in the Social Sciences**. [s.l.: s.n.].
- CARDOSO, J. P. et al. Aspectos psicossociais do trabalho e dor musculoesquelética em professores. **Cadernos de Saúde Pública**, v. 27, n. 8, p. 1498–1506, 2011.
- CHENG, H. M. S. et al. Disabilities of the arm, shoulder and hand (DASH): Factor analysis of the version adapted to Portuguese/Brazil. **Disability and Rehabilitation**, v. 30, n. 25, p. 1901–1909, 2008.
- CHURCHILL, S. E. et al. The upper limb of Australopithecus sediba. **Science**, v. 340, n. 6129, p. 1233477, 2013.
- CICONELLI, R. M. et al. **Tradução para a língua portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36 (Brasil SF-36)** Revista Brasileira De Reumatologia, 1999. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-0032808921&partnerID=40&md5=8e372f8e7feece5ae4fc33228a55d3a8>>

- CUESTA-VARGAS, A. I.; GABEL, P. C. Cross-cultural adaptation, reliability and validity of the Spanish version of the upper limb functional index. **Health and Quality of Life Outcomes**, v. 11, n. 1, p. 1, 2013.
- DA SILVA, N. C. et al. Reliability, validity and responsiveness of Brazilian version of QuickDASH. **Musculoskeletal Science and Practice**, v. 48, p. 102163, 2020.
- DAMASIO, B. F. Uso da análise fatorial exploratória em psicologia. **Avaliação psicológica**, v. 11, n. 2, p. 213–228, 2012.
- DAVIDSON, M.; KEATING, J. Patient-reported outcome measures (PROMs): How should I interpret reports of measurement properties? A practical guide for clinicians and researchers who are not biostatisticians. **British Journal of Sports Medicine**, v. 48, n. 9, p. 792–796, 2014.
- DAVIS, A. M. et al. Measuring disability of the upper extremity: A rationale supporting the use of a regional outcome measure. **Journal of Hand Therapy**, v. 12, n. 4, p. 269–274, 1999.
- DE KLERK, S.; BUCHANAN, H.; JEROSCH-HEROLD, C. The validity and clinical utility of the Disabilities of the Arm Shoulder and Hand questionnaire for hand injuries in developing country contexts: A systematic review. **Journal of Hand Therapy**, v. 31, n. 1, p. 80–90.e1, 2018.
- EMERY, M.-P.; PERRIER, L.-L.; ACQUADRO, C. Patient-reported outcome and quality of life instruments database (PROQOLID): frequently asked questions. **Health and quality of life outcomes**, v. 3, p. 12, 2005.
- FAIRBANK, J. et al. The Oswestry low back pain disability questionnaire. **Physiotherapy**, v. 66, n. 8, p. 271–3, 1980.
- FERNANDES, R. D. C. P.; CARVALHO, F. M.; ASSUNÇÃO, A. Á. Prevalence of musculoskeletal disorders among plastics industry workers. **Cad Saude Publica**, v. 27, n. 1, p. 78–86, 2011.
- FERRAZ, M. B. et al. Upper-extremity Musculoskeletal Disorders in Keyboard Operators in Brazil: A Cross-sectional Study. **Int J Occup Environ Health**, v. 1, n. 3, p. 239–244, 1995.
- FERREIRA-VALENTE, M. A.; PAIS-RIBEIRO, J. L.; JENSEN, M. P. Validity of four pain intensity rating scales. **Pain**, v. 152, n. 10, p. 2399–2404, 2011.
- FLEISS, J. L. **The design and analysis of clinical experiments**. New York: Wiley, 1986.
- FRANCHIGNONI, F. et al. Suggestions for refinement of the disabilities of the arm, shoulder and hand outcome measure (DASH): A factor analysis and Rasch validation study. **Archives of Physical Medicine and Rehabilitation**, v. 91, n. 9, p. 1370–1377, 2010.
- FROUD, R. et al. Responsiveness, Reliability, and Minimally Important and Minimal Detectable Changes of 3 Electronic Patient-Reported Outcome Measures for Low Back Pain: Validation Study. **J Med Internet Research**, v. 20, n. 10, p. e272, 2018.
- GABEL, C. P. et al. The Upper Limb Functional Index: Development and Determination of Reliability, Validity, and Responsiveness. **Journal of Hand Therapy**, v. 19, n. 3, p. 328–349, 2006.
- GABEL, C. P. et al. A modified QuickDASH-9 provides a valid outcome instrument for upper limb function. **BMC Musculoskeletal Disorders**, v. 10, p. 1–11, 2009.
- GABEL, C. P. et al. Modification of the Upper Limb Functional Index to a Three-point Response Improves Clinimetric Properties. **Journal of Hand Therapy**, v. 23, n. 1, p. 41–52, 2010.
- GABEL, C. P. et al. Lower Limb Functional Index: Development and clinimetric properties.

**Physical Therapy**, v. 92, n. 1, p. 98–110, 2012.

GABEL, C. P. et al. The Spine Functional Index: development and clinimetric validation of a new whole-spine functional outcome measure. **Spine Journal**, v. 19, n. 2, p. e19–e27, 2013.

GUILLEMIN, F.; BOMBARDIER, C.; BEATON, D. Cross-cultural adaptation of health-related quality of life measures: literatura review and proposed guidelines. **J Clin Epidemiol**, v. 46, n. 12, p. 1417–1432, 1993.

HAMASAKI, T. et al. A cross-cultural adaptation of the Upper Limb Functional Index in French Canadian. **Journal of Hand Therapy**, v. 27, n. 3, p. 247–253, 2014.

HAMASAKI, T.; DEMERS, L.; FILIATRAULT, J. Test-retest reliability and responsiveness of a French Canadian Upper Limb Functional Index (ULFI-FC). **Disability and Rehabilitation**, v. 37, n. 12, p. 1090–1096, 2015.

HAMILTON, C. B.; CHESWORTH, B. M. A Rasch-Validated Version of the Upper Extremity Functional Index for Interval-Level Measurement of Upper Extremity Function. **Physical therapy**, v. 93, n. 11, p. 1507–1519, 2013.

HAYS, R. D. et al. Effects of mode and order of administration on generic health-related quality of life scores. **Value Health**, v. 12, n. 6, p. 1035–1039, 2009.

HAYTON, J. C.; ALLEN, D. G.; SCARPELLO, V. Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. **Organizational Research Methods**, v. 7, n. 2, p. 191–205, 2004.

HEMBECKER, P. K. et al. Investigation of musculoskeletal symptoms in a manufacturing company in Brazil: a cross-sectional study. **Brazilian Journal of Physical Therapy**, v. 21, n. 3, p. 175–183, 2017.

HU, L.-T.; BENTLER, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. **Structural Equation Modeling**, v. 6, n. 1, p. 1–55, 1999.

HUDAK, P. L.; AMADIO, P. C.; BOMBARDIER, C. Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder, and head). **American Journal of Industrial Medicine**, v. 29, n. 6, p. 602–608, 1996.

IN, T.-S. et al. The reliability and validity of the Korean version of the Upper Limb Functional Index. **Journal of Physical Therapy Science**, v. 29, n. 6, p. 1062–1065, 2017.

JAYAKUMAR, P. et al. What factors are associated with disability after upper extremity injuries? A systematic review. **Clinical Orthopaedics and Related Research**, v. 476, n. 11, p. 2190–2215, 2018.

KYTE, D. G. et al. An introduction to patient-reported outcome measures (PROMs) in physiotherapy. **Physiotherapy (United Kingdom)**, v. 101, n. 2, p. 119–125, 2015a.

KYTE, D. G. et al. An introduction to patient-reported outcome measures (PROMs) in physiotherapy. **Physiotherapy (United Kingdom)**, v. 101, n. 2, p. 119–125, 2015b.

LEVINE, D. W. et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome A Self-Administered Questionnaire for the Assessment of Severity of Symptoms and Functional in Carpal Tunnel Syndrome \*. 2011.

LUKE, A. et al. Musculoskeletal Injuries. In: **CURRENT Diagnosis & Treatment: Occupational & Environmental Medicine**. 5th. ed. New York: McGraw-Hill, 2013.

MARX, R. G. et al. A comparison of two time intervals for test-retest reliability of health status instruments. **Journal of Clinical Epidemiology**, v. 56, n. 8, p. 730–735, 2003.

- METCALF, C. et al. A review of clinical upper limb assessments within the framework of the WHO ICF. **Musculoskeletal Care**, v. 5, n. 3, p. 160–173, 2007.
- MICHENER, L. A.; LEGGIN, B. G. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. **Journal of Hand Therapy**, v. 14, n. 2, p. 68–76, 2001.
- MOKKINK, L. B. et al. COSMIN methodology for systematic reviews of Patient - Reported Outcome Measures ( PROMs ). User Manual. n. February, p. 1–78, 2018a.
- MOKKINK, L. B. et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. **Quality of Life Research**, v. 27, n. 5, p. 1171–1179, 2018b.
- MOKKINK, L. B. et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. **Department of Epidemiology and Biostatistics Amsterdam Public Health research institute Amsterdam University Medical Centers, location VUmc**, n. July, p. 1–32, 2019.
- ORFALE, A. G. et al. Translation into brazilian portuguese, cultural adaptation and evaluation of the reliability of the Disabilities of the Arm, Shoulder and Hand questionnaire. **Brazilian Journal of Medical and Biological Research**, v. 38, n. 2, p. 293–302, 2005.
- PRANSKY, G. et al. Measuring Functional Outcomes in Work-Related Upper Extremity Disorders. Development and Validation of the Upper Extremity Function Scale. **Journal of Occupational & Environmental Medicine**, v. 39, n. 12, p. 1195–1202, 1997.
- PRINSEN, C. A. C. et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. **Quality of Life Research**, v. 27, n. 5, p. 1147–1157, 2018.
- PUGA, V. O. D. O. et al. Clinimetric Testing Supports the Use of 5 Questionnaires Adapted Into Brazilian Portuguese for Patients With Shoulder Disorders. **Journal of Orthopaedic & Sports Physical Therapy**, v. 43, n. 6, p. 404–413, 2013.
- RABIN, R.; CHARRO, F. DE. EQ-SD: a measure of health status from. **Ann Med**, v. 33, p. 337–343, 2001.
- RODRIGUES, E. K. DA S.; FONSECA, M. DE C. R.; MACDERMID, J. C. Brazilian version of the Patient Rated Wrist Evaluation (PRWE-BR): Cross-cultural adaptation, internal consistency, test-retest reliability and construct validity. **Journal of Hand Therapy**, v. 28, n. 1, p. 69–76, 2015.
- SALERNO, D. F. et al. A review of functional status measures for workers with upper extremity disorders. **Occupational and Environmental Medicine**, v. 59, n. 10, p. 664–670, 2002.
- SANTOS FILHO, S. B.; BARRETO, S. M. Atividade ocupacional e prevalência de dor osteomuscular em cirurgiões-dentistas de Belo Horizonte, Minas Gerais, Brasil: contribuição ao debate sobre os distúrbios osteomusculares relacionados ao trabalho. **Cadernos de Saude Publica**, v. 17, n. 1, p. 181–193, 2001.
- SARTORIO, F. et al. [CROSS-CULTURAL ADAPTATION, AND VALIDITY OF THE ITALIAN VERSION OF THE UPPER LIMB FUNCTIONAL INDEX (ULFI-I)]. **Giornale italiano di medicina del lavoro ed ergonomia**, v. 37, n. 2, p. 115–119, 2015.
- SCHERMELLEH-ENGEL, K.; MOOSBRUGGER, H.; MÜLLER, H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. **MPR-online**, v. 8, n. May, p. 23–74, 2003.
- SCHREIBER, J. B. et al. Reporting structural equation modeling and confirmatory factor analysis results: A review. **Journal of Educational Research**, v. 99, n. 6, p. 323–338, 2006.
- SOUZA, A. C. DE; ALEXANDRE, N. M. C.; GUIRARDELLO, E. DE B. Propriedades

- psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. **Epidemiologia e serviços de saúde : revista do Sistema Unico de Saude do Brasil**, v. 26, n. 3, p. 649–659, 2017.
- STOCK, S. R.; STREINER, D.; REARDON, R. The impact of neck and upper limb musculoskeletal disorders on the lives of affected workers: development of a new functional status index. **Qual Life Res**, v. 4, p. 491, 1995.
- STRATFORD, P. W.; BINKLEY, J. M.; STRATFORD, D. M. Development and initial validation of the Upper Extremity Functional Index. **Physiother Can**, v. 53, p. 259–267, 2001.
- TERWEE, C. B. et al. Quality criteria were proposed for measurement properties of health status questionnaires. **Journal of Clinical Epidemiology**, v. 60, n. 1, p. 34–42, 2007.
- TERWEE, C. B. et al. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. **Quality of Life Research**, v. 21, n. 4, p. 651–657, 2012.
- TERWEE, C. B. et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. **Quality of Life Research**, v. 27, n. 5, p. 1159–1170, 2018.
- TIMMERMAN, M. E.; LORENZO-SEVA, U. Dimensionality assessment of ordered polytomous items with parallel analysis. **Psychological Methods**, v. 16, n. 2, p. 209–220, 2011.
- TONGA, E. et al. Cross-cultural adaptation, reliability and validity of the Turkish version of the Upper Limb Functional Index (ULFI). **Journal of Hand Therapy**, v. 28, n. 3, p. 279–285, 2015.
- VAN ARK, M. et al. Cross-cultural adaptation and reliability and validity of the Dutch Patient-Rated Tennis Elbow Evaluation (PRTEE-D). **BMC Musculoskeletal Disorders**, v. 15, n. 1, 2014.
- VAN KAMPEN, D. A. et al. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). **Journal of Orthopaedic Surgery and Research**, v. 8, n. 1, p. 1–9, 2013.
- VARNI, J. W.; LIMBERS, C. A.; NEWMAN, D. A. Using factor analysis to confirm the validity of children's self-reported health-related quality of life across different modes of administration. **Clinical Trials**, v. 6, n. 2, p. 185–195, 2009.
- WEIR, J. P. QUANTIFYING TEST-RETEST RELIABILITY USING THE INTRAClass CORRELATION COEFFICIENT AND THE SEM. **Journal of Strength and Conditioning Research**, v. 19, n. 1, p. 231–240, 2005.
- WORMALD, J. C. R. et al. Site-specific Patient-reported Outcome Measures for Hand Conditions. **Plastic and Reconstructive Surgery - Global Open**, v. 7, n. 5, p. e2256, 2019.
- XIA, Y.; YANG, Y. RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. **Behavior Research Methods**, v. 51, n. 1, p. 409–428, 2019.

## ANEXO 1 - PARECER CONSUBSTANCIADO DO CEP



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Adaptação Transcultural e Validação do Upper Limb Functional Index (ULFI) para o Português Brasileiro

**Pesquisador:** Henrique Yuji Takahasi

**Área Temática:**

**Versão:** 2

**CAAE:** 94496318.4.0000.0022

**Instituição Proponente:** ASSOCIACAO DAS PIONEIRAS SOCIAIS

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

**Número do Parecer:** 2.990.249

#### Apresentação do Projeto:

Trata-se de um estudo de validação do questionário Upper Limb Functional Index (ULFI) realizado conforme as Guidelines for the Process of Cross-cultural Adaptation of Self-Report Measures (BEATON et al., 2000) e Consensus-based Standards for the Selection of Health Measurement Instruments (MOKKINK et al., 2010). O ULFI, assim é um questionário regional-específico para o membro superior que apresenta boas propriedades psicométricas e alta correlação com o DASH, mas com maior praticidade pelos tempos de preenchimento e pontuação reduzidos (GABEL et al., 2006). Sua versão original é composta por 25 questões e opções dicotômicas de resposta. Posteriormente, foi desenvolvida a versão modificada com 3 opções de resposta que demonstrou ainda maior confiabilidade sem perder sua praticidade (GABEL et al., 2010). O Upper Limb Functional Index (ULFI) é uma PROM regional-específica, mas não há versão em português.

A autorização para realização da tradução do ULFI para o português brasileiro foi concedida via e-mail por um dos autores do questionário (Dr. Charles Philip Gabel). Medidas de resultados relatados pelo paciente ou patient-reported outcome measures (PROMs) são instrumentos de avaliação centrados na percepção do paciente de seus sintomas que têm sido cada vez mais utilizados na prática clínica e em pesquisas.

O estudo será realizado em três fases: (1) tradução do questionário, (2) teste da versão pré-final da versão traduzida do ULFI para o português brasileiro, e (3) validação da versão final do ULFI

Endereço: SMHS Quadra 301 Bloco B nº 45 Entrada A Edifício Pioneiras Sociais - Terceiro Andar  
 Bairro: SMHS CEP: 70.334-900  
 UF: DF Município: BRASILIA  
 Telefone: (61)3319-1494 E-mail: comiteeticapesquisa@sarah.br

## ANEXO 2 – ULFI ORIGINAL

<b>UPPER LIMB FUNCTIONAL INDEX</b>		DATE: _____	
NAME: _____	INJURY: _____	LEFT ARM	RIGHT ARM
<p><b>PLEASE COMPLETE:</b> Your arm may make it difficult to do some things you normally do. This list contains sentences people use to describe themselves with such problems. Think of yourself over the last few days. <b>If an item describes you, mark the Box 'Partly' or 'Yes'. If an item does not describe you, Mark the Box 'NO'.</b></p> <p style="text-align: center;"><b>DUE TO MY ARM:</b></p>			
<b>NO Part YES</b>			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1. I stay at home most of the time.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2. I change position frequently for comfort.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3. I avoid heavy jobs e.g. cleaning, lifting more than 5kg or 10lbs, gardening, etc.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4. I rest more often.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5. I get others to do things for me.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. I have the pain / problem almost all the time.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7. I have difficulty lifting and carrying (e.g. bags, shopping up to 5kg or 10lbs).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8. My appetite is now different.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9. My walking or normal recreation or sporting activity is affected.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10. I have difficulty with normal home or family duties and chores.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11. I sleep less well.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12. I need assistance with personal care e.g. washing and hygiene
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13. My regular daily activities (work, social contact) are affected.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14. I am more irritable and / or bad tempered.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	15. I feel weaker and / or stiffer.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	16. My transport Independence is affected (driving, public transport).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	17. I have difficulty putting my arm into a shirt sleeves or need assistance dressing.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	18. I have difficulty writing or using a key board and / or 'mouse'.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	19. I am unable to do things at or above shoulder height.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	20. I have difficulty eating and / or using utensils (e.g. knife, fork, spoon, chop sticks).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	21. I have difficulty holding and moving dense objects (e.g. mugs, jars, cans).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	22. I tend to drop things and / or have minor accidents more frequently.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	23. I use the other arm more often.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	24. I have difficulty with buttons, keys, coins, taps / faucets, containers or screw-top lids.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	25. I've difficulty opening, holding, pushing or pressing (eg. triggers, levers, heavy doors ...).
<b>ULFI SCORE: To score the upper part - add the marks:</b>			
<input style="width: 50px;" type="text"/>	<b>TOTAL (ULFI points). (x4) = _____ 100 Scale</b>		<b>FINAL TOTAL (100 – ULFIx4) = <input style="width: 50px;" type="text"/> %</b>
<b>MDC (90% confidence): 7.9 % or 1.9 ULFI points. Change less than this may be due to error.</b>			

## ANEXO 3 – QUICKDASH

### **QuickDASH**

Por favor meça sua habilidade para realizar as seguintes atividades na semana passada circulando o número apropriado da resposta:

	Não houve dificuldade	Houve pouca dificuldade	Houve dificuldade moderada	Dificuldade severa	Não consegui fazer
1. Abrir um vidro novo ou com a tampa muito apertada.	1	2	3	4	5
2. Fazer tarefas domésticas pesadas (por exemplo: lavar paredes, lavar o chão).	1	2	3	4	5
3. Carregar uma sacola ou uma maleta.	1	2	3	4	5
4. Lavar suas costas.	1	2	3	4	5
5. Usar uma faca para cortar alimentos.	1	2	3	4	5
6. Atividades recreativas que exigem alguma força ou impacto nos braços, ombros ou mãos (por exemplo: jogar vôlei, martelar).	1	2	3	4	5

	Não afetou	Afetou pouco	Afetou Moderadamente	Afetou muito	Afetou Extrema Mente
7. Durante a semana passada, em que ponto o seu problema com braço, ombro ou mão afetaram suas atividades normais com família, amigos, vizinhos ou colegas?	1	2	3	4	5



	Não limitou	Limitou pouco	Limitou moderadamente	Limitou muito	Não conseguiu fazer
8. Durante a semana passada, o seu trabalho ou outras atividades diárias regulares foram limitadas devido ao seu problema com braço, ombro ou mão?	1	2	3	4	5

Por favor meça a gravidade dos seguintes sintomas na semana passada. (circule o número)	Nenhuma	Pouca	Moderada	Severa	Extrema
9. Dor no braço, ombro ou mão.	1	2	3	4	5
10. Desconforto na pele (alfinetadas) no braço, ombro ou mão.	1	2	3	4	5
	Não houve dificuldade	Pouca dificuldade	Dificuldade moderada	Dificuldade severa	Tão difícil que eu não pude dormir
11. Durante a semana passada, quanto de dificuldade você teve para dormir por causa da dor no seu braço, ombro ou mão? (circule o número)	1	2	3	4	5

## ANEXO 4 – SF-36

### Versão Brasileira do Questionário de Qualidade de Vida -SF-36

1- Em geral você diria que sua saúde é:

Excelente	Muito Boa	Boa	Ruim	Muito Ruim
1	2	3	4	5

2- Comparada há um ano atrás, como você se classificaria sua saúde em geral, agora?

Muito Melhor	Um Pouco Melhor	Quase a Mesma	Um Pouco Pior	Muito Pior
1	2	3	4	5

3- Os seguintes itens são sobre atividades que você poderia fazer atualmente durante um dia comum. Devido à sua saúde, você teria dificuldade para fazer estas atividades? Neste caso, quando?

Atividades	Sim, dificulta muito	Sim, dificulta um pouco	Não, não dificulta de modo algum
a) Atividades Rigorosas, que exigem muito esforço, tais como correr, levantar objetos pesados, participar em esportes árduos.	1	2	3
b) Atividades moderadas, tais como mover uma mesa, passar aspirador de pó, jogar bola, varrer a casa.	1	2	3
c) Levantar ou carregar mantimentos	1	2	3
d) Subir vários lances de escada	1	2	3
e) Subir um lance de escada	1	2	3
f) Curvar-se, ajoelhar-se ou dobrar-se	1	2	3
g) Andar mais de 1 quilômetro	1	2	3
h) Andar vários quarteirões	1	2	3
i) Andar um quarteirão	1	2	3
j) Tomar banho ou vestir-se	1	2	3

4- Durante as últimas 4 semanas, você teve algum dos seguintes problemas com seu trabalho ou com alguma atividade regular, como consequência de sua saúde física?

	Sim	Não
a) Você diminui a quantidade de tempo que se dedicava ao seu trabalho ou a outras atividades?	1	2
b) Realizou menos tarefas do que você gostaria?	1	2
c) Esteve limitado no seu tipo de trabalho ou a outras atividades.	1	2
d) Teve dificuldade de fazer seu trabalho ou outras atividades (p. ex. necessitou de um esforço extra).	1	2

5- Durante as últimas 4 semanas, você teve algum dos seguintes problemas com seu trabalho ou outra atividade regular diária, como consequência de algum problema emocional (como se sentir deprimido ou ansioso)?

	Sim	Não
a) Você diminui a quantidade de tempo que se dedicava ao seu trabalho ou a outras atividades?	1	2
b) Realizou menos tarefas do que você gostaria?	1	2
c) Não realizou ou fez qualquer das atividades com tanto cuidado como geralmente faz.	1	2

6- Durante as últimas 4 semanas, de que maneira sua saúde física ou problemas emocionais interferiram nas suas atividades sociais normais, em relação à família, amigos ou em grupo?

De forma nenhuma	Ligeiramente	Moderadamente	Bastante	Extremamente
1	2	3	4	5

7- Quanta dor no corpo você teve durante as últimas 4 semanas?

Nenhuma	Muito leve	Leve	Moderada	Grave	Muito grave
1	2	3	4	5	6

8- Durante as últimas 4 semanas, quanto a dor interferiu com seu trabalho normal (incluindo o trabalho dentro de casa)?

De maneira alguma	Um pouco	Moderadamente	Bastante	Extremamente
1	2	3	4	5

9- Estas questões são sobre como você se sente e como tudo tem acontecido com você durante as últimas 4 semanas. Para cada questão, por favor dê uma resposta que mais se aproxime de maneira como você se sente, em relação às últimas 4 semanas.

	Todo Tempo	A maior parte do tempo	Uma boa parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nunca
a) Quanto tempo você tem se sentindo cheio de vigor, de vontade, de força?	1	2	3	4	5	6
b) Quanto tempo você tem se sentido uma pessoa muito nervosa?	1	2	3	4	5	6
c) Quanto tempo você tem se sentido tão deprimido que nada pode anima-lo?	1	2	3	4	5	6
d) Quanto tempo você tem se sentido calmo ou tranqüilo?	1	2	3	4	5	6
e) Quanto tempo você tem se sentido com muita energia?	1	2	3	4	5	6
f) Quanto tempo você tem se sentido desanimado ou abatido?	1	2	3	4	5	6
g) Quanto tempo você tem se sentido esgotado?	1	2	3	4	5	6
h) Quanto tempo você tem se sentido uma pessoa feliz?	1	2	3	4	5	6
i) Quanto tempo você tem se sentido cansado?	1	2	3	4	5	6

10- Durante as últimas 4 semanas, quanto de seu tempo a sua saúde física ou problemas emocionais interferiram com as suas atividades sociais (como visitar amigos, parentes, etc)?

Todo Tempo	A maior parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nenhuma parte do tempo
1	2	3	4	5

**11- O quanto verdadeiro ou falso é cada uma das afirmações para você?**

	Definitivamente verdadeiro	A maioria das vezes verdadeiro	Não sei	A maioria das vezes falso	Definitivamente falso
a) Eu costumo adoecer um pouco mais facilmente que as outras pessoas	1	2	3	4	5
b) Eu sou tão saudável quanto qualquer pessoa que eu conheço	1	2	3	4	5
c) Eu acho que a minha saúde vai piorar	1	2	3	4	5
d) Minha saúde é excelente	1	2	3	4	5

**ANEXO 5 – ESCALA NUMÉRICA DA DOR**

0 1 2 3 4 5 6 7 8 9 10  
Sem dor Pior dor que se  
pode imaginar

## APÊNDICE 1 – VERSÃO FINAL DO ULFI-BR

<b>UPPER LIMB FUNCTIONAL INDEX (ULFI-Br)</b>			DATA: _____
NOME: _____	LESÃO: _____	<input type="checkbox"/> BRAÇO ESQUERDO	<input type="checkbox"/> DIREITO

**POR FAVOR, COMPLETE:** seu braço pode dificultar a realização de algumas coisas que você normalmente faz. Essa lista contém algumas frases que as pessoas usam para descrever tais problemas. Pense em você nos últimos dias. Se um item o(a) descreve, marque a caixa “Parcialmente” ou “SIM”. Se um item não descreve você, marque a caixa “NÃO”.

### POR CAUSA DO MEU BRAÇO:

NÃO Parcialmente SIM

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1. Eu fico em casa a maior parte do tempo.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2. Eu mudo de posição frequentemente para ter mais conforto.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3. Eu evito tarefas pesadas, por ex. limpeza, levantar mais de 5 kg, jardinagem, etc.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4. Eu descanso com mais frequência.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5. Eu peço que outras pessoas façam as coisas para mim.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. Eu tenho dor/problema quase o tempo todo.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7. Eu tenho dificuldade em levantar e carregar (por ex. sacolas, compras até 5 kg).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8. Meu apetite agora está diferente.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9. Caminhada, atividade recreativa ou esportiva está prejudicada.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10. Eu tenho dificuldade com tarefas domésticas ou familiares.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11. Eu durmo mal.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12. Eu preciso de ajuda com cuidados pessoais, por ex. com banho e higiene pessoal.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13. Minhas atividades diárias (trabalho, contato social) estão prejudicadas.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14. Eu estou mais irritado(a) e/ou mal-humorado(a).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	15. Eu sinto mais fraqueza e/ou rigidez.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	16. Minha independência para locomoção está prejudicada (dirigir, usar transporte público).

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	17. Eu tenho dificuldade para colocar meu braço na manga da camisa ou blusa ou preciso de ajuda para me vestir.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	18. Eu tenho dificuldade para escrever ou utilizar o teclado e/ou o mouse de computador.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	19. Eu não consigo realizar tarefas na altura do ombro ou acima dele.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	20. Eu tenho dificuldade em comer e/ou usar utensílios (por ex. faca, garfo, colher).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	21. Eu tenho dificuldade em segurar e mover objetos compactos (por ex. canecas, potes, latas).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	22. Eu costumo deixar coisas caírem e/ou sofrer pequenos acidentes com mais frequência.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	23. Eu uso o outro braço com mais frequência.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	24. Eu tenho dificuldade com botões, chaves, moedas, torneiras, recipientes ou tampas de rosca.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	25. Eu tenho dificuldade em abrir, segurar, empurrar ou pressionar (por ex. gatilhos, alavancas, portas pesadas).

**ESCORE ULFI:** para pontuar a parte superior, adicione os pontos:

**TOTAL** (pontos ULFI).

$100 - (\text{TOTAL} \times 4) =$   %

**Diferença mínima detectável (90% de confiança):** 14,26. Mudanças menores podem ser devido a erro.