



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Escola Nacional de Saúde Pública Sergio Arouca



“Instrumentos para mensuração de satisfação de usuários de serviços de saúde - uma adaptação transcultural do instrumento MedRisk para o contexto brasileiro”

por

Renato Santos de Almeida

Tese apresentada com vistas à obtenção do título de Doutor em Ciências na área de Saúde Pública.

Orientador: Prof.^a Dr.^a Mônica Silva Martins

Rio de Janeiro, maio de 2015.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Escola Nacional de Saúde Pública Sergio Arouca



Esta tese, intitulada

“Instrumentos para mensuração de satisfação de usuários de serviços de saúde - uma adaptação transcultural do instrumento MedRisk para o contexto brasileiro”

apresentada por

Renato Santos de Almeida

foi avaliada pela Banca Examinadora composta pelos seguintes membros:

Prof. Dr. Michael Eduardo Reichenheim

Prof.^a Dr.^a Anke Bergmann

Prof.^a Dr.^a Cláudia Tartaglia Reis

Prof.^a Dr.^a Margareth Crisóstomo Portela

Prof.^a Dr.^a Mônica Silva Martins – Orientadora

Tese defendida e aprovada em 14 de maio de 2015.

Catálogo na fonte
Instituto de Comunicação e Informação Científica e Tecnológica
Biblioteca de Saúde Pública

- A447i Almeida, Renato Santos de
Instrumentos para mensuração de satisfação de usuários de serviços de saúde - uma adaptação transcultural do instrumento MedRisk para o contexto brasileiro. / Renato Santos de Almeida. -- 2015.
162 f. : tab.
- Orientador: Mônica Silva Martins
Tese (Doutorado) – Escola Nacional de Saúde Pública Sergio Arouca, Rio de Janeiro, 2015.
1. Qualidade da Assistência à Saúde. 2. Assistência Centrada no Paciente. 3. Satisfação do Paciente. 4. Serviços de Saúde. 5. Fisioterapia. 6. Estudos de Validação. 7. Psicometria. I. Título.

CDD – 22.ed. – 362.1

À minha família e meus pacientes, que são o alicerce de minha passagem por esta vida.

AGRADECIMENTOS

Inicio agradecendo aos meus pais, que conseguiram, ao longo de suas vidas, fazer com que eu me orgulhasse sempre da família na qual nasci. Obrigado minha mãe, Maria da Conceição Santos de Almeida, pelo amor incondicional a mim dedicado. Agradeço ao meu pai, Gilberto Lima de Almeida, *in memoriam*, pelo homem incansável que fora para nos criar, educar, dar amor e conforto. Agradeço ainda aos meus avós, pelo carinho com que me trataram sempre. Ao meu irmão, pela amizade de todos esses anos e aos sobrinhos lindos que me deu.

Um agradecimento especial à minha esposa, Bruna Couri. Companheira para todas as horas e fundamental para meu crescimento como homem. Sem ela certamente minha vida teria menos brilho!

Ao meu amigo e mentor, Leandro Alberto Calazans Nogueira. Muito do que conquistei devo a esta pessoa.

À minha orientadora Mônica Martins, que, sem me conhecer, aceitou minha proposta de tese e esteve sempre ao meu lado me encorajando e me ensinando. Ao professor Michael Reichenheim, por sua orientação técnica, gentileza e ensinamentos valiosos durante toda construção desta tese.

Aos amigos da fisioterapia do Hospital Universitário Gaffrée e Guinle. Tenho um carinho enorme por este hospital, pois nestes 13 anos em que lá trabalhei, venho colhendo conquistas profissionais e laços de amizade. Aos estagiários do serviço, que foram fundamentais e colaboraram no trabalho de campo.

Aos pacientes, que fizeram parte deste trabalho e também aos demais, que fazem parte da minha história como fisioterapeuta clínico, pesquisador e ser-humano.

À Coordenação do curso de fisioterapia do UNIFESO em Teresópolis, pela compreensão nos momentos em que precisei me ausentar por conta da tese.

RESUMO

A satisfação dos pacientes é um eixo importante relacionado ao cuidado centrado no paciente e, conseqüentemente, está imbricada nas discussões referentes à qualidade do cuidado. Existem diferentes instrumentos para mensuração da satisfação disponíveis para diversas especialidades do campo da saúde. Especificamente na fisioterapia, o instrumento MedRisk (composto por duas dimensões e 12 itens) se destaca atualmente pelo interesse de diferentes grupos de pesquisadores. O objetivo do estudo foi conduzir o processo de adaptação transcultural do instrumento Medrisk para mensuração da satisfação do paciente para o contexto brasileiro. Assim, a metodologia foi desenvolvida em duas etapas: “(i) Equivalência conceitual, equivalência de itens e equivalência semântica; (ii) Equivalência de mensuração - psicometria (avaliação da validade dimensional, itens componentes e avaliação da confiabilidade). Na primeira etapa, para avaliação das traduções e pertinência dos itens, foi utilizado um painel de especialistas e um pré-teste do instrumento com 55 pacientes do Hospital Universitário Gaffrée e Guinle. Na segunda etapa, foi realizada a análise fatorial exploratória/confirmatória para identificação das dimensões e itens componentes do instrumento, com dados coletados de 325 indivíduos do mesmo serviço onde foi realizado o pré-teste. Foi realizada ainda a avaliação da confiabilidade do instrumento, índice de correlação intra-classe e medidas de erro. Na etapa 1, o painel de especialistas propôs a retirada de dois itens devido ao contexto de aplicação do instrumento. O pré-teste identificou dificuldade de compreensão de um item e após modificação semântica, o problema foi minimizado. O resultado das análises indicaram um instrumento com 6 itens específicos e duas medidas globais de satisfação. Três itens foram retidos na dimensão 1- “Interpessoal” e três na dimensão 2- “Estrutura física e apoio ao serviço”. A confiabilidade mostrou valores excelentes para as medidas de teste/reteste (0,96 para dimensão 1 e 0,98 para dimensão 2), entretanto o alpha de Cronbach mostrou valores moderados (0,60 para dimensão 1 e 0,63 para dimensão 2). As medidas de erro também apresentaram valores adequados quanto à concordância. Embora o plano de análise tenha seguido os mesmos passos dos estudos anteriores sobre a validade do MedRisk, os achados são discordantes de versões precedentes, incluindo a adaptação para contexto brasileiro. Limitações no aspecto processual do histórico psicométrico do instrumento foram identificadas, indicando a necessidade de outras avaliações.

Palavras-chave: Qualidade de cuidado, Cuidado centrado no paciente, Satisfação do paciente; Fisioterapia; Estudos de validação, Psicometria.

ABSTRACT

Patient satisfaction is an important construct for patient-centered care concept and it is usually embedded in discussions regarding the quality of care. There are different instruments to measure patient satisfaction available in various specialties of healthcare. Specifically in the physiotherapy the MedRisk instrument gain currently notoriety because the common interest form different researchers groups. The objective of the study was to conduct the MedRisk cross-cultural adaptation for measuring patient satisfaction in the Brazilian context. Therefore, methods have been developed in two stages: (i) Conceptual equivalence, item equivalence, and semantic equivalence; (ii) Measurement Equivalence - psychometry (dimensional validity evaluation, items and reliability). In the first stage it were evaluated the pertinence of the items, relevance of the content for the target population and also the comprehension of the items after the translation to the portuguese language. For this we used a panel of experts and a pre-testing of the instrument in 55 patients from Gaffrée and Guinle Universtity Hospital. In the second stage it was performed an exploratory/confirmatory factor analysis in 325 patients from the same health service of pre-testing to identify the dimensions and items that should remain in the instrument. The reliability of the MedRisk was evaluated also in this phase. Another aspect of reliability was measure with intra-class correlation indices and error measurements. The expert panel proposed the exclusion of two items due to the context of instrument application. The pre-testing identified only one item with difficult comprehension and after semantic adjustment, this issue was resolved. The results proposed an instrument with six specific items and two global measures of satisfaction. Three items were retained in Interpersonal dimension and three other items in Physical structure and service support dimension. The reliability showed excellent values for the test/retest (0,96 for dimension 1-Interpersonal and 0,98 for dimension 2-Structure and service support), however the Cronbach's alpha showed moderate values (0,60 for dimension 1 and 0,63 for dimension2). Error measures also showed excellent values for agreement (SEM = 0,13 for dimension 1 and 0,11 for dimension 2). The findings do not support the other version recently proposed to Brazil and also disagrees with the original instrument proposed in the United States. It was identified some limitation in the psychometric analysis history of the instrument which stress the need of further assessments.

Keywords: Quality of care, Patient-centered care, Patient satisfaction; Physiotherapy, Validation studies; Psychometry .

LISTA DE FIGURAS

Figura 1: Composição original do instrumento MedRisk: estrutura dimensional, itens e medidas globais de satisfação do paciente	40
Figura 2: Cronologia dos estudos de validação do MedRisk: instrumento inicial, análises já realizadas e versões propostas	43
Figura 3: Fluxograma do número de casos incluídos nas etapas do estudo	54
Figura 4: Representação esquemática do modelo para o instrumento MedRisk obtido na Análise Fatorial Confirmatória	69
Figura 5: Modelo proposto para o MedRisk - etapas da tomada de decisão no presente estudo.....	78

LISTA DE QUADROS

Quadro 1: Definições de qualidade do cuidado em saúde	20
Quadro 2: Avaliação qualitativa das propriedades de medida dos instrumentos para mensuração da satisfação de pacientes com serviços de fisioterapia	36
Quadro 3: Comparação do MedRisk e demais instrumentos para mensuração da satisfação do paciente na fisioterapia – dimensões e itens dos demais instrumentos presentes ou ausentes no MedRisk	37
Quadro 4: Versão original e síntese das traduções e retro-traduições do MedRisk	57
Quadro 5: Itens da versão utilizada na fase de pré-teste e síntese dos ajustes do painel de especialistas	58
Quadro 6: Resultados do pré-teste e versão do MedRisk proposta após esta fase	62
Quadro 7: Comparação semântica entre as duas versões do instrumento MedRisk para a língua portuguesa.....	75
Quadro 8: Comparação dos aspectos conceituais que moldaram as dimensões apresentadas nos estudos de adaptação transcultural do MedRisk.....	82

LISTA DE TABELAS

Tabela 1: Características clínicas e socioeconômicas da população de estudo	63
Tabela 2: Valores médios dos itens do MedRisk e respectivos coeficientes de correlações (ρ) com as duas medidas globais de satisfação (I17 e I18).....	65
Tabela 3: Análise fatorial exploratória: autovalores e percentual da variância explicada nos modelos com 3 e 2 fatores.....	66
Tabela 4: Análise fatorial exploratória; itens e respectivas cargas fatoriais.....	67
Tabela 5: Modelo proposto do MedRisk: dimensão, valor médio da satisfação e correlação entre os itens	68
Tabela 6: Análise fatorial confirmatória da estrutura dimensional do instrumento MedRisk e resíduos da Análise fatorial exploratória	68
Tabela 7: Comparação dos índices de ajuste do modelo de três estruturas dimensionais propostas para o MedRisk	70
Tabela 8: Valores das medidas de confiabilidade e reprodutibilidade do instrumento MedRisk	70

LISTA DE ABREVIATURAS e SIGLAS

AFC – Análise fatorial confirmatória

AFE – Análise fatorial exploratória

CEP – Comitê de ética em pesquisa

CFI - *Comparative Fit Index*

COSMIN – *Consensus-based Standard for the selection of health status Measurement Instrument*

SURCA/SMSDC – Secretaria municipal de saúde e defesa civil

DP – Desvio padrão

ENSP - Escola Nacional de Saúde Pública Sérgio Arouca

FIOCRUZ – Fundação Oswaldo Cruz

HUGG/UNIRIO – Hospital Universitário Gaffrée e Guinle / Universidade Federal do Estado Rio de Janeiro

IC – Intervalo de confiança

IOM – *Institute of Medicine*

ISO - *International Standardisation Organisation*

ICC – *Intraclass correlation*

MDC - *Minimal Detectable Change*

OMS – Organização Mundial de Saúde

PRO - *patient-related outcome*

SD – *Standard deviation*

SEM - *Standard Error of Measurement*

SUS – Sistema Único de Saúde

RMSEA - *Root Mean Square Error of Approximation*

RN – Rio Grande do Norte

TLI - *Tucker-Lewis Index*

TRI – Teoria de resposta ao item

TTC – Teoria do teste clássico

T1 – Tradução 1

T2 – Tradução 2

TR 1 – Retrotradução 1

TR 2 – Retrotradução 2

WRMR - *Weighted Root Mean Square Residual*

WLSMV - Mínimos quadrados robustos ajustados

MCAR – *Missing completely at random*

MAR – *Missing at random*

MNAR – *Missing not completely at random*

NA – Não se aplica

ND – não demonstrado

SUMÁRIO

1 INTRODUÇÃO.....	13
1.1 Problema e questão de pesquisa	15
1.2 Objetivos.....	18
2 REVISÃO DA LITERATURA	19
2.1 Qualidade do Cuidado	19
2.2 Cuidado centrado no paciente.....	21
2.3 Satisfação do usuário de serviços de saúde	24
2.4 Satisfação dos pacientes com serviços de fisioterapia	27
2.5 Mensuração da satisfação de usuários de serviços de saúde – instrumentos de medida e considerações metodológicas	28
2.5.1 Instrumentos para mensuração da satisfação do paciente: revisão sistemática ...	31
2.5.2 Instrumentos para avaliação da satisfação do paciente na fisioterapia	33
2.5.3 O instrumento MedRisk	38
3 METODOLOGIA.....	44
3.1 Delineamento do estudo	44
3.2 Etapa 1 – Equivalência conceitual, equivalência de itens e equivalência semântica	45
3.2.1 Equivalência Conceitual e de itens	45
3.2.2 Equivalência Semântica	45
3.3 Etapa 2: Equivalência de mensuração: análise psicométrica	47
3.3.1 Validade dimensional e itens componentes	48
3.3.2 Confiabilidade	49
3.3.3 Contexto do estudo, população, coleta dos dados e digitalização das informações	50
3.3.4 Análise de Dados.....	52
3.4 Considerações éticas.....	55
4 RESULTADOS	56
4.1 Etapa 1 - Equivalência conceitual, equivalência de itens e equivalência semântica	56
4.2 Etapa 2 - Equivalência de mensuração – análise psicométrica	63
4.2.1 Análise descritiva e correlações entre os itens	63
4.2.2 Análise fatorial exploratória (AFE) e análise fatorial confirmatória (AFC)	66
4.2.3 Reprodutibilidade – confiabilidade e concordância	70
5 DISCUSSÃO	71

6 CONSIDERAÇÕES FINAIS	86
7 REFERÊNCIAS	88
ANEXO A - Critérios para análise da qualidade metodológica dos instrumentos propostos pelo Grupo COSMIN	98
ANEXO B - Documento de aprovação do comitê de ética da ENSP/Fiocruz	117
ANEXO C - Autorização para uso do instrumento MedRisk	118
APÊNDICE A - Artigo publicado: revisão sistemática sobre instrumentos para mensuração da satisfação do paciente	119
APÊNDICE B - Versão do MedRisk utilizada na fase de pré-teste	141
APÊNDICE C - Perguntas anexadas ao instrumento apenas na fase do pré-teste para análise da compreensão do instrumento	143
APÊNDICE D - Versão MedRisk pré-testada e aplicada para análise psicométrica (N 325)	144
APÊNDICE E - Termo de consentimento livre e esclarecido (TCLE)	146
APÊNDICE F - Análise fatorial exploratória realizadas com o programa SPSS 17.0	148
APÊNDICE G - Análise fatorial confirmatória realizadas com o programa Mplus 5.0	150
APÊNDICE H - Versão do MedRisk proposta pelo presente estudo	162

1 INTRODUÇÃO

A área da saúde apresentou muitos avanços nas últimas décadas, com mudanças tecnológicas e nos processos de gestão. No Brasil, particularmente na esfera pública, constantes adaptações às condições sócio-demográficas da população vêm sendo realizadas, principalmente após a implementação do Sistema Único de Saúde (SUS) no início dos anos 90 (PAIM et al., 2001). O sistema de saúde brasileiro possui grande complexidade e esse fato concorre para a morosidade na implementação de diretrizes nas esferas Municipal, Estadual e Federal. No contexto atual, dentre os principais desafios que deverão ser enfrentados, destacam-se: o acesso à saúde igualitário para as diferentes classes sociais; condições de trabalho justas para profissionais da área; política de medicamentos eficiente; estímulo à produção tecnológica voltada para saúde; e melhora da qualidade do cuidado (VICTORA et al., 2011).

Especialmente a qualidade do cuidado no sistema de saúde brasileiro, comparativamente a outros países, adquiriu relevância mais recentemente, mesmo em termos de publicações (MACHADO et al., 2013; MENDES et al., 2013). O conceito de qualidade pode apresentar diversas definições, de acordo com a área do conhecimento que está inserido. Tal conceito, quando oriundo da área da engenharia, geralmente está ligado à padronização de processos. Entretanto, na área de saúde, assume algumas particularidades e diversas perspectivas (DONABEDIAN 1980, 1988). A atenção aos processos e resultados são abordagens usualmente privilegiadas na avaliação da qualidade do cuidado em saúde (DONABEDIAN 1980; ALMEIDA et al., 2011; RASHIDE e JUSOFF, 2009).

A consolidação da importância dos processos de certificação e acreditação são exemplos de iniciativas voltadas para garantir a qualidade e as intervenções preconizadas buscam alcançar melhorias nos indicadores de saúde e desempenho dos serviços (SHAW et al., 2010). Embora não haja ainda consenso na literatura, a implementação de ambos os processos apresentaria associação positiva com gestão de mudanças e desenvolvimento profissional da equipe (FENTON et al., 2012). Atualmente, estas normas, baseadas em padronizações internacionais para certificação (*International Standardisation Organisation* – ISO) ou nos requerimentos de organizações internacionais de acreditação, buscam além de padronizar processos,

implementar melhorias centradas nas necessidades e preferências dos pacientes, incluindo os valores e preferências do paciente e seus familiares (SHAW et al., 2010).

Do ponto de vista da Organização Mundial de Saúde (OMS), certamente inspirada pelo documento do Instituto de Medicina dos Estados Unidos da América (IOM, 2001), a busca pela qualidade do cuidado prestado pelos serviços de saúde passa por produzir um cuidado que seja efetivo, seguro e centrado no paciente. Considerando a complexidade e âmbito do conceito de qualidade, Berwick (2002) aponta a necessidade de mudanças no sistema de saúde em quatro níveis: (a) na experiência do paciente; (b) no funcionamento das unidades onde acontece o cuidado; (c) no funcionamento das organizações que apoiam as unidades onde acontece a assistência; (d) no ambiente externo das políticas públicas, regulação do sistema e financiamento. Esses níveis podem ser vistos como aninhados, e, se existe qualidade no primeiro nível (experiência dos pacientes), esta poderá refletir nos demais.

Desta forma, considerando o cuidado centrado no paciente, torna-se imprescindível também o foco sobre as experiências dos usuários e de suas famílias. Assim, um conceito que ganhou espaço nas discussões referentes a esta temática é a satisfação do paciente, e, conseqüentemente, o estudo do grau de satisfação se apresenta como uma maneira de mensurar a percepção e experiência do paciente com determinado serviço ou tratamento (SOFAER e FIRMINGER, 2005; VINAGRE e NEVES, 2008).

A satisfação dos pacientes é um importante eixo dentro do campo da qualidade do cuidado, que se debruça sobre a dimensão do cuidado centrado no paciente (BLEICH et al., 2009; IOM, 2001). A satisfação em si é um conceito diferenciado quando comparado às definições de qualidade do cuidado descritas na literatura. Embora haja diversas definições, a satisfação geralmente envolve a avaliação do indivíduo baseada nas expectativas, experiências prévias, e percepções com o serviço utilizado (PARASURAMAN et al., 1988).

Este constructo multidimensional e os fatores associados podem variar de acordo com a natureza do serviço, especialidade clínica e até mesmo condição de saúde dos usuários. Devido a esta complexidade, o desenho dos estudos sobre satisfação do usuário é variado e muitas vezes com resultados paradoxais, não havendo um consenso sobre quais seriam os principais fatores determinantes da satisfação (RASHID E JUSOFF, 2009; SHAW et al., 2010).

Estudos brasileiros sobre a satisfação do usuário de serviços de saúde têm início na década de 1990, mas são ainda escassos quando comparados ao cenário internacional. Essa produção é ainda mais restrita no que concerne ao desenvolvimento ou validação de instrumentos de mensuração específicos para satisfação de serviços de saúde. A aplicação de instrumentos genéricos, oriundos de outras áreas do conhecimento ou até mesmo utilizado em diferentes especialidades clínicas, nem sempre garante que o conteúdo destes esteja teoricamente embasado no conceito de satisfação dos pacientes (ROCHA et al., 2013).

No cenário internacional, há diversos instrumentos para mensuração de satisfação do usuário, disponíveis na literatura (HUSH et al., 2011; ALMEIDA et al., 2015). O conteúdo destes varia, principalmente, com o tipo e complexidade do serviço onde são aplicados, e também com a população alvo. Consequentemente, ainda não há um padrão único em relação aos itens e dimensões constituintes de um instrumento de mensuração deste constructo (BOYER et al., 2009; MANAF et al., 2012; SJETNE et al., 2011; ORDEN et al., 2010). Entretanto, a busca por instrumentos robustos e validados em contextos específicos ajuda a fomentar o processo de adaptação transcultural de instrumentos desenvolvidos em outros países (REICHENHEIM e MORAES, 2007).

1.1 Problema e questão de pesquisa

O recorte desta tese se debruçou sobre os instrumentos de mensuração da satisfação de usuários com serviços de saúde. O delineamento do estudo circunscreveu à abordagem referente aos instrumentos especificamente desenvolvidos para uma especialidade clínica. Considerando a especialidade e interesse do pesquisador principal, os estudos voltados para a área da Fisioterapia foram privilegiados. Portanto, o objeto do estudo foi a validação de instrumento de avaliação da satisfação de pacientes usuários de serviços fisioterapêuticos.

Esta especialidade clínica está ligada aos processos de reabilitação funcional do corpo humano e influencia diretamente a qualidade de vida dos pacientes. Especificamente nos ambulatorios que oferecem atendimentos fisioterapêuticos, além da oferta assistencial, garantir serviços de qualidade também deve ser preocupação dos gestores e profissionais. Outro ponto que corrobora com a decisão de abordar o tema na área da fisioterapia, é o perfil epidemiológico e a carga da doença no país, caracterizado

pela magnitude das doenças crônicas e degenerativas e das causas externas. Tal fenômeno, provavelmente devido ao envelhecimento da população e padrões comportamentais socioculturais da sociedade moderna (SCHRAMM et al., 2004), possivelmente pressionará o aumento da demanda por cuidado e serviços fisioterápicos para responder a essa necessidade de saúde crescente.

Desta forma, novos modelos de atenção passam a ser requeridos. Abordagens para incrementar a qualidade de vida e bem estar, com foco na relação profissional-paciente é uma tendência (IOM, 2001). A proximidade entre esses atores é uma marca da fisioterapia. Apesar da regulamentação profissional ser recente, sua importância no contexto atual do SUS é clara e a oferta desse tipo de cuidado está difundida na rede em todos os níveis assistenciais e com grande demanda nos serviços ambulatoriais.

Com relação aos instrumentos para mensuração, considerando os estudos que privilegiaram a avaliação da satisfação no âmbito da área de fisioterapia (GOLDSTAIN et al., 2000; MONNIN e PERNEGER, 2002; MENDONÇA e GUERRA, 2007; MOREIRA et al., 2007; BEATTIE et al., 2002) o instrumento denominado MedRisk foi o escolhido para adaptação ao contexto brasileiro e mensuração da satisfação de usuários de serviços de fisioterapia. Além de concebido para esta especialidade, este instrumento tem sido utilizado em diversos contextos culturais e suas propriedades psicométricas foram reavaliadas em diferentes momentos (BEATTIE et al., 2002; BEATTIE et al., 2005; BEATTIE et al., 2007; HUSH et al., 2012; HUSH et al., 2013). A justificativa da escolha deste instrumento passa ainda pela maior facilidade de aplicação quando comparado aos outros (ALMEIDA et al., 2013; ALMEIDA et al., 2015; GOLDSTAIN et al., 2000; MONNIN e PERNEGER, 2002; MENDONÇA e GUERRA, 2007; MOREIRA et al., 2007); pela metodologia de desenvolvimento e validação aceitável segundo os critérios incluídos no *checklist* COSMIN - *Consensus-based standards for the selection of health measurement instruments* (TERWEE et al., 2011); e ainda por se propor a mensurar o constructo teórico, objeto deste estudo.

Concomitantemente a este trabalho, tal instrumento foi recentemente adaptado para língua portuguesa e sua equivalência de mensuração foi avaliada em um cenário brasileiro (Oliveira et al, 2014). Contudo, o cenário de aplicação foi circunscrito e distinto do selecionado no presente estudo. Além disso, a diversidade sociocultural brasileira e as particularidades do sistema de saúde aportam senões para conformar um único e validado instrumento para amplo uso nesse contexto. Por outro lado, a utilização

de instrumentos de mensuração com pouca ou nenhuma avaliação de suas propriedades de mensuração, ou revisitações metodológicas, é uma prática usual na fisioterapia clínica que representa uma limitação importante dos estudos realizados. Nesse sentido, os resultados do estudo de Oliveira et al. (2014) permitem a comparação não somente dos testes psicométricos, mas um avanço; pois oferece uma oportunidade de buscar maior sintonia do MedRisk com o contexto brasileiro. Esses elementos retratam a relevância e contribuição do presente estudo.

Assim, com a perspectiva de avaliar a satisfação dos pacientes e visando consolidar o processo de validação do instrumento, a presente tese revisou de forma sistemática a literatura e apresenta um artigo de revisão publicado (Apêndice A) referente aos instrumentos validados de mensuração de satisfação dos usuários de serviços de saúde que ajudou a selecionar o MedRisk como objeto de estudo da adaptação transcultural para o contexto brasileiro. O presente estudo é apresentado em quatro seções, seguidas dessa introdução, onde foram primeiramente apresentados os marcos contextuais e os principais temas e conceitos correlatos; as definições do constructo satisfação do paciente, as questões teórico-metodológicas referentes aos instrumentos para mensuração da satisfação de pacientes com serviços de saúde e o desenvolvimento e validação do instrumento MedRisk. A segunda seção se detém sobre a abordagem metodológica utilizada e descreve o processo de adaptação transcultural. A terceira apresenta os resultados e a quarta discute e compara os achados do presente estudo. Por fim, são expostas as considerações finais, limites, críticas e recomendações para futuros estudos e utilização do MedRisk em iniciativas de melhoria da qualidade do cuidado na fisioterapia.

1.2 OBJETIVOS

Objetivo Geral

Conduzir o processo de adaptação transcultural do instrumento MedRisk para mensuração da satisfação do usuário de serviços de saúde em contexto brasileiro.

Objetivos específicos

- Realizar uma revisão sistemática da literatura científica sobre instrumentos utilizados e validados para mensuração da satisfação de pacientes com serviços de saúde.
- Analisar a equivalência conceitual, de itens, operacional e semântica do instrumento MedRisk.
- Avaliar a equivalência de mensuração instrumento quanto à validade de constructo e confiabilidade.

2 REVISÃO DA LITERATURA

2.1 Qualidade do cuidado

A área da saúde possui diversas peculiaridades, entretanto, assim como outras áreas do conhecimento, a busca por qualidade e consequentemente sua conceituação, cada vez mais toma espaço nas pautas de discussão. Especificamente, a qualidade do cuidado reporta-se tanto ao âmbito gerencial dos serviços quanto à linha de frente do atendimento, envolvendo profissionais da saúde e demais colaboradores envolvidos na assistência (ROZENBLUM, 2011).

Historicamente, no panorama internacional podem-se encontrar relatos na literatura sobre iniciativas mais sistemáticas e contínuas para melhoria da qualidade do cuidado em diversos países: no Reino Unido, a partir de 1950, na Austrália, a partir de 1960 e nos Estados Unidos, em 1970 com a criação do *Institute of Medicine*. Contudo, o pioneirismo de Florence Nightingale na Inglaterra nos anos 1860 e de Ernest Codman nos Estados Unidos nos anos 1910, certamente serviram de inspiração a essas iniciativas mais contemporâneas (FELDMAN et al., 2005; MENDES et al., 2013).

No contexto brasileiro, entretanto, preocupações com a qualidade são relativamente mais recentes (NORONHA e ROSA, 1999; MENDES et al., 2013). Embora tenha havido alguns esforços pontuais na década de 50, como o Colégio de Cirurgiões estabelecendo o padrão mínimo para planta física de centros cirúrgicos; ou o Ministério do Trabalho e Previdência Social em 1974 aprovando proposições de “Relatório de Classificação Hospitalar” e de “Comprovante de Classificação Hospitalar”, abordando processos e aspectos físicos; apenas na década de 1980, surgem programas para acreditação de hospitais. Somente na década de 1990, surge um programa para melhoria da qualidade do cuidado formulado pelo Ministério da Saúde (NORONHA, 1999; MENDES et al., 2013).

O conceito de qualidade do cuidado é representado por uma série de definições (Quadro 1). Um dos primeiros autores a formular teoricamente este campo foi Donabedian (1980), que, em sua obra, cunhou vários conceitos centrais (DONABEDIAN, 1988). Nos primórdios, esse autor definiu qualidade do cuidado como: “cuidado no qual se espera haver uma maximização nos resultados de mensuração do bem-estar do paciente, após o balanço entre ganhos e perdas inerente ao

processo de cuidado de todos os pacientes” (DONABEDIAN, 1980; QUIGLEY et al., 2008).

Quadro 1: Definições de qualidade do cuidado em saúde

Autor / Organização	Definição
<i>Donabedian (1980)</i>	É o cuidado no qual se espera haver uma maximização nos resultados de mensuração do bem-estar do paciente, após o balanço entre ganhos e perdas inerente a processo de cuidado de todos os pacientes.
<i>“Institute of medicine” (IOM,1990)</i>	É o grau de cuidado em que serviços de saúde para indivíduos e populações aumentam a probabilidade de ocorrência de resultados desejados e consistentes com o conhecimento profissional
<i>Departamento de Saúde, Reino Unido (1997)</i>	Qualidade no cuidado é: realizar a coisa de maneira correta; para as pessoas certas; no tempo adequado; e realizando as ações adequadas de primeira.
<i>Conselho Europeu (1998)</i>	Qualidade no cuidado é o grau em que o tratamento dispensado aumenta as chances do paciente em alcançar os resultados desejados e diminui a probabilidade dos indesejados, levando em consideração os estágios atuais do conhecimento.
<i>Organização Mundial de Saúde (OMS, 2000)</i>	Qualidade do cuidado é o nível de realização dos objetivos intrínsecos do sistema de saúde em melhorar a saúde e os níveis de pronta resposta para legitimar as expectativas dos pacientes

Fonte: Quigley et al (2008).

Com base no conhecimento acumulado ao longo de décadas, a publicação do relatório *Crossing the Quality Chasm* pelo *Institute of Medicine* dos Estados Unidos, em 2001, representa um marco para estudos e intervenções na área da qualidade do cuidado. Neste relatório, a qualidade do cuidado é compreendida como um conceito multidimensional, composto por seis dimensões: efetividade, eficiência, oportunidade, segurança, centralidade do cuidado no paciente e equidade (BERWICK, 2002; IOM,

2001). Mais que dimensões da qualidade, essas seis características devem estar presentes para garantia da qualidade do cuidado nos serviços, identificando como desafios: evitar que o tratamento proposto cause qualquer tipo de dano ao paciente (mais segurança); oferecer cuidado baseado no conhecimento científico para resolver o problema de saúde do paciente (ações *efetivas*); considerar em todas as etapas da prestação do cuidado as preferências, necessidades e valores do paciente (*centralidade do paciente*), reduzir esperas e atrasos no início e continuidade do tratamento (*oportunidade temporal*), evitar desperdício de insumos e equipamentos (*eficiência*), reduzir variação no cuidado devido a características dos usuários: gênero, etnia ou estado socioeconômico (respeito a *equidade*) (IOM, 2001).

Para garantir a operacionalização e implementação destas recomendações, é importante primeiramente mensurar o grau de qualidade dos serviços de saúde e do cuidado prestado. Um dos modelos para avaliação de qualidade foi proposto por Donabedian (1980), no qual a mensuração de elementos chaves do tripé *estrutura-processo-resultado*, inerentes à produção dos serviços, fornece elementos para esse julgamento. Nesse modelo, a estrutura abrange aspectos físicos do serviço como ambiente, insumos, recursos humanos, materiais de consumo e equipamentos. O processo diz respeito às ações realizadas para cuidar do paciente e pressupõe uma estrutura adequada para que os recursos necessários sejam empregados. Os resultados são os efeitos, desejáveis ou adversos, sobre o estado de saúde do indivíduo ou população que podem ser atribuídos ao cuidado prestado (DONABEDIAN, 1988).

No nível do microssistema de saúde (unidade onde realmente acontece o cuidado e a inter-relação entre o paciente e os profissionais envolvidos neste processo), a qualidade do cuidado é mais perceptível, portanto problemas relacionados às seis dimensões da qualidade podem ser mais perceptíveis e mensuráveis. Nesse âmbito, o redesenho do microssistema deveria se apoiar em três princípios norteadores: cuidado baseado em evidência; cuidado centrado no paciente; cuidado integrado ao sistema (BERWICK, 2002; MENDES et al., 2013).

2.2 Cuidado centrado no paciente

Dentre os eixos da qualidade do cuidado definidos pelo IOM (2001), o referente à “centralidade do cuidado no paciente” está, também, relacionado à satisfação do usuário do serviço de saúde. O conceito da centralidade do cuidado no indivíduo refere-

se ao cuidado respeitoso e responsivo às preferências, necessidades e valores individuais dos pacientes, e assegura que seus valores orientem todas as decisões clínicas (TRAVASSOS e CALDAS, 2013). Dessa maneira, essa dimensão evidencia a relevância de entender as necessidades dos usuários, incluindo as informações e os meios para comunicar as etapas de tratamento, suas expectativas com relação aos resultados, empatia com os profissionais e participação nas decisões.

O conceito de cuidado centrado no paciente foi introduzido por volta de 1969 como “uma nova maneira de pensar” a prestação de serviços de saúde (GROENE, 2011). Esta nova perspectiva propunha que os profissionais de saúde tivessem uma visão holística do ser humano e levassem em consideração tudo que sabiam sobre seus pacientes antes de fechar o diagnóstico clínico. Esta abordagem permitiria ao clínico identificar o perfil do paciente - aquele mais ativo e participativo das decisões clínicas e o mais passivo, que questiona menos o profissional quanto à sua condição de saúde. Entretanto, um elemento paradoxal permeia tal abordagem, pois pacientes com maior nível sociocultural, que usualmente participam das decisões clínicas, necessitam de menos explicações referentes à sua condição de saúde. Por sua vez, aqueles com menor nível sociocultural, que necessitam de educação em saúde, muitas vezes não participam das decisões e apresentam dificuldades para adesão ao tratamento (HOLMSTROM e ROING, 2010; GROENE, 2011).

Quando o profissional de saúde enxerga determinada enfermidade pela perspectiva do paciente e considera suas crenças, a adesão e a efetividade do tratamento podem ser maximizadas. Esta abordagem influenciaria ainda discussões e decisões institucionais voltadas para a garantia dos direitos dos pacientes, consequentemente aportando insumos para a aprendizagem organizacional (GROENE, 2011)

Esse paradigma ajudou a estruturar o modelo biopsicossocial na medicina, que leva em consideração não só variáveis biomédicas, como também características do paciente e do contexto social a sua volta. Bardes (2012) sublinha que o conceito de cuidado centrado no paciente representa uma metáfora, pois o cuidado não pode ser centrado totalmente nos desejos e perspectivas dos usuário dos serviços, visto que o profissional de saúde exerce um papel de gerência clínica da condição de saúde. Desta forma, estes atores devem coexistir, com igual participação nas decisões.

Relatos da literatura evidenciam que o cuidado centrado no paciente, moldado pela decisão partilhada, está associado também a satisfação do paciente, melhor

prognóstico e resultado clínico, assim como redução de reincidivas, dos níveis de dor e estresse (GROENE, 2011; ROZENZVEIG, 2014).

A lacuna existente entre a área gerencial e clínica, na maioria dos serviços de saúde, é uma grande barreira para a implementação de melhorias nesses processos. Desta forma, a mudança da cultura organizacional e o diálogo entre os gestores e a equipe clínica são algumas abordagens que podem facilitar a implementação de programas de melhoria de qualidade (ROZENBLUN et al., 2011). Rozenblun et al. (2013) relataram ainda que equipes clínicas que obtiveram algum tipo de *feedback*, após aplicação em seus departamentos de questionários voltados para um dos aspectos do cuidado centrado no paciente, apresentavam maior interesse em relação às percepções do paciente no decorrer do tratamento e ainda maior aderência às ações propostas para melhoria de qualidade.

A mensuração de problemas na qualidade associados à centralidade do cuidado no paciente traz especificidades dadas por este constructo. Avaliar o grau de centralidade do cuidado no paciente requer a apreensão da percepção deste indivíduo acerca da organização, sua estrutura, dos processos e ainda dos resultados clínicos. Elementos como habilidade de comunicação dos profissionais, satisfação dos pacientes e medidas clínicas do resultado são usualmente empregados nessas avaliações. Entretanto, a preferência dos pacientes quanto ao grau de envolvimento na decisão sobre tratamento e o nível educacional são menos utilizados (GROENE, 2011; EPSTEIN e STREET, 2011; ROZENZVEIG, et al., 2014).

Portanto, o eixo “cuidado centrado no paciente” deve se apoiar em diversas óticas e medidas. Recente revisão sistemática (DOYLE et al., 2013) definiu alguns fatores relevantes para a chamada “experiência do paciente”, incluindo apoio emocional, participação nas decisões clínicas, clareza e transparência nas informações. Essa revisão destacou ainda que a adequação do cuidado às experiências dos indivíduos está associada à efetividade clínica e segurança do paciente.

O cuidado centrado no paciente aborda também aspectos clínicos. Rozenzveig et al. (2014), realizaram uma revisão sistemática da literatura e esta apontou que a aplicação de questionários para identificação da auto percepção do paciente sobre o nível de dor, fadiga, estresse, depressão, ansiedade, sono e auto avaliação global do estado de saúde vêm sendo utilizadas como medidas do resultado do tratamento. Estas abordagens refletem a preocupação dos clínicos quanto à percepção dos pacientes frente

aos resultados dos seus tratamentos e a tentativa de compartilhar com os próprios pacientes o gerenciamento clínico das patologias (ROZENZVEIG et al., 2014; OLIVEIRA et al., 2014).

Hundon et al. (2011), baseados na literatura, propuseram quatro dimensões fundamentais para o cuidado centrado no paciente: o doente como pessoa e sua experiência com a patologia; abordagem do indivíduo de forma holística (modelo biopsicossocial); compartilhamento das responsabilidades entre paciente-profissional de saúde; relação terapeuta-paciente. Assim, o reconhecimento dessas dimensões, agregando elementos externos ao modelo biomédico, vislumbra a satisfação do paciente como uma importante ferramenta para avaliar o cuidado centrado no paciente (ROZENZVEIG et al., 2014).

2.3 Satisfação do usuário de serviços de saúde

As pesquisas sobre satisfação do paciente acontecem de longa data. Orientadas pela sociologia, relatos sobre a percepção de tais pacientes referentes à relação médico-paciente já aconteciam antes mesmo da discussão em torno do conceito mais amplo de cuidado centrado no usuário. Desde a década de 50, alguns estudos discutiam aspectos da relação entre os pacientes e seus médicos, e, gradativamente, estes achados passaram a ser disseminados para ajudar diretamente os profissionais envolvidos com assistência em saúde (WILLIAMS, 1994). Nas décadas seguintes, 60 e 70, este movimento continuou a crescer e contribuiu para a introdução do conceito de satisfação do paciente nas preocupações e publicações da área de saúde.

Esse conceito foi gradativamente abarcado pelo conceito de “cuidado centrado no paciente”. Essa revisitação contemporânea do conceito renovou a relevância dos estudos de satisfação do usuário nos serviços de saúde (PASCOE, 1983; SITZIA, 1997; IOM, 2001). Este movimento foi também fundamental para inserir a satisfação, ancorada pela perspectiva e preferências dos pacientes, como um dos tópicos extremamente valorizados para garantia da qualidade do cuidado (WILLIAMS, 1994; PASCOE, 1983).

Ware et al. (1977) definiram a satisfação do paciente como um construto composto pelas seguintes dimensões: a arte do cuidado, a qualidade técnica, a acessibilidade/conveniência, a financeira/custo, o ambiente físico, a prontidão no

atendimento, a continuidade do cuidado e os resultados. Os autores, já àquela época, estavam preocupados ainda com a mensurabilidade de cada dimensão da satisfação do usuário e com métodos adequados e práticos para sua mensuração.

Ao longo do tempo, o conceito de satisfação do paciente foi discutido com base em diversos modelos teóricos explicativos. Em 1983, Pascoe descreveu um leque de quadros conceituais, abarcando desde modelos debruçados sobre aspectos cognitivos dos indivíduos até aqueles que destacavam aspectos socioculturais relacionados a este constructo. Um dos modelos mais utilizados é o de *Linder-Pelz (1982)*, neste, a satisfação foi definida como “uma atitude positiva”, isto é, levava em consideração que o comportamento do indivíduo frente ao serviço é baseado nos seus valores e julgamento do cuidado recebido. Acorados neste conceito, quatro dimensões associadas à satisfação do paciente foram postuladas: acesso, eficácia, resultado no tratamento, custo e conveniência (PASCOE, 1983).

Outros quadros conceituais também foram propostos. Um dos mais empregados foi o modelo calcado no balanço entre o que o paciente “recebeu” (percepção do serviço) e o que ele “esperava” (expectativa) – *Fulfilment Theory* (PASCOE, 1983). A aplicação desta base teórica – lacuna entre expectativa e percepções, resultou inclusive no desenvolvimento do instrumento SERVQUAL. Este foi um dos primeiros instrumentos genéricos, aplicável a diversas áreas e setores, utilizado para mensuração de satisfação. Sua configuração aponta que a satisfação é composta por cinco dimensões: tangibilidade, confiabilidade, prontidão, segurança e empatia (PARASURAMAN et al., 1988).

Em contraposição à escolha de um modelo explicativo, Fitzpatrick (1984) postulou que a satisfação não poderia ser vista como um conceito único relacionado a diversos determinantes. Esse autor defendia que três modelos deveriam ser considerados e cada grupo de determinantes estaria ligado à (1) necessidade de familiaridade – os indivíduos se comportam como grupos homogêneos, dependendo da cultura onde estão inseridos; ao (2) resultado - a satisfação estaria relacionada aos resultados finais do tratamento; e à (3) importância emocional – dado que os pacientes não teriam como julgar aspectos técnicos, assim dependeria da relação com o profissional de saúde envolvido (SITZIA e WOOD, 1997; FITZPATRICK, 1984).

Paralelamente à discussão teórico-conceitual na década de 1980, a percepção e expectativa do indivíduo (satisfação) foram consideradas para avaliar a qualidade do

cuidado oferecido, sobretudo nos Estados Unidos. Já a esta época, a operacionalização dos estudos de satisfação era realizada com base em questionários, pois estes eram fáceis de aplicar, baratos e quantificavam percepção e satisfação (WILLIAMS, 1994; SITZIA e WOOD, 1997).

Contudo, ao longo dos anos, nenhum destes modelos conseguiu se sustentar absolutamente, havendo, portanto, uma automática remodelagem neste conceito. Embora ainda existam discussões quanto à definição teórica-conceitual que embasa tais modelos, em geral a literatura atual se detém sobre abordagens empíricas dos fatores associados à satisfação do usuário (FENTON et al., 2012). Assim, há evidência que as características individuais dos pacientes (idade, estado de saúde e níveis de escolaridade) são relevantes na explicação do grau de satisfação e, muitas vezes, tais características são mais importantes que alguns aspectos intrínsecos do serviço de saúde (tipo de hospital, tamanho do hospital e cultura organizacional) (HEKKERT et al., 2009; RAHMQVIST e BARA, 2010).

Os níveis de satisfação e insatisfação podem ser ainda vistos como tipos de reações humanas, que surgem de um processo de influências culturais e experiências prévias pessoais. A percepção dos serviços varia de acordo com os valores culturais de cada país ou grupo de países (ELLEUCH, 2007). Desta forma, a satisfação é descrita também como um processo dinâmico, pois as percepções dos indivíduos, frente a um mesmo serviço ou tratamento, podem variar de acordo com o momento pessoal de cada indivíduo, isto é, as expectativas podem mudar em função de aspectos intrínsecos (PASCOE, 1983; HILLS e KITCHEN, 2007; RASHID e JUSOFF, 2009). Outro tópico associado à satisfação é a condição clínica do paciente. Hills e Kitchen (2007) observaram, num estudo utilizando “grupos focais” com usuários de serviços de saúde, que indivíduos com enfermidades agudas (recentes) são mais otimistas com o tratamento e, portanto, tendem a demonstrar maiores níveis de satisfação. Por outro lado, indivíduos com disfunções crônicas tendem a apresentar um comportamento geralmente mais desconfiado e receoso quanto ao sucesso do tratamento, fato que pode impactar negativamente na satisfação.

Bleich et al. (2009) realizaram um estudo multicêntrico em 21 países da União Europeia e identificaram, dentre os fatores preditivos da satisfação de usuários de serviços ambulatoriais de saúde, a experiência prévia dos indivíduos com o serviço como o fator mais importante. Já outros fatores como a expectativa do paciente, o

estado de saúde e o tipo de cuidado quando acrescentados agregaram juntos 7,1% à capacidade explicativa do modelo de regressão testado para explicar presença ou não de satisfação. Entretanto, o resultado do cuidado também influi na Satisfação do paciente. Schoenfelder et al. (2011) estudando a satisfação entre 8.428 indivíduos em 39 hospitais na Alemanha, identificaram a variável “resultado do tratamento” como o principal fator preditivo da satisfação, seguido pelo “comportamento dos profissionais”. O achado do estudo alemão (SCHOENFELDER et al., 2011) exemplifica a diferença de fatores preditivos da satisfação entre pacientes hospitalizados e usuários de serviços ambulatoriais (BLEICH et al. 2009).

2.4 Satisfação dos pacientes com serviços de fisioterapia

A importância relativa dos fatores associados à satisfação do paciente pode se distinguir entre diferentes especialidades médicas. Na área da fisioterapia, uma revisão sistemática realizada por Hush et al. (2011), relatou que a boa relação entre o fisioterapeuta e o paciente é o principal fator para a satisfação. Contudo, assim como em outras especialidades da saúde, as discussões referentes a este constructo na área da fisioterapia corroboram os modelos de estruturas multidimensionais (CURRY e SINCLAIR, 2002; BEATTIE et al., 2002; MOREIRA et al., 2007; MENDONÇA e GUERRA, 2007).

Estudos empíricos, voltados à mensuração da satisfação dos pacientes, apontam ainda a relevância da avaliação de fatores como a atenção do terapeuta dedicada ao paciente (tempo de consulta e acessibilidade para diálogo), a qualidade de comunicação do terapeuta (explicações referentes ao prognóstico e diretrizes para tratamento), e por último os resultados esperados. Outras variáveis relacionadas ao processo de atendimento (tempo de espera, acompanhamento do paciente, conveniência); e expectativas prévias também foram apontadas como variáveis importantes. Tais achados são consistentes entre estudos realizados em diferentes regiões geográficas e culturas no mundo (CURRY e SINCLAIR, 2002; GOLDSTEIN, 2000; HILLS e KITCHEN, 2007; HUSH et al., 2011).

Um estudo internacional, referente à satisfação de pacientes com serviços de fisioterapia, mostrou que os níveis de satisfação encontrados em indivíduos com disfunções musculoesqueléticas são usualmente altos. Em uma escala de 5 pontos, onde

5 representa satisfação com o serviço, os valores médios encontrados na Austrália giraram em torno de 4,6; enquanto no Canadá foi 4,7, nos Estados Unidos 4,5, na Suécia 4,4, na Irlanda 4,3 e na Inglaterra 4,7. Embora os instrumentos utilizados nesses estudos sejam distintos, dificultando a comparação, todos os valores apontaram uma tendência à satisfação com os serviços (HUSH et al., 2012).

No cenário brasileiro, Machado e Nogueira (2008) estudaram a satisfação com o atendimento em Teresina (PI). Esses autores observaram que na rede privada sem vínculo com o SUS, 53,6% dos usuários entrevistados relataram estar satisfeitos com o atendimento recebido. Já nas redes municipal e estadual esse percentual foi ainda maior, respectivamente de 73,2% e 82,3%. Curiosamente, apesar de toda dificuldade encontrada por muitos pacientes para conseguir uma consulta e garantir a continuidade de tratamento, a grande maioria dos estudos demonstram haver satisfação por parte dos pacientes tratados na rede SUS (HERCOS e BEREZOVSKY et al., 2006; MOREIRA et al., 2007; MENDONÇA e GUERRA, 2007, BORGES et al., 2010).

O fenômeno de superestimação da satisfação pode ser um problema que enviesará esse tipo de estudo. Baker et al. (2009) relatam que indivíduos que possuem alto grau de envolvimento com o serviço ou seu terapeuta, tendem a atribuir valores elevados quando julgam qualidade do cuidado. Tal fato contribui com possível viés também na área da fisioterapia, talvez mesmo em maior grau, dada a natureza do cuidado fisioterápico, que envolve o toque no paciente e intervenção em aspectos psicossociais relacionados ao quadro clínico.

O caráter complexo confere a esta temática um status de constante busca por fatores, globais e específicos, determinantes da satisfação, além de instrumentos com robustez metodológica capazes de mensurar tal complexidade. Desta maneira, a mensuração do grau de satisfação e suas principais dimensões devem perpassar antes pela utilização de instrumentos de aferição confiáveis e válidos (ALMEIDA et al., 2015).

2.5 Mensuração da satisfação de usuários de serviços de saúde – instrumentos de medida e considerações metodológicas

Paralelamente às discussões conceituais sobre “cuidado centrado no paciente” e “satisfação do usuário”, há preocupações referentes aos métodos de mensuração destes constructos. O desenvolvimento de novas ferramentas ou adaptação de instrumentos validados para o contexto onde serão utilizados é uma decisão fundamental; que deve basear-se na literatura referente ao constructo em questão, discussão entre especialistas e análise psicométrica. Após concepção ou adaptação de um instrumento, o trabalho de campo para o teste deve ser conduzido em amostras representativas da população alvo, para que a equivalência de mensuração do instrumento possa ser realizada. O processo de adaptação de ferramentas, já validadas também, deve passar por etapas específicas, imprescindíveis para calibração e utilização de tais ferramentas em larga escala (HERDMAN et al., 1997; HERDMAN et al., 1998; REICHENHEIM e MORAES, 2007; BRÉDART et al., 2014).

Uma importante iniciativa, que padronizou os principais passos necessários para desenvolvimento e adaptação de instrumentos, é a lista de verificação [*checklist*] do grupo COSMIN : *Consensus-based standards for the selection of health measurement instruments* (<http://www.cosmin.nl>). Para validação de instrumentos esta lista propõe a avaliação de cinco atributos. O primeiro atributo é a (1) consistência interna – inter-relação entre os itens, isto é, determinar se grupos de itens são parte de um modelo que reflete o constructo que se pretende medir e se todos os itens formam um mesmo constructo. O segundo atributo é a (2) validade de conteúdo – o quanto um instrumento realmente reflete o constructo a ser medido, considerando a relevância dos itens para população, para o constructo e para o objetivo da medida e ainda a compreensão dos itens pela população. O terceiro atributo é a (3) validade de constructo – o quanto os resultados encontrados realmente representam o constructo a ser medido; a validade de constructo deve considerar três aspectos: *teste de hipótese* (magnitude e direção das correlações esperadas entre o instrumento e outra ferramenta comparável que mensura o mesmo constructo); *relações internas do instrumento* (relacionadas com a validade estrutural) e *validade transcultural*. O quarto atributo é a (4) validade de critério – o quanto os valores obtidos com o instrumento se assemelham ao resultado de um instrumento de referência (“*gold standard*”). Por fim o quinto atributo é a (5)

capacidade de resposta – habilidade do instrumento para detectar uma mudança do constructo mensurado dentro de um intervalo de tempo. Esta abordagem deve ainda incluir medidas de confiabilidade (teste-reteste) e de erro (MOKKINK et al., 2010; TERWEE et al., 2011).

Os instrumentos de medida que se baseiam nas perspectivas dos pacientes para avaliar o resultado do tratamento, sintomas, grau de incapacidade, qualidade de vida associada à condição de saúde e satisfação, são denominados “*patient-related outcome (PRO)*”. Estes trazem informações das condições de saúde ou variáveis a elas associadas pela ótica do paciente, sem interferência direta do profissional de saúde responsável pelo caso (BRÉDART et al., 2014).

Neste contexto, a satisfação do paciente, vista não apenas como uma medida gerencial que interfere na qualidade do cuidado, mas também como uma medida que pode interferir sobre as condições clínicas, vem cada vez mais sendo estudada por instrumentos *PRO* (HUDAK e WRIGHT, 2000). Entretanto, a escolha da ferramenta adequada é difícil, pois depende da proposta principal de uso - avaliar um serviço de saúde ou tratamento específico. A escolha do instrumento errado pode impactar na acurácia das informações, assim como a utilização de um instrumento, cuja propriedade psicométrica tenha sido parcialmente avaliada, pode enviesar os resultados (MOKKINK et al., 2010; TERWEE et al., 2011).

As medidas de satisfação do usuário podem ser classificadas segundo dois eixos: o conteúdo e o método. O conteúdo diz respeito ao foco principal da medida: instrumento global ou multidimensional; genérico ou específico para uma condição de saúde; medida de resultado de tratamento ou do processo de cuidado em saúde. Já o método refere-se à maneira como a medida é realizada: perguntas abertas ou fechadas, auto administradas ou entrevista, tipo do formato da resposta. Deve-se ressaltar que é comum encontrar na literatura instrumentos que mesclam elementos destes dois eixos em sua composição, principalmente quanto aos elementos do conteúdo (HUDAK e WRIGHT, 2000). Quanto ao conteúdo dos instrumentos, deve-se apontar ainda que a satisfação relacionada ao resultado do tratamento é mensurada por instrumentos distintos daqueles desenvolvidos para medir satisfação com o processo de cuidado. Processo e resultado são fenômenos distintos, além disso, há evidências que a correlação entre satisfação e resultado final do tratamento é fraca (KAMPER et al., 2010; HUDAK et al., 1999).

O uso de instrumentos multidimensionais para análise da satisfação do usuário é amplamente encontrado na literatura. Estes parecem mais indicados para avaliar a satisfação com um determinado serviço de saúde (ALMEIDA et al., 2015). Hall e Dornan (1988), realizaram uma meta-análise e encontraram que 25% dos estudos analisados media apenas uma dimensão, enquanto que a maioria (75%) media duas ou mais dimensões. Mais recentemente, outras revisões sistemáticas confirmaram tais achados (BOYER et al., 2009; HUSH et al., 2011; ALMEIDA et al., 2015).

Além do conteúdo e apresentação dos itens, o formato das opções de respostas também é de grande relevância para a precisão do instrumento em captar aquilo que realmente se pretende medir. As escalas com pontuação crescentes (*likert*) são amplamente utilizadas nos questionários dirigidos aos pacientes (MEYER, 2007). Entretanto, Khadka et al. (2012) relataram que a maneira como a pergunta é formulada influencia mais a calibração do item do que o formato de resposta. Willis e Artino (2013) relatam que a “entrevista cognitiva” pode ser um método mais preciso para abordar questões de um instrumento novo ou em desenvolvimento. A entrevista cognitiva é uma técnica oriunda da psicologia que estuda a maneira com que os indivíduos processam e respondem mentalmente aos questionários. É um método qualitativo desenvolvido para investigar se questionários desenvolvidos para identificação de perfis comportamentais, atitudes ou de tomada de decisão, realmente medem o que se propõem a medir.

Existem diversos instrumentos para mensuração da satisfação de pacientes com serviços de saúde. A mensuração deste constructo pode ser realizada de diversas maneiras, entretanto a aplicação de instrumentos estruturados e validados vem sendo a abordagem mais objetiva e recomendada (HUDAK e WRIGHT, 2000; BOYER et al., 2009; HUSH et al., 2011; ALMEIDA et al., 2015). Nessa perspectiva, o estudo do processo de validação de tais instrumentos é de extrema relevância, visando conhecer e minimizar viés de aferição e melhorar a consistência dos achados obtidos (MOKKINK et al., 2010; ALMEIDA et al., 2015).

2.5.1 Instrumentos para mensuração da satisfação do paciente: revisão sistemática

No âmbito deste estudo, foi realizada uma revisão sistemática da literatura para identificar os principais instrumentos validados para mensuração da satisfação dos

pacientes¹. Os estudos incluídos nesta revisão indicam que a aplicação de instrumentos para mensuração de satisfação dos pacientes com serviços de saúde constitui uma prática em todo o mundo (Europa, Ásia, América do Norte, América do Sul, Austrália e África).

Nesta revisão, foram encontrados 34 instrumentos diferentes (ALMEIDA et al., 2015). Tal fato chamou atenção para a inexistência de um instrumento de referência (“*gold standard*”) para mensuração de satisfação do usuário de serviços de saúde. Tanto em serviços de maior complexidade, como hospitais, quanto em ambulatorios ou clínicas, o uso de diferentes instrumentos transparece na literatura (OTONELLO et al., 2012; ARRARAS et al., 2009; GERKENSMEYER e AUSTIN, 2005). Ou seja, mesmo para um determinado nível assistencial há variedade no tipo de instrumento aplicado.

Embora, também, não haja um consenso na literatura quanto aos fatores constituintes da satisfação dos pacientes, a grande maioria dos estudos relatou que o constructo “satisfação do usuário de serviços de saúde” é multidimensional. Apenas três estudos demonstraram um modelo com estrutura unidimensional após avaliação do instrumento de mensuração da satisfação (GERKENSMEYER e AUSTIN, 2005; WILDE e HENDRIKS, 2000; MULTINOVIC´ et al., 2012).

Nos estudos revistos (ALMEIDA et al., 2015), as dimensões relativas à relação profissional-paciente, estrutura física e processos internos administrativos estavam presentes na estrutura dimensional tanto dos novos instrumentos como daqueles adaptados (GERKENSMEYER e AUSTIN, 2005; EVANS et al., 2012; WILDE e HENDRIKS, 2005).

Modalidade de cuidado ou serviços de saúde diferenciados, como cuidado domiciliar, *out-of-hours*, assistência psiquiátrica, pediátrica, unidades com maior complexidade e abordagem multidisciplinar, geralmente incorporam outras dimensões para adequá-los a estas especificidades (EVANS et al., 2012; SHEA et al., 2008; DANIELSEN et al., 2010; POINSOT et al., 2006; ALMEIDA et al., 2015).

Com relação aos aspectos metodológicos, a maioria dos estudos se baseou na teoria do teste clássico (TTC) para investigar o comportamento psicométrico dos instrumentos. As propriedades psicométricas mais comumente descritas foram a validade estrutural (dimensional) e consistência interna. A validade de conteúdo (validade de face), teste de hipóteses (comparação com outros instrumentos de

¹ O artigo completo produto desta revisão sistemática, publicado nos Cadernos de Saúde Pública em janeiro de 2015, é apresentado no Apêndice A.

referência), confiabilidade (teste-reteste) e validação transcultural foram menos frequentemente examinados.

Todos os estudos incluídos utilizaram o alpha de Cronbach como uma medida de consistência interna. A grande maioria deles apresentou a correlação entre os itens, sub-escalas e com outros instrumentos para mensuração da confiabilidade. Também na grande maioria dos estudos, a validade do instrumento foi testada com base na análise fatorial exploratória (EFA) e análise fatorial confirmatória (CFA). A aplicação dos critérios incluídos no *checklist* COSMIN demonstrou que a maioria dos instrumentos apresentou avaliação mediana (“*Good*” ou “*Fair*”) quanto à qualidade metodológica das etapas realizadas para validação dos instrumentos (MONKKINK et al., 2010 e TERWEE et al., 2011).

2.5.2 Instrumentos para avaliação da satisfação do paciente na fisioterapia

A revisão sistemática realizada (ALMEIDA et al, 2015) incluiu apenas um instrumento da área da fisioterapia (Vanti et al. 2013), provavelmente explicado pela escolha das palavras-chave e também recorte do estudo: estudos publicados entre 2002 e 2013 que apresentassem alguma propriedade psicométrica do instrumento descrita. A análise qualitativa do instrumento adaptado por Vanti et al (2013), apontou alguma dificuldade operacional para sua aplicação: demasiadamente extenso (26 itens inicialmente) e avaliação metodológica regular segundo os critérios do *checklist* COSMIN (TERWEE et al., 2011).

Assim, foi realizado um levantamento por busca livre para identificar instrumentos na área de fisioterapia. Esta busca revelou seis instrumentos para mensuração de satisfação do paciente; sendo dois em uso no cenário nacional e quatro fora do país. Portanto, o desenvolvimento de um novo instrumento não parece necessário na área.

Dentre os instrumentos aplicados no Brasil, destacam-se pela qualidade metodológica os estudos de Moreira et al. (2007) e Mendonça e Guerra (2007). O primeiro trabalho (MOREIRA et al., 2007), realizado na rede pública do município de Natal (RN), após discussão do constructo por um painel de especialistas e avaliação das relações internas por meio de análise fatorial, propôs um instrumento com cinco dimensões (interação terapeuta-paciente, ambiente físico e conveniência, marcação de consulta, acesso, estrutura física). Já o trabalho de Mendonça e Guerra (2007), também

realizado no Rio Grande do Norte, propôs um instrumento com uma estrutura com quatro dimensões (interação terapeuta-paciente, ambiente físico, acesso e atendimento da recepcionista e do pessoal de apoio e ainda conveniência). Em ambos os estudos, a dimensão terapeuta-paciente foi a mais relevante para explicar a satisfação do paciente.

No panorama internacional, alguns estudos também já propuseram instrumentos com aplicabilidade nesta área (KNIGHT et al., 2010; CURRY e SINCLAIR, 2002; GOLDSTEIN et al., 2000; FEENEY et al., 2008). Todos apresentaram estruturas multidimensionais com mais de quatro fatores ligados à satisfação, com características similares às citadas nos instrumentos nacionais.

Entretanto, tais ferramentas apresentam dificuldade de aplicação devido ao excessivo número de itens, ou por serem genéricos a todas as áreas da saúde (ALMEIDA et al., 2013), ou até mesmo por possuírem alguma fragilidade metodológica baseando-se nos critérios do *checklist* proposto pelo grupo COSMIN (TERWEE et al., 2011; MOKKINK et al., 2010). Vale destacar que este *checklist* pode ser utilizado também para avaliar a qualidade metodológica do processo de validação dos instrumentos de medição, e sua utilidade já foi descrita em revisões sistemáticas recentes (TERWEE et al., 2011; ALMEIDA et al., 2015). Trata-se de uma ferramenta padronizada para avaliar qualitativamente o processo de validação e análise psicométrica de estudos voltados a adaptação ou desenvolvimento de questionários. Esta ferramenta contém nove propriedades de medida, com 5 a 18 critérios em cada uma para avaliação do delineamento do estudo e do tratamento estatístico. Como citado anteriormente, as nove propriedades de medida avaliadas são: consistência interna, confiabilidade, erro de medição, validade de conteúdo, validade estrutural, testes de hipóteses, validade transcultural, validade de critério e capacidade de resposta. Cada propriedade é avaliada separadamente. Isto significa que, se várias propriedades da medida são apresentadas no estudo, todas essas deverão ser consideradas na avaliação qualitativa. O conteúdo avaliado em cada item gera um resultado para cada propriedade avaliada, classificadas em: Excelente, Bom, Regular, Ruim (TERWEE et al., 2011).

Na perspectiva de utilização no contexto brasileiro, baseando-se na avaliação obtida pela aplicação do *checklist* COSMIN e na facilidade para operacionalização dos instrumentos, o MedRisk foi o que apresentou melhores atributos. O quadro 2 apresenta a síntese da avaliação das propriedades da medida dos instrumentos, encontrados na literatura, voltados para mensuração da satisfação de pacientes com serviços de

fisioterapia. A descrição completa da avaliação metodológica sugerida pelo *checklist* COSMIN é apresentada no Anexo A.”

O quadro 3 apresenta as dimensões e os itens dos instrumentos encontrados e as compara com o MedRisk - instrumento escolhido no presente. Quando comparado aos outros cinco instrumentos, o MedRisk além de abarcar conceitualmente a maioria das dimensões incluídas nos demais instrumentos, apresenta a vantagem operacional de conter menor número de itens.

Quadro 2: Avaliação qualitativa das propriedades de medida dos instrumentos para mensuração da satisfação de pacientes com serviços de fisioterapia.

Propriedades para validação avaliadas com base no <i>checklist</i> COSMIN*	QUALIDADE DOS INSTRUMENTOS					
	MedRisk (Beatti et al.,2002)	PTOPS** (Roush and Sonstroem, 1999)	Goldstain (2000)	Monnin e Perneger (2002)	Mendonça e Guerra (2007)	Moreira et al. (2007)
Consistência interna	Regular	Regular	Regular	Regular	Ruim	Regular
Confiabilidade	Regular	Regular	Não descrito	Regular	Não descrito	Não descrito
Medida de Erro	Regular	Ruim	Não descrito	Não descrito	Não descrito	Não descrito
Validade de Conteúdo	Bom	Regular	Ruim	Regular	Regular	Regular
Validade Estrutural	Regular	Regular	Regular	Regular	Regular	Regular
Teste de hipóteses	Ruim	Ruim	Ruim	Ruim	Regular	Não descrito
Validade Transcultural	Regular [¶]	Regular ^{¶¶}	Regular ^{¶¶¶}	Não Descrito	Não descrito	Não descrito
Validade de critério	Não descrito	Não descrito	Não descrito	Não Descrito	Não descrito	Não descrito
Responsividade	Não descrito	Não descrito	Não descrito	Não Descrito	Não descrito	Não descrito

Legenda: *As propriedades para validação dos instrumentos podem ser avaliadas como: excelente, bom, regular, ruim (TERWEE et al, 2012); **PTOPS - *physical therapy outpatient satisfaction questionnaire*; [¶]A validade transcultural foi avaliada com base no estudo de Oliveira et al.,(2014), que propôs uma versão adaptada para a cultura brasileira; ^{¶¶}A validade transcultural foi avaliada com base no estudo de Vanti et al. (2013), que propôs uma versão adaptada para cultura italiana; ^{¶¶¶}A validade transcultural foi avaliada com base no estudo de Mendonça (2004), que propôs uma versão adaptada para cultura brasileira.

Quadro 3: Comparação do MedRisk e demais instrumentos para mensuração da satisfação do paciente na fisioterapia – dimensões e itens dos demais instrumentos presentes ou ausentes no MedRisk.

MedRisk (Beattie et al, 2002)	PTOPS (Roush and Sonstroem, 1999)	Goldstain et al (2000)	Momnin e Perneger (2002)	Mendonça e Guerra (2007)	Moreira <i>et al</i> (2007)
D1- Ambiente físico	D1 - Potencializadores da satisfação	D1 - Tratamento	D1 - Admissão	D1 - Interação paciente-terapeuta	D1 - Relação Terapeuta Paciente
i1-i3	i2 / i12 / i16 / i18 / i22 / i26 - presentes	i11 / i17 / i20 / i21 - presentes	i1 / i2 / i3 - presentes	i1 / i3 / i4 / i5 / i7 / i8 - presentes	i1 / i3 / i4 / i6 / i8 / i9 / i11 / i12 - presente
D2- Interações do paciente	i7 / i24 / i29 / i34 - ausentes	i19 - ausente	D2 - Tratamento	i2 / i6 - ausentes	i2 / i5 / i7 / i10 / i13 / i16 / i29 / i30 - ausentes
i4-i10	D2 - Opositores da satisfação	D2 - Privacidade	i5 / i6 - presentes	D2-Ambiente Físico	D2 - Ambiente físico e conveniência
Medida Global de Satisfação	i6 / i8 / i15 / i27/ i33 - presentes	i 7 - ausente	i4 / i7 / i8 - ausentes	i17 / i18 - presentes	i19 / i20 / i21 / i22 - presentes
i11-i12	i3 / i13 / i20 / i23 / i31 - ausentes	D3 - Conveniência do atendimento	D3 - Logística	i19 / i20 - ausentes	i23 / i26 - ausentes
TOTAL - 12 ITENS	D3 - Localização	i15 / i18 - presentes	i9 / i11 / i12 - presentes	D3 - Acesso e atendimento da recepção e apoio	D3 - Marcação de consulta
	i10 / i4 / i32 - presentes	D 4 - Custos	i10 - ausente	i9 / i10 / i13 / i14 - presentes	i17 / i18 - presentes
	i5 / i19 / i21 / i28 - ausentes	i24 / i25 - ausentes	D4 - Satisfação Global	i11 / i12 - ausentes	D4 - Acesso
	D4 - Custos	D5- Faturamento	i13 / i14 - ausentes	D4 - Conveniência	i27 / i28 - ausentes
	i14 / i9 / i11 / i17 / i25 / i30 - ausentes	i16 - ausente	TOTAL - 14 ITENS	i15 / i16 - presentes	D5 - Estrutura física
	TOTAL - 34 itens	D 6 - Facilidade no agendamento		TOTAL - 20 ITENS	i14 / i15 - presentes
		i 10 - ausente			TOTAL - 30 ITENS
		D7 - Agenda disponível			
		i 12 - presente			
		i 13 - ausente			
		D8 - Tempo de espera			
		i14 - presente			
		D9 - Amabilidade da equipe			
		i9 - presente			
		D 10 - Amabilidade do Fisioterapeuta			
		i8 - presente			
		Satisfação Geral			
		i22 / i23 - presentes			
		TOTAL - 25 itens			

Legenda: PTOPS – Physical Therapy Outpatient Satisfaction; D – dimensão; i – itens presentes ou não no MedRisk.

2.5.3 O instrumento MedRisk

O MedRisk é uma ferramenta desenvolvida em língua Inglesa, nos Estados Unidos, com o objetivo de avaliar a satisfação de usuários de serviços de saúde. Especificamente concebido para a área da fisioterapia no ano de 2002 por Beattie et al., este instrumento já foi utilizado em culturas diferentes e reavaliado por seus autores (BEATTIE et al., 2002; BEATTIE et al 2005; BEATTIE et al., 2007). Apoiados pela *World Confederation for Physical Therapy*, uma cooperação internacional de pesquisadores e interessados buscaram a adequação deste instrumento em vários países, disponível para as línguas Alemã, Mandarim, Árabe, Norueguesa, Coreana e Indonésia (OLIVEIRA et al., 2014).

Inicialmente, durante seu processo de desenvolvimento, 25 itens foram listados considerando a opinião de especialistas, a discussão com pacientes e a revisão de literatura. Essa lista de itens foi examinada por dois fisioterapeutas clínicos, com título de doutor e experiência acadêmica. Ao final dessa fase, 18 itens foram selecionados para o estudo piloto. Esses 18 itens foram aleatoriamente ordenados visando controlar possível viés na resposta; também a construção gramatical incluiu sentenças afirmativas e negativas. Adicionalmente, foram ainda incluídas como variáveis dependentes, quatro medidas da satisfação global: “(1) *Overall, I am completely satisfied with the services I receive from my therapist,*” (2) “*I would return to this office for future care,*” (3) “*I would not recommend this therapist to a friend,*” and (4) “*This therapy has helped me as much as expected.*”

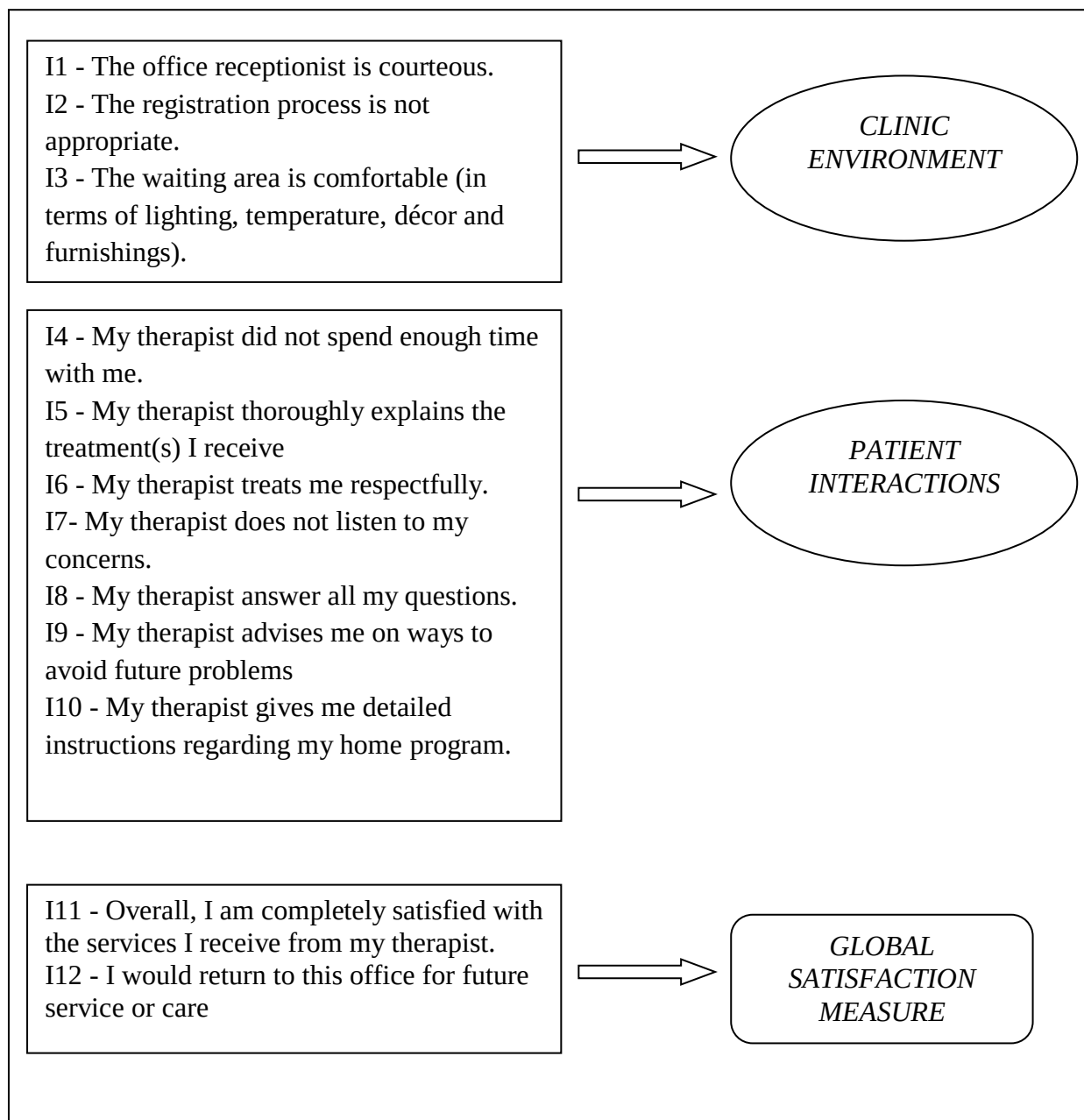
O estudo piloto aplicou o instrumento proposto com 18 itens em 192 pacientes de clínicas de fisioterapia. A análise de correlação entre os itens e as medidas globais de satisfação variou entre 0,35 e 0,74. Posteriormente, a análise multivariada (regressão linear) indicou que duas medidas globais de satisfação (“*I would not recommend this therapist to a friend,*” and (4) “*This therapy has helped me as much as expected.*”) não eram fortemente associadas aos itens propostos. A regressão linear múltipla, contendo duas medidas globais de satisfação e os 18 itens propostos, explicou 76% da variação na satisfação do paciente (desfecho). Portanto, o resultado do estudo piloto foi um questionário com 18 itens específicos e duas medidas globais de satisfação: (1) “*Overall, I am completely satisfied with the services I receive from my therapist*”; (2) “*I would return to this office for future care*”.

Na etapa de validação, o instrumento foi aplicado em 1.868 indivíduos, sem distinção de gênero ou idade, com perfis clínicos e regiões geográficas diferenciados. Uma análise de componentes principais (*principal componente - PCA*), seguida pela análise fatorial exploratória foram conduzidas para determinar o número de dimensões do instrumento. Os resultados indicaram dois modelos possíveis: (i) modelo com três dimensões e 12 itens; (ii) modelo com duas dimensões e 10 itens. Com base na revisão das cargas fatoriais de cada dimensão e discussão com especialistas, foi escolhido o modelo 2. Este apresentava duas dimensões denominadas: *Patient Interactions* (ex. atendimento do especialista, explicação do tratamento) e *Clinic environment* (ex. acessibilidade, localização, estrutura), contendo um total de 10 itens e duas medidas globais de satisfação (Figura 1). Dos 10 itens, três (itens 4, 7 e 11) foram invertidos para controlar respostas automáticas, ou seja, evitar que os pacientes preencham todo o questionário de maneira padrão e sem atenção (BEATTIE et al., 2002). A Figura 1 apresenta os itens e a estrutura dimensional do instrumento final proposto pelos autores (BEATTIE et al., 2002).

Três anos após o desenvolvimento do instrumento, as propriedades psicométricas foram reavaliadas pelos autores (BEATTIE et al., 2005) para confirmar a capacidade discriminativa das duas dimensões. Nessa etapa, foi realizada uma análise confirmatória do modelo bidimensional proposto, com base em uma amostra de 1.448 pacientes com perfis clínicos diferenciados. Os resultados mostraram um ajuste entre bom a excelente, com as cargas fatoriais distribuídas nas dimensões *Patient interactions* (7 itens) e *Clinical environment* (3 itens). As medidas de erro padrão (*standard error of measurement - SEM*)² obtidas mostraram-se aceitáveis para as duas dimensões (*Clinical environment*: 0,24 e *Patient interaction*: 0,19). Face à inexistência de um padrão de referência para instrumentos de satisfação do paciente em fisioterapia, a análise da validade concorrente (fase de *teste de hipótese segundo o checklist COSMIN*) baseou-se no cálculo da correlação entre os dois fatores (dimensões) com as duas medidas globais de satisfação constituintes do instrumento (BEATTIE, 2005). Esta análise também mostrou valores aceitáveis ($r = 0,71 - 0,83$).

² Auxilia o julgamento da confiabilidade em estudos transversais; indica o grau de erro de uma dada medida, valores mais próximos do zero são desejáveis (Beattie et al, 2002)

Figura 1: Composição original do instrumento MedRisk; estrutura dimensional, itens e medidas globais de satisfação do paciente.



Há ainda a avaliação da estrutura do instrumento realizada em um centro de fisioterapia especializado em quiropraxia. O instrumento proposto inicialmente, antes da modelagem final, com 18 itens e duas medidas globais de satisfação, foi aplicado numa amostra de 333 pacientes. A análise fatorial confirmatória mostrou uma estrutura bidimensional (dois fatores), porém com 13 itens, dos quais 5 destes relativos à dimensão *Patient interactions* e 7 outros na dimensão *Clinical environment* (BEATTIE et al., 2011).

Esse instrumento foi também submetido a processos de adaptação transcultural (Beattie et al, 2007). Foram realizadas traduções e retrotraduções para língua espanhola e esta versão foi aplicada em 203 indivíduos que possuíam o espanhol como língua nativa, mas residentes nos Estados Unidos.

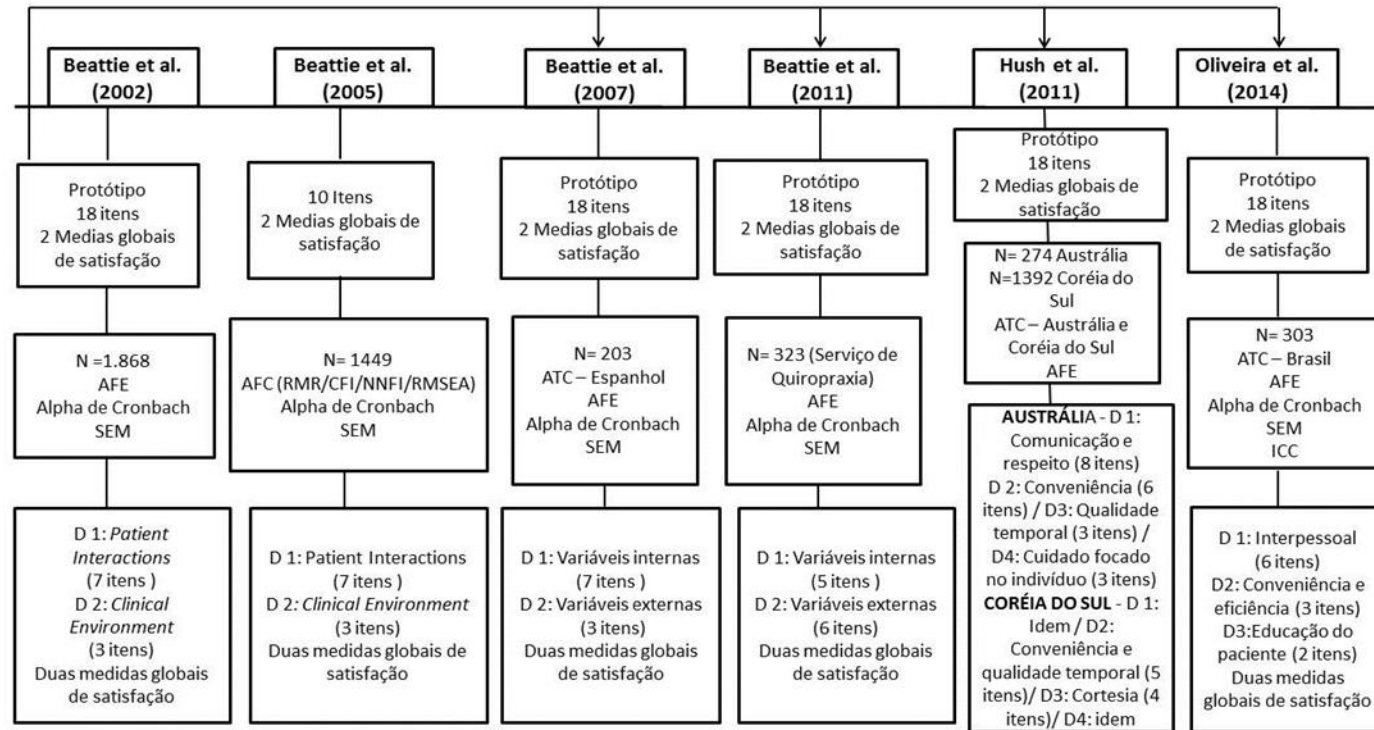
Em outro estudo, Beattie et al (2005), usando o instrumento original com 18 itens, realizaram uma análise fatorial confirmatória que apresentou um modelo aceitável para uma estrutura com dois fatores (7 itens na dimensão *Clinical environment* e 3 na *Patient interactions*), ou seja, confirmou o modelo inicialmente proposto. Igualmente, as medidas de erro (SEM) e a validade concorrente mostraram valores semelhantes aos obtidos para a versão em Inglês.

A utilização deste instrumento nos contextos Australiano e Coreano também foi descrita. Em um estudo multicêntrico, Hush et al 2013 conduziram um processo de adequação semântica para a língua Coreana e avaliação das propriedades psicométricas em serviços de fisioterapia da Coreia e da Austrália. Ambas adaptações transculturais apresentaram resultados semelhantes na análise fatorial exploratória, indicando um modelo de quatro fatores; no contexto australiano, o modelo proposto explicou 61% da variância da satisfação do paciente e no contexto coreano, o modelo explicou 55%.

Recentemente, o instrumento passou por um processo de adaptação transcultural para o contexto brasileiro. O estudo conduzido por Oliveira et al. (2014) propôs uma versão do MedRisk vertida para a língua portuguesa após adequação semântica (tradução, retro tradução e discussão em painel de especialistas) e análise psicométrica. Oliveira et al. (2014) avaliaram o instrumento baseados numa amostra de 303 pacientes de serviços de fisioterapia inseridos nas redes do SUS e da saúde suplementar na cidade de Belo Horizonte (MG). A análise fatorial exploratória indicou uma estrutura dimensional com 3 fatores, nos quais a carga fatorial foi aceitável para a maioria dos itens testados. Assim, seis itens apresentaram carga no fator denominado “*Interpessoal*”; três itens foram retidos no fator denominado “*Conveniência e eficiência*”; e o terceiro fator denominado “*Educação do paciente*” é constituído por dois itens. A análise da consistência interna apresentou valores do alpha de Cronbach de 0,77, 0,63 e 0,67, respectivamente para os fatores 1, 2 e 3. Em geral, as sub-escalas (dimensões) demonstraram moderados níveis de reprodutibilidade (analisados pelo coeficiente de correlação intra-classe [*intraclass correlation coefficient*]), com bom nível de concordância para duas sub-escalas e baixa concordância para sub-escala

“Patient education” (analisados pelo minimal detectable change e standard error of measurement). Os resultados apontaram ainda para um elevado efeito de teto (*ceiling effect*) (Oliveira et al., 2014). A figura 2 apresenta uma síntese dos resultados e versões propostas pelos estudos voltados para avaliação das propriedades psicométricas do MedRisk. Assim, o instrumento foi submetido a algumas etapas importantes do processo de validação (Figura 2), entretanto, para sua utilização, em larga escala, é ainda imprescindível seu teste e avaliação em diferentes cenários de uma mesma cultura.

Figura 2: Cronologia dos estudos de validação do MedRisk: instrumento inicial, análises já realizadas e modelos propostos.



Legenda: AFE – análise fatorial exploratória; SEM – *standard error of measurement*; D – dimensão; AFC- análise fatorial confirmatória; RMSEA - *Root Mean Square Error of Approximation*; CFI - *Comparative Fit Index*, NNFI – *non normed fit index*; RMR – *root mean square residual*; ATC – adaptação transcultural; ICC – *intraclass correlation*.

3 METODOLOGIA

3.1 Delineamento do estudo

O delineamento deste estudo se insere no campo do desenvolvimento e validação de instrumentos e particularmente focou na adaptação transcultural do instrumento MedRisk. Este instrumento se propõe a mensurar a satisfação do paciente, importante conceito relacionado à dimensão da qualidade do cuidado denominada “centralidade do cuidado no paciente” (TZELEPIS et al., 2014; ALMEIDA et al., 2015).

O presente estudo adotou a perspectiva universalista para avaliação da equivalência transcultural do MedRisk (REICHENHEIM e MORAES, 2007; OLIVEIRA et al., 2014; HUSH et al., 2013). A escolha desse instrumento baseou-se na sua especificidade ao cuidado fisioterápico, no amplo uso e validação em diferentes contextos culturais e acesso livre e sem ônus ao instrumento. Embora a versão do MedRisk validada apresente 10 itens, optou-se, no presente estudo, por examinar os 18 itens específicos originalmente testados pelos autores do instrumento MedRisk e duas medidas globais de satisfação também contidas no instrumento original (BEATTIE et al., 2002). Essa opção buscou alcançar a melhor adaptação possível do instrumento para a sua aplicação no contexto brasileiro específico: pacientes do SUS submetidos a cuidado fisioterápico na cidade do Rio de Janeiro, RJ.

Alguns autores sintetizaram os principais passos metodológicos necessários ao processo de adaptação transcultural (HERDMAN et al., 1998; REICHENHEIM e MORAES, 2007; MOKKINK et al., 2010). Considerando os principais passos elencados, este estudo foi desenvolvido em duas etapas: “(i) Equivalência conceitual, equivalência de itens, equivalência semântica; (ii) Equivalência de mensuração – análise psicométrica (avaliação da validade dimensional e itens componentes, avaliação da confiabilidade). A Figura 3 indica o número de casos incluídos nas distintas etapas do presente estudo.

3.2 Etapa 1 - Equivalência conceitual, equivalência de itens, equivalência semântica

A adequação conceitual e de itens explora o constructo em questão e se apoia em discussões com especialistas e revisão bibliográfica, focando sobre a pertinência dos domínios intrínsecos ao instrumento na nova cultura e se os itens são capazes de identificar tais dimensões (REICHENHEIM e MORAES, 2007). Com base na equivalência conceitual e de itens, avalia-se a validade de conteúdo do instrumento, referindo-se à pertinência dos itens para população-alvo e cultura em que o instrumento será utilizado. A equivalência semântica é outro aspecto importante que avaliará a compreensão dos itens por parte da população-alvo (MOKKINK et AL, 2010).

3.2.1 Equivalência Conceitual e de itens

A revisão conceitual fortalece a discussão da validade de conteúdo da escala, isto é, se os itens realmente são relevantes para o construto, para a população e para o propósito de aplicação (MOKKINK et al., 2010). Apoiados na revisão bibliográfica sobre o tema, que sobretudo destacava a pertinência de se abordar a “satisfação do paciente” como um conceito multidimensional, a discussão sobre a equivalência conceitual e de itens foi realizada no âmbito de um painel de especialistas. Este painel foi constituído por quatro especialistas com experiência nas áreas de Avaliação de Serviços de Saúde e Qualidade do Cuidado, Adaptação transcultural e Avaliação Psicométrica de Questionários, Epidemiologia e Fisioterapia. Assim, na primeira reunião do painel de especialistas, discutiu-se a relevância dos conceitos, dimensões do constructo e adequação de cada item do instrumento, considerando sua aplicação num contexto distinto do qual fora desenvolvido.

3.2.2 Equivalência Semântica

A adequação semântica busca harmonizar o sentido denotativo e conotativo das sentenças do instrumento original com o traduzido para nova língua (REICHENHEIM E MORAES, 2007; PAIXÃO et al., 2007).

Como recomendado (BEATON, 2000), o processo de equivalência semântica envolveu quatro etapas: (i) tradução; (ii) retro-tradução; (iii) discussão no painel de especialistas; (iv) pré-teste do instrumento. Na primeira etapa, duas traduções

independentes do instrumento original em inglês para o português (T1 e T2) foram realizadas por um profissional de saúde com domínio do inglês e um tradutor profissional, ambos brasileiros. Na segunda etapa, foram realizadas duas retro-traduições (RT1 e RT2), também de forma independentes, onde um tradutor possuía como língua materna o inglês. Na terceira etapa, um painel de quatro especialistas, composto por três participantes da etapa de avaliação da equivalência conceitual e por um tradutor anglófono com proficiência em português e experiência nesse tipo de abordagem metodológica, foi constituído para comparação das traduções, retro-traduições e original. O painel discutiu o sentido denotativo de cada palavra e o sentido conotativo, ou seja, o que determinada palavra/expressão traz de representação conceitual para os diversos cenários em que podem ser aplicados. Nesta etapa, foram discutidos e propostos ajustes semânticos entre o instrumento original e a versão em português, pois palavras utilizadas no instrumento original pareciam, segundo percepção dos membros do painel, inadequadas ao contexto de adaptação. Também foram elaboradas sentenças alternativas e escolha de palavras similares para aqueles itens julgados como potencialmente mais difíceis de compreensão. O produto do painel de especialista foi a versão pré-teste do instrumento na língua portuguesa (Apêndice B).

A quarta etapa foi o pré-teste para avaliar a compreensão da versão brasileira na população-alvo e essa etapa incluiu duas rodadas: o pré-teste em si e o reteste (Figura 3). Para o pré-teste, foram escolhidos aleatoriamente 30 pacientes do serviço de fisioterapia ambulatorial do Hospital Universitário Gaffrée e Guinle (HUGG/UNIRIO). Este serviço foi escolhido por ser um hospital escola e pela facilidade e autorização de acesso ao serviço pelo pesquisador principal. Após responder ao instrumento, foi solicitado aos indivíduos que sinalizassem os itens não compreendidos e que parafraseassem aquele item. Foram ainda apresentadas sentenças alternativas sugeridas pelo painel para julgar melhorias no entendimento (Apêndice C). Todas as palavras ou expressões identificadas pelo painel como passíveis de dupla interpretação foram levadas aos respondentes do pré-teste para apreciação. Os itens não compreendidos foram rediscutidos pelo painel de especialista para ajustes semânticos e retestados na população alvo. O critério de seleção do item para retorno ao campo para novo pré-teste foi a sinalização de incompreensão por ao menos 4 entrevistados (~15%) (REICHENHEIM e MORAES, 2007; PAIXÃO et al, 2007). Para o reteste foram selecionados outros 25 pacientes (Figura 3) para avaliar a percepção e entendimento do

significado conotativo e denotativo daqueles itens considerados incompreensíveis, e portanto, modificados.

3.3 Etapa 2: Equivalência de mensuração: análise psicométrica

Para a avaliação da equivalência de mensuração, a versão do MedRisk, vertida para língua portuguesa, após avaliação da adequação semântica, foi aplicada em um grupo populacional maior da população-alvo. A avaliação da equivalência de mensuração (psicometria) do instrumento pode ser realizada por modelagens que se apoiam em duas bases teóricas: a Teoria do Teste Clássico (TTC) e a Teoria de Resposta ao Item (TRI). A TTC foi fundamental para a aceitação da possibilidade de quantificar aspectos da psique dos indivíduos, originando a psicometria. A evolução neste campo, inicialmente aplicado às ciências sociais, indicou que os instrumentos de mensuração deveriam ter propriedades/significados equivalentes, independentemente do objeto de medida (THOMAS, 2011). Esta teoria serve de base para avaliação da validade e confiabilidade de um determinado instrumento de medida e permitiu que as análises não se concentrassem apenas na “finalidade do teste”, mas também nos itens constituintes destes testes e nos possíveis erros de medida (CAPELLERI et al., 2014). Por sua vez, a evolução na área psicométrica legitimou a Teoria de resposta ao item. Ambas as teorias, TTC e TRI, se debruçam em premissas semelhantes, como por exemplo, a existência de erros de medida, de variáveis observadas e de variáveis latentes (fatores) – não observadas. Entretanto, na TRI, os modelos são ajustados em função de dois principais parâmetros: do indivíduo e do item (THOMAS, 2011; TENNANT e CONAGHAN, 2007). Assim, os modelos baseados na TRI permitem uma melhor avaliação escalar do instrumento, isto é, maior acurácia quanto a capacidade discriminativa e dificuldade que cada item representa dentro do instrumento de medida (THOMAS, 2011; GULLER et al., 2014).

De forma geral, pode-se assumir que os parâmetros da TRI estão diretamente associados aos parâmetros de uma análise fatorial com variáveis categóricas (como a análise das cargas e limiares dos itens) também utilizados na metodologia proposta pela TTC. Nesta condição, a equivalência entre os modelos baseados na TRI e na análise fatorial baseados na TTC é destacada (THOMAS, 2011). Entretanto, deve-se ressaltar que existem certas distinções, visto que as análises baseadas na TRI são usualmente

empregadas para a avaliação do item e de parâmetros escalares do instrumento, enquanto que a análise fatorial geralmente é considerada nas abordagens de análise configuracional do instrumento (dimensões) e métrica (cargas fatoriais, erros e relações entre os itens) (THOMAS, 2011; CAPELLERI et al., 2014). Considerando esses pontos, o presente trabalho se propôs a analisar a equivalência de mensuração (psicometria) do instrumento MedRisk, abarcando a análise de sua estrutura configuracional e métrica - relações entre os itens, seus fatores e cargas fatoriais (REICHENHEIM et al., 2014). Desta forma, o presente estudo utiliza parâmetros também oriundos da TRI, entretanto não emprega modelagens exclusivas da TRI, excluindo a análise escalar do instrumento. Assim, as análises realizadas detiveram-se sobre a avaliação da validade dimensional e itens componentes do instrumento, assim como avaliação da confiabilidade do MedRisk.

3.3.1 Validade dimensional e itens componentes

Na ausência de um instrumento de referência [*gold standard*], optou-se, neste estudo, por analisar a correlação entre os itens e as duas medidas globais de satisfação do instrumento original (I19 – “Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta” e I20 – “Eu voltaria a este serviço para futuros atendimentos ou cuidados”. Por se tratarem de medidas diretas e descritas na literatura como úteis para identificação de satisfação (BEATTIE et al., 2002; HUSH et al., 2012; ALMEIDA et al., 2015), apenas os itens que apresentaram correlação estatisticamente significativa com essas medidas globais de satisfação foram incluídos para análise da estrutura dimensional do instrumento. Tal abordagem foi realizada em outros estudos voltados para validação de instrumentos de satisfação (BEATTIE et al., 2002; OLIVEIRA et al., 2014; ALMEIDA et al., 2015).

Para a avaliação da estrutura dimensional do instrumento, inicialmente foi empregada a análise fatorial exploratória (AFE). Nesta técnica, os parâmetros são estimados para construção exploratória do modelo. Assim, a definição dos fatores (dimensões) e seus respectivos itens agregados (cargas) podem ser examinados sem um modelo/teoria prévios (REICHENHEIM e MORAES, 2007; SCHREIBER et al., 2006).

Sequencialmente, foi utilizada a análise fatorial confirmatória (AFC), aplicada a partir do modelo obtido na AFE. Esta técnica parte de um pressuposto teórico para estimativa de um modelo determinado, permitindo comparar diferentes modelos para o

mesmo instrumento. Este procedimento visa analisar o ajuste do modelo encontrado e auxilia o aprimoramento da estrutura dimensional do instrumento (SCHREIBER et al., 2006). A modelagem compara a matriz de covariância de uma dada população (estimada pelo modelo) com a matriz observada no estudo empírico. Assim, é possível confirmar ou rejeitar a estrutura dimensional proposta para os instrumentos de aferição na área da saúde (BROWN, 2006). Na AFC, existem diferentes testes para avaliar a adequação de modelo; no presente estudo, optou-se por adotar aqueles frequentemente utilizados para identificar três tipos de ajustes: (i) o absoluto - *Weighted Root Mean Square Residual (WRMR)*; (ii) o parcimonioso - *Root Mean Square Error of Approximation (RMSEA)*; (iii) o comparativo - *Comparative Fit Index (CFI)* e *Tucker-Lewis (TLI)* (SCHREIBER et al., 2006; REICHENHEIM, PAIXÃO e MORAES, 2009). A literatura aponta que valores aceitáveis para estes índices são: $WRMR < 0,90$; $RMSEA < 0,06$; $CFI \geq 0,95$; $TLI \geq 0,95$ (SCHREIBER et al, 2006).

Ainda na AFC, também foi testado o modelo proposto pelos autores do instrumento original – 10 itens e 2 dimensões (BEATTIE et al., 2002) e o modelo obtido por Oliveira et al. (2014) – 11 itens e 3 dimensões - que adaptou e aplicou o MedRisk no contexto brasileiro.

3.3.2 Confiabilidade

Análise da Consistência Interna

A consistência interna é parte da análise da confiabilidade do instrumento e pode ser definida como a inter-relação dos itens de uma mesma escala. Deve-se levar em consideração ainda se o grupo de itens conforma um modelo que reflete o construto a ser mensurado. Embora existam diversas maneiras de analisar a consistência interna de um item (HENSON, 2001; BROWN, 2006), esta usualmente é realizada com base no cálculo do alpha de Cronbach. O alpha deve sempre ser realizado para cada dimensão/subescala constituinte do instrumento, onde são aceitáveis valores maiores a 0,70 e desejáveis valores maiores a 0,90 (HENSON, 2001; MOKKINK et al., 2010). No presente estudo, foi calculada a consistência interna para cada dimensão constituinte do instrumento.

Medidas repetidas / teste-reteste

A confiabilidade é também avaliada com base em medidas realizadas em um intervalo de tempo. No presente trabalho, o instrumento foi aplicado para o mesmo indivíduo em dois momentos distintos. Os primeiros 150 respondentes do MedRisk preencheram novamente este instrumento, presencialmente, em ambiente semelhante, após o intervalo de 5 dias. Este intervalo foi definido no intuito de minimizar que possíveis mudanças no estado de saúde afetassem o escore de satisfação com o serviço. Assim, a reprodutibilidade do MedRisk, mensuração da consistência de medidas repetidas, foi avaliada utilizando a análise do coeficiente de correlação intra-classe (*intraclass correlation* - ICC), onde valores acima de 0,70 são considerados bons e acima de 0,90 excelentes (Vet et al, 2006). Também, foi avaliada a reprodutibilidade por meio da concordância entre as medidas, utilizando medidas de erro. Para tal foram utilizados o *Standard Error of Measurement* (SEM) - desvio padrão dos erros intrínsecos da medida - e o *Minimal Detectable Change* (MDC), que pode ser definido como o intervalo de mudança necessária para ultrapassar o erro e realmente modificar o valor verdadeiro obtido (*true score*). Valores mais próximos do zero são desejáveis para estas medidas (VET et al., 2006; TERWEE et al., 2007).

3.3.3 Contexto do estudo, população, coleta dos dados e digitalização das informações

Foi realizado um breve diagnóstico sobre o padrão de utilização do cuidado ambulatorial fisioterápico prestado aos pacientes no Sistema Único de Saúde (SUS) no estado do Rio de Janeiro. Com base nas informações disponíveis no sistema de informação ambulatorial e hospitalar do Rio de Janeiro, no ano de 2014, foram realizados 415.798 procedimentos ambulatoriais, totalizando um montante de R\$ 2.062.478,44. Estes recursos foram distribuídos por 55 estabelecimentos cadastrados na rede SUS do estado do Rio de Janeiro (SURCA/SMSDC-Rio, 2014), e, dentre estes serviços, foi selecionado o serviço de fisioterapia do Hospital Universitário Gaffrée e Guinle (HUGG/UNIRIO) para compor o universo de estudo. A escolha dessa unidade para a aplicação do instrumento MedRisk se apoiou na sua representatividade dado o volume de atendimentos realizados (11.321 procedimentos); que em 2014 ocupava o sétimo lugar em termos do total de atendimentos ambulatoriais. Por estar inserido em

um hospital-escola, esse ambulatório apresenta perfil clínico diversificado, maior complexidade no cuidado oferecido e maior variabilidade de procedimentos diagnóstico e terapêutico. Por fim, também foi considerado o critério de conveniência e acesso a este serviço.

Assim, a aplicação dos questionários MedRisk vertidos para a língua portuguesa (APÊNDICE D) teve a população de estudo composta por 325 pacientes do ambulatório de fisioterapia do HUGG/UNIRIO (Figura 3). A população de estudo foi selecionada baseada em uma amostra de conveniência, incluindo todos os pacientes registrados e que estavam em atendimento no ambulatório (~ 400) no período de Agosto de 2014 a Janeiro de 2015. O número de indivíduos incluídos (325 casos) é semelhante ao tamanho médio amostral utilizado em demais estudos de validação na área de fisioterapia (OLIVEIRA et al., 2014; MENDONÇA e GUERRA, 2007) e encontra-se dentro do número mínimo amostral considerando na diretriz do grupo COSMIN (TERWEE et al., 2011; MOKKINK et al., 2010).

Além dos itens constituintes no MedRisk, foram coletadas informações sócio-demográficas e clínicas dos pacientes, como: idade, gênero, área de tratamento, perfil nosológico (doenças crônicas), escolaridade e renda familiar média.

A aplicação do instrumento MedRisk vertido para o português seguiu a mesma formatação proposta pelos autores do instrumento original: formato impresso, auto-aplicável, com uma escala *Likert* com graduação de concordância de cinco pontos (discordo fortemente, discordo, neutro, concordo, concordo fortemente) (BEATTIE et al., 2002). Os pacientes foram abordados antes do atendimento e neste momento era perguntado se queriam participar do estudo. Em seguida, era apresentado o termo de consentimento livre e esclarecido (TCLE) e, caso soubessem ler e escrever, era solicitada a leitura do termo e o preenchimento do formulário para coleta dos dados sócio-culturais e resposta aos itens do MedRisk (APÊNDICE E). O tempo estimado deste processo foi 20 minutos para cada paciente. Os dados foram coletados em dias úteis semanais durante todo horário de funcionamento do ambulatório (das 8 às 17 horas).

Os questionários preenchidos foram digitados e nesse arquivo os pontos da escala *Likert*, referentes às respostas do instrumento, foram transformadas em valores numéricos, com valores variando de 1 (discordo fortemente) a 5 (concordo fortemente)

para cada item. Os itens, cuja escala de resposta era reversa (itens negativados na construção gramatical), foram recodificados (1=5; 2=4; 5=1; 4=2).

3.3.4 Análise dos dados

Para as análises estatísticas descritivas e o cálculo de correlações (coeficiente de Spearman) entre itens e as duas medidas globais de satisfação, o nível de significância aceito foi valor de $p < 0,05$.

Para avaliação da consistência interna, foi calculado o valor do alpha de Cronbach para cada dimensão separadamente, e o valor adotado, como aceitável, foi $\alpha > 0,70$.

Na avaliação da reprodutibilidade do instrumento, analisou-se a confiabilidade teste-reteste, com base no coeficiente de correlação intra-classe (*intraclass correlation*: ICC_{Agreement}) e seu intervalo de confiança (CI – 95%); e ainda a concordância, utilizando duas medidas de erro: *Standard Error of Measurement* (SEM) e *Minimal Detectable Change* (MDC), calculados segundo as formulas abaixo (VET et al., 2006):

- $SEM = DP_{DIF} / \sqrt{2}$; (formula 1) - onde: DP_{DIF} é o desvio padrão da diferença entre as médias das duas medidas.
- $MDC = 1.96 \times \sqrt{2} \times SEM$; (fórmula 2) - onde SEM é o valor obtido na fórmula 1

Foi realizada análise fatorial exploratória (AFE) por componente principal para redução das variáveis (AFE), com rotação oblíqua *promax* para depuração das cargas fatoriais. A rotação oblíqua, *a priori*, não estipula relação entre os fatores, diferentemente da ortogonal, que assume não haver correlação entre os fatores. Assim, a rotação oblíqua pode extrair resultados mais precisos se houver correlação entre os fatores. A utilização, em especial da rotação oblíqua *promax*, foi escolhida por ser usualmente utilizada em instrumentos com estruturas fatoriais mais simples (DAMÁSIO, 2012), estrutura suposta para o MedRisk (BEATTIE et al., 2002; OLIVEIRA et al., 2014). O critério de inclusão no modelo foi inicialmente a correlação estatisticamente significativa dos itens com as medidas globais de satisfação (I17 e I18), além do fator de “comunalidade” dos itens ser $\geq 0,50$. O critério para definição das cargas fatoriais de cada item foi a relação 0,60/0,40; ou seja, para o item ser retido em um determinado fator, a carga deveria ser $\geq 0,60$ em um fator e $\leq 0,40$ nos demais

fatores. Apenas fatores com autovalores maiores do que 1,0 foram considerados (BROWN, 2006; HAIR et al., 1987).

As análises descritivas, a correlação, a consistência interna, a reprodutibilidade e a fatorial exploratória acima descritas foram realizadas no pacote estatístico SPSS 17.0.

Para a confirmação do modelo estrutural encontrado na AFE e comparação com resultados dos estudos de Beattie et al. (2002) e de Oliveira et al. (2014), foi empregada a análise fatorial confirmatória (AFC) utilizando o pacote MPLUS 5.0. O modelo foi estimado com base nos “mínimos quadrados robustos ajustados” (WLSMV), indicado quando há distribuição dos dados moderadamente não-normal (Skewness < 2 e Kurtosis < 7). Na AFC, foram utilizadas 4 medidas de adequação do modelo, e os pontos de corte para indicar a adequação de modelo foram $RMSEA < 0,60$; $CFI \geq 0,95$; $TLI \geq 0,95$; $WRMR < 0,90$. Estes mesmos índices foram utilizados para aceite ou não de todos os modelos avaliados na AFC, incluindo a presente tese (BEATTIE et al., 2002; OLIVEIRA et al., 2014).

Dados faltantes (missing data)

O tratamento de dados faltantes deve ser realizado de acordo com o motivo que desencadeou tal resultado. Resumidamente, essas perdas podem ser (i) completamente aleatória (MCAR) - quando o dado faltante acontece realmente ao acaso; (ii) aleatória (MAR) - quando o dado faltante depende das variáveis preenchidas; e (iii) não aleatória (MNAR), quando as perdas ocorrem em variáveis que o pesquisador não consegue observar/controlar (SCHAFFER e GRAHAM, 2002). Algumas técnicas estatísticas podem ser adotadas, tais como imputação simples, imputação múltipla e *Full Information Maximum Likelihood* e outros (NUNES et al., 2010), desde que respeite alguns pressupostos referentes aos motivos que levam à perda do item.

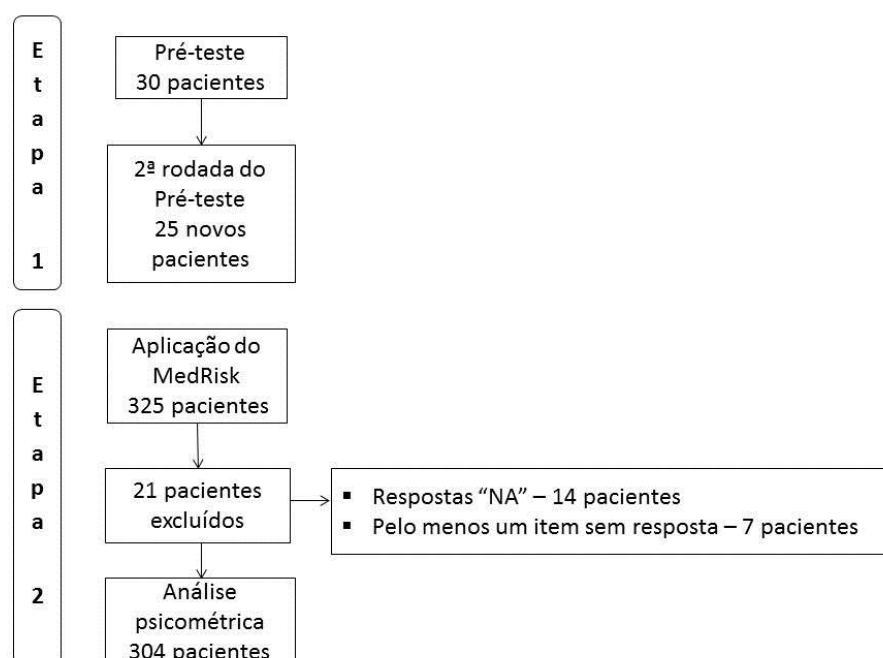
No presente estudo, optou-se por utilizar o método *Listwise deletion*, no qual assume-se que as perdas são aleatórias (MAR) e assim, o indivíduo, com pelo menos um dado faltante, é excluído. Tal escolha deveu-se à dificuldade de identificar os motivos que levaram a não resposta (esquecimento, dificuldade de entendimento do item, falta de vontade para responder). Consequentemente, houve perda no poder estatístico dos testes devido à diminuição do número de casos, entretanto sem grandes prejuízos à veracidade da informação. Ademais, o próprio instrumento inclui a opção

“não se aplica” entre as respostas possíveis, o que pode ser compreendido pelos respondentes como semelhante a não obrigatoriedade de responder a todos os itens.

Analogamente, os questionários, com pelo menos um item assinalado com “não se aplica” (NA), também foram excluídos. Embora o NA seja uma resposta prevista no instrumento, em última instância sua seleção representa um dado faltante. Neste caso específico, imputar o dado seria “forçar uma resposta não verdadeira”, já que o respondente assinalou esta opção. A opção de retirar apenas o item também foi considerada um risco, pois poderia impactar na confiabilidade e validade das informações. Como a fase de pré-teste e as discussões no painel de especialistas não sinalizaram nenhum item incompatível ou desarmônico com o cenário de aplicação do instrumento, o número de respostas “NA” esperado era baixo. Em suma, os questionários com alguma resposta NA ou com, pelo menos, um item não respondido, foram considerados perda e excluídos da análise.

Dentre os 325 indivíduos, para os quais o instrumento foi aplicado, 14 assinalaram, pelo menos, uma opção de resposta “não se aplica” (NA) em algum item, e 7 deixaram, pelo menos, um item sem resposta. Assim, seguindo a abordagem *Listwise deletion* para dados faltantes, 21 observações foram excluídas, logo as análises fatoriais contaram com 304 questionários (Figura 3).

Figura 3: Fluxograma do número de casos incluídos nas etapas do estudo



3.4 Considerações éticas

A pesquisa foi conduzida respeitando a resolução nº 466/12 do Conselho Nacional de Saúde, e o projeto de pesquisa foi aprovado pelo Comitê de Ética em Pesquisa (CEP) da Escola Nacional de Saúde Pública Sérgio Arouca (ENSP/FIOCRUZ) sob parecer de número 427.839 (Anexo B).

Após aprovação do estudo pelo referido comitê, foi realizado um contato oficial com o HUGG para solicitação formal da autorização e realização do projeto e também encaminhamento do memorando de aprovação do CEP. Todos os participantes foram esclarecidos quanto aos objetivos, os riscos e benefícios da pesquisa, explicitados de maneira clara e simples no Termo de Consentimento Livre e Esclarecido (Apêndice E).

O presente estudo não apresentou grandes riscos aos pacientes. Caso o indivíduo declarasse constrangimento com alguma pergunta, ele poderia não respondê-la ou solicitar não participar do estudo. Os benefícios do presente estudo estão relacionados à divulgação dos resultados da pesquisa junto ao serviço de fisioterapia do Hospital Universitário Gaffrée e Guinle, que poderão ajudar a melhorar a qualidade do cuidado prestado aos usuários de fisioterapia da rede SUS. Igualmente, a disponibilização do instrumento MedRisk, após a avaliação da equivalência transcultural, poderá contribuir para a avaliação e melhoria da qualidade também em outras unidades, em especial, ampliaria o papel da percepção do paciente no julgamento do padrão de qualidade do cuidado prestado.

A utilização do MedRisk foi autorizada pelo instituto criado pelos autores que desenvolveram o instrumento – *Expert Clinical Benchmarks* (Anexo C).

4 RESULTADOS

4.1 Etapa 1 – Equivalência conceitual, equivalência de itens e equivalência semântica

A discussão no painel de especialistas sobre os conceitos e dimensões incluídas no MedRisk apontou a pertinência destes para a população e contexto brasileiros. A revisão bibliográfica sobre o constructo satisfação do paciente corrobora os pressupostos teóricos empregados no desenvolvimento do MedRisk. Estes mostram-se coerentes com a cultura e contexto para os quais o instrumento foi inicialmente proposto. Para avaliação da pertinência dos itens e construção gramatical na língua portuguesa, foram utilizadas as versões de síntese das traduções e retro-traduições (Quadro 4, colunas 2 e 3).

O painel de especialistas optou pela exclusão de dois itens. O item 5 (*“O serviço oferece estacionamento conveniente”*) foi retirado em virtude do contexto de aplicação do instrumento na rede SUS. O item é irrelevante, portanto inadequado, em um contexto onde grande parte da população-alvo utiliza o transporte público. Esta estratégia buscou ainda evitar um número elevado de respostas “Não se aplica”, que poderiam gerar grande quantidade de dados faltantes. O item 12 (*“A(o) assistente/ajudante foi respeitosa(o)”*) foi suprimido por referir-se à presença de assistente no tratamento, fato pouco usual no contexto de utilização. O painel de especialistas destacou ainda cinco itens com possível dificuldade de compreensão (itens: 2,4,6,8,11). Após as duas exclusões e apontamentos feitos pelos membros do painel de especialista (Quadro 5), uma versão do instrumento com 16 itens específicos e duas medidas globais de satisfação foi submetida à fase do pré-teste, que verificou a compreensão do instrumento na população-alvo

Dos 30 pacientes do serviço de fisioterapia ambulatorial selecionados para a fase do pré-teste, 60% eram mulheres e a idade média foi 57 (± 13) anos. A maioria dos pacientes possuía ensino básico, ainda que 27%, ensino básico incompleto. Cinquenta e três por cento declarou ganhar de 1 a 3 salários mínimos.

Quadro 4: Versão original e síntese das traduções e retro-traduições do MedRisk.

MedRisk – Original	Tradução (T1 + T2)	Retro-tradução (RT1 + RT2)
1-The office receptionist was courteous.	1-A recepcionista do serviço foi cortês.	1-The service reception was courteous.
2-The registration process was appropriate.	2-O processo de registro foi adequado.	2-The registration process was appropriate.
3-The waiting area was comfortable (lighting, temperature, furnishings).	3-A sala de espera era confortável (iluminação, temperatura, mobília).	3-The waiting area was comfortable (lighting, temperature, furniture).
4-The office location was not convenient.	4-A localização do serviço não era conveniente.	4-The service location was not convenient.
5-This office provided convenient parking.	5-Este serviço proporcionou estacionamento conveniente.	5-The service offers convenient parking.
6-I waited too long to see my chiropractor.	6-Esperei demais para ver meu fisioterapeuta.	6-I had to wait too long to see my physiotherapist.
7-The office hours were convenient for me.	7-O horário de funcionamento foi apropriado para mim.	7-Opening hours were convenient for me.
8-My chiropractor did not spend enough time with me.	8-Meu fisioterapeuta não dedicou tempo suficiente a mim.	8-My physiotherapist did not spend long enough with me.
9-My chiropractor thoroughly explained the treatment(s) I received.	9-Meu fisioterapeuta explicou atenciosamente o tratamento recebido.	9-My physiotherapist gave me a complete explanation of the treatment I received.
10-My chiropractor treated me respectfully.	10-Meu fisioterapeuta me tratou com todo o respeito.	10-My physiotherapist treated me respectfully
11-The office staff was respectful.	11-O pessoal do consultório foi respeitoso.	11-The staff were respectful.
12-The chiropractor's assistant/aide was respectful.	12-O assistente/auxiliar do fisioterapeuta foi respeitoso.	12-The assistant/auxiliary was respectful.
13- My chiropractor did not listen to my concerns.	13-Meu fisioterapeuta não ouviu as minhas preocupações.	13-The physiotherapist did not listen to my concerns.
14-My chiropractor answered all my questions.	14-Meu fisioterapeuta respondeu a todas as minhas perguntas.	14-My physiotherapist answered all my questions.
15-My chiropractor advised me on ways to avoid future problems.	15-Meu fisioterapeuta me aconselhou sobre como evitar problemas futuros.	15-My physiotherapist advised me on how to avoid future problems.
16-The office and its facilities were clean.	16-O serviço e as instalações estavam limpos.	16-The service and premises were clean.
17-The office used up-to-date equipment.	17-Os equipamentos do consultório eram atualizados.	17-The service used modern equipment.
18-My chiropractor gave me detailed instructions regarding my home program.	18-Meu fisioterapeuta me deu instruções detalhadas sobre o meu programa de casa.	18-My physiotherapist gave me detailed guidance about my home programme.
19-Overall, I am completely satisfied with the services I received from my chiropractor.	19-No geral, estou completamente satisfeito com os serviços recebidos do meu fisioterapeuta.	19-Overall, I am completely satisfied with the services I received from my physiotherapist.
20-I would return to this office for future services or care.	20-Retornaria a este escritório para futuros serviços ou cuidados.	20-I would return to this service for other care or treatment.

Quadro 5: Itens da versão utilizada na fase de pré-teste e síntese dos ajustes do painel de especialistas.

Versão pré-teste	Decisões / Ajustes do painel de especialistas
1-A(o) recepcionista do serviço foi atenciosa(o).	1-Foi escolhida a palavra <i>serviço</i> para traduzir a palavra <i>office</i> em todos os itens onde esta aparece, pois foi considerada mais adequada ao cenário de aplicação do instrumento: ambulatório de fisioterapia; <i>Atencioso</i> foi a palavra escolhida para traduzir a palavra <i>courtesous</i> , visando melhor entendimento por parte da população alvo e respeitando o uso coloquial da língua portuguesa no contexto brasileiro. Além disso, no pré-teste buscou-se extrair da população-alvo o que se entendia por <i>atencioso</i> .
2- O processo de registro foi adequado.	2.-A palavra <i>adequado</i> foi escolhida para traduzir a palavra <i>appropriate</i> . No pré-teste buscou-se extrair da população-alvo o que se entendia por <i>apropriado</i> .
3-A sala de espera era confortável (iluminação, temperatura, os móveis).	3-Tradução literal, sem observações.
4-A localização do serviço não era adequada.	4-No pré-teste considerou-se ainda o emprego da palavra <i>conveniente</i> , ao invés da palavra <i>adequada</i> .
5- Item excluído: O serviço oferece estacionamento conveniente.	5-Considerando o contexto brasileiro e cenário de aplicação do instrumento, o item 5, por se referir ao local de estacionamento de automóvel, foi excluído.
6-Eu esperei demais para ser atendido pelo meu fisioterapeuta.	6-Foi escolhido traduzir a palavra <i>chiropractor</i> por <i>fisioterapeuta</i> em todos os itens onde este termo está presente, devido ao cenário de aplicação do instrumento vertido para o português: serviços de fisioterapia e não de quiropraxia - serviço não comum no Brasil e na rede SUS. O pré-teste considerou ainda a utilização ou não do pronome <i>meu</i> – pois interrogava-se a existência ou não do paciente ser acompanhado por um único profissional no cenário de aplicação. Ou seja, sempre o mesmo profissional para os todo os atendimentos?
7-O horário de funcionamento era adequado para mim.	7-Foi considerada ainda o emprego da palavra <i>conveniente</i> , ao invés da palavra <i>adequado</i> na população alvo.

Versão pré-teste	Decisões / Ajustes do painel de especialistas
8-Meu fisioterapeuta não dedicou tempo suficiente para mim.	8-Tradução literal. No pré-teste foi confirmado haver entendimento da afirmativa: “dedicou tempo suficiente a mim”.
9-Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi.	9-Foi usada também outra construção gramatical como alternativa no pré-teste: “Meu fisioterapeuta me explicou todos os detalhes do tratamento que recebi”.
10-Meu fisioterapeuta me tratou com respeito.	10-Tradução literal. O pré-teste confirmou haver entendimento da expressão: “tratou com respeito”.
11-Os funcionários foram respeitosos.	11-Tradução literal. O pré-teste confirmou haver entendimento da expressão: “foram respeitosos”.
12- Item excluído: A(o) assistente/ajudante foi respeitosa(o).	12-Em função do cenário de aplicação do instrumento, o item 12 foi excluído. Este, refere-se a presença de assistente no tratamento, fato não usual no contexto proposto para sua utilização.
13-Meu fisioterapeuta não escutou as minhas preocupações.	13-Foi usado ainda outra construção gramatical no pré-teste: “...não deu atenção às minhas preocupações”.
14-Meu fisioterapeuta respondeu a todas as minhas perguntas.	14-Tradução literal, sem observações.
15. Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros.	15-Tradução literal, sem observações.
16-O serviço e suas instalações eram limpos.	16 – A palavra <i>facilities</i> foi traduzida por <i>instalações</i> . O pré-teste confirmou haver compreensão pela população alvo do termo <i>instalações</i> empregado na versão do pré-teste.
17-O serviço utiliza equipamentos modernos.	17-Tradução literal, sem observações.
18-Meu fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar.	18-Tradução literal. O pré-teste confirmou haver compreensão do termo: “programa domiciliar”.
19-Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta.	19-Tradução literal, sem observações.

Versão pré-teste	Decisões / Ajustes do painel de especialistas
20-Eu voltaria a este serviço para futuros atendimentos ou cuidados.	20- A expressão <i>future services</i> foi traduzida por <i>futuros atendimentos</i> para manter coerência com a opção anterior da tradução de <i>office</i> por <i>serviço</i> - referindo-se ao local do tratamento, evitando assim redundância na versão vertida para o português.

O instrumento utilizado no pré-teste apresentou boa compreensão pela população alvo (Quadro 6). Com exceção do item 2, todos os demais sinalizados, com possível problema de interpretação dúbia pelo painel de especialistas, foram compreendidos pela população-alvo com a mesma interpretação pretendida pelo painel. Da mesma forma, todas as terminologias e construções gramaticais alternativas formuladas pelo painel não impactaram ou melhoraram a compreensão das sentenças.

Na fase do pré-teste, 50% dos pacientes tiveram dificuldade para entender o item 2 (“O processo de registro foi adequado”), possivelmente devido aos termos “processo” e “registro”. O termo processo, talvez por ser vago e impreciso; o termo registro, por incluir uma gama de procedimentos (como agendamento ou inscrição no serviço) e abarcar distintas interpretações no âmbito do serviço hospitalar. Tal fato suscitou o painel de especialistas a retomar a discussão e propor reformulação do item. Assim, alternativamente foi proposta a assertiva: “*A matrícula no serviço de fisioterapia foi adequada*”. O instrumento foi retestado, incluindo a nova construção gramatical do item 2 e os demais itens idênticos a primeira rodada, em outros 25 pacientes, com perfil clínicos e socioeconômico semelhante. Houve significativa melhora no entendimento do item, com apenas um paciente sinalizando alguma dificuldade. A versão do MedRisk para o português proposta após a equivalência semântica é apresentada no Quadro 6.

Quadro 6: Resultados do pré-teste e versão do MedRisk proposta após esta fase.

Resultados pré-teste de cada item	MedRisk Versão para a língua portuguesa Após pré-teste
1-Item com boa compreensão pela população alvo	1-A(o) recepcionista do serviço foi atenciosa(o)
2- Este item não foi compreendido por 50% da população e foi revisto pelo painel de especialistas. Na segunda rodada do pré-teste, após modificação, apenas 5% relataram dificuldade de compreensão.	2-A matrícula no serviço de fisioterapia foi adequada
3-Item com boa compreensão pela população alvo	3-A sala de espera era confortável (iluminação, temperatura, os móveis)
4-Item com boa compreensão pela população alvo	4-A localização do serviço não era adequada
Item 5 excluído pelo painel de especialista	---
6-Item com boa compreensão pela população alvo	6-Eu esperei demais para ser atendido pelo meu fisioterapeuta
7-Item com boa compreensão pela população alvo	7-O horário de funcionamento era adequado para mim
8-Item com boa compreensão pela população alvo	8-Meu fisioterapeuta não dedicou tempo suficiente para mim
9-Item com boa compreensão pela população alvo	9-Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi
10-Item com boa compreensão pela população alvo	10-Meu fisioterapeuta me tratou com respeito
11-Item com boa compreensão pela população alvo	11-Os funcionários foram respeitosos
Item 12 excluído pelo painel de especialista	---
13-Item com boa compreensão pela população alvo	13-Meu fisioterapeuta não escutou as minhas preocupações
14-Item com boa compreensão pela população alvo	14-Meu fisioterapeuta respondeu a todas as minhas perguntas
15-Item com boa compreensão pela população alvo	15-Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros
16-Item com boa compreensão pela população alvo	16-O serviço e suas instalações eram limpos
17-Item com boa compreensão pela população alvo	17-O serviço utiliza equipamentos modernos
18-Apenas 5% da população teve dificuldade com o termo “programa domiciliar”. Como houve entendimento pela grande maioria, foi mantida a sentença	18-Meu fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar
19- Item com boa compreensão pela população alvo	19-Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta
20- Item com boa compreensão pela população alvo	20-Eu voltaria a este serviço para futuros atendimentos ou cuidados

4.2 Etapa 2 - Equivalência de mensuração – análise psicométrica

4.2.1 Análise descritiva e correlações entre os itens

A média de idade da população incluída no estudo foi de 62,1 anos ($\pm 11,7$), dos quais 65% eram do gênero feminino. Quanto ao perfil clínico, a grande maioria foi admitida no setor de fisioterapia ortopédica (72%) e tinham como principais queixas: poliartralgia (26%), disfunção em joelho (18%) e disfunção em coluna lombar (13%). Foram também incluídos pacientes de outras especialidades fisioterapêuticas (Tabela 1). A grande maioria dos pacientes relatou ainda possuir alguma doença associada (52%), além do diagnóstico que o levou a procurar a fisioterapia.

As variáveis socioeconômicas revelaram que 38% dos respondentes estudaram até o ensino básico (Tabela 1). Do ponto de vista do impacto sobre as atividades laborais, a grande maioria (31%) não havia retornado ao trabalho no momento da aplicação do questionário e 29% não trabalhavam antes da lesão (Tabela 1).

Tabela 1: Características clínicas e socioeconômicas da população de estudo

Características da população	
Idade em anos (média; DP)	62,1 ($\pm 11,7$)
Número de casos (%)	
Gênero (Feminino)	212 (65)
Especialidade Fisioterapêutica	
<i>Traumato-ortopedia</i>	234 (72)
<i>Cardio-respiratória</i>	26 (8)
<i>Uro-genecologia</i>	32 (10)
<i>Drenagem linfática</i>	26 (8)
<i>Neurologia</i>	7 (2)
Escolaridade	
<i>Até 5 anos de estudo</i>	91 (28)
<i>De 5 à 9 anos de estudo</i>	124 (38)
<i>De 9 à 12 anos de estudo</i>	74 (23)
<i>Acima de 12 anos de estudo</i>	36 (11)
Impacto sobre as atividades laborais	
<i>Retornaram as atividades</i>	68 (21)
<i>Não retornaram as atividades</i>	101 (31)
<i>A lesão não provocou afastamento</i>	62 (19)
<i>Não trabalhavam</i>	94 (29)
Total	325 (100)

Legenda: DP – desvio padrão.

Quanto à satisfação dos pacientes com o serviço, os valores médios variaram entre 3,88 e 4,68 (Tabela 2). Considerando que, quanto mais próximo de cinco mais satisfeito está o paciente, os valores médios observados indicam satisfação do paciente com o cuidado.

Na análise de correlação entre os itens ($N=304$), 3 não apresentavam correlação estatisticamente significativa ($p>0,05$) com as duas medidas globais de satisfação (I17 e I18): o item 5; item 7 e o item 11 (Tabela 2). Como o modelo conceitual que originou o MedRisk propõe as medidas globais de satisfação (itens 17 e 18) como variáveis dependentes, os itens estatisticamente não correlacionados a estas (5, 7, 11) foram excluídos da análise da estrutura dimensional do instrumento. Além disso, essas medidas globais não representam itens específicos do instrumento MedRisk, portanto não foram incluídas na análise dimensional. Consequentemente, o resultado dessa etapa foi um modelo com 13 itens, submetido à análise fatorial.

Tabela 2: Valores médios dos itens do MedRisk e respectivos coeficientes de correlações (ρ) com as duas medidas globais de satisfação (I17 e I18).

Itens Medrisk	Valores médios (DP)	Correlação I17	Correlação I18
1.A(o) recepcionista do serviço foi atenciosa(o)	4,48 ($\pm 0,63$)	0,20	0,26
2.A matrícula no serviço de fisioterapia foi adequada	4,43($\pm 0,63$)	0,15	0,12
3.A sala de espera era confortável (iluminação, temperatura, os móveis)	3,88($\pm 1,09$)	0,18	0,24
4.A Localização do Serviço não era Adequada **	3,96($\pm 1,17$)	0,17	0,18
5.Eu esperei demais para ser atendido pelo meu fisioterapeuta **	4,01($\pm 1,31$)	-0,10*	-0,11*
6.O horário de funcionamento era adequado para mim	4,30($\pm 0,96$)	0,23	0,14
7.Meus fisioterapeuta não dedicou tempo suficiente para mim **	4,15($\pm 1,24$)	-0,10*	-0,10*
8.Meus fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi	4,53($\pm 0,71$)	0,23	0,18
9.Meus fisioterapeuta me tratou com respeito	4,57($\pm 0,69$)	0,21	0,13
10.Os funcionários foram respeitosos	4,50($\pm 0,67$)	0,13	0,18
11.Meus fisioterapeuta não escutou as minhas preocupações **	4,19($\pm 1,13$)	0,02*	0,07*
12.Meus fisioterapeuta respondeu a todas as minhas perguntas	4,46 ($\pm 0,74$)	0,20	0,22
13.Meus fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	4,50 ($\pm 0,61$)	0,24	0,32
14.O serviço e suas instalações eram limpos	4,32 ($\pm 0,74$)	0,31	0,38
15.O serviço utiliza equipamentos modernos	4,03 ($\pm 0,98$)	0,13	0,22
16.Meus fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar	4,41($\pm 0,73$)	0,41	0,28
17.Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta	4,64 ($\pm 0,63$)	1,0	0,69
18.Eu voltaria a este serviço para futuros atendimentos ou cuidados	4,68 ($\pm 0,58$)	0,69	1,0

Legenda: DP – desvio padrão.

* correlação estatisticamente não significativa: $p > 0,05$;

** Item reverso – valores apresentados após recodificação/inversão.

4.2.2 Análise fatorial exploratória (AFE) e análise fatorial confirmatória (AFC)

O resultado da AFE do modelo com 13 itens (1, 2, 3, 4, 6, 8, 9, 10, 12, 13, 14, 15, 16) levou à exclusão de dois itens devido à baixa “comunalidade” (item 14 que apresentou valor igual a 0,29 e item 16 com valor de 0,24). Após nova AFE, sem estes itens, obteve-se um modelo com três fatores e 11 itens, que explicava 51,4 % da variância. Entretanto, a análise fatorial confirmatória (AFC) revelou índices de ajuste para este modelo abaixo do recomendado (CLI = 0,72; TLI = 0,69; RSMEA = 0,08). Assim, o modelo foi retestado com base na AFE. Nessa etapa, 3 outros itens com valores ainda mais baixos de “comunalidade” foram então excluídos: item 2 “A matrícula no serviço de fisioterapia foi adequada”; item 8 “Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi” e item 12 “Meu fisioterapeuta respondeu a todas as minhas perguntas”.

Desta forma, um modelo com três fatores e 8 itens foi novamente submetido à análise exploratória. Este novo modelo apresentava todas as “comunalidades” maiores do que 0,5 e explicava 58,9% da variância (Tabela 6)³. O fator 1 reteve os itens 9,10,13; o fator 2 reteve os itens 1,3,15; e o fator 3, os itens 6 e 4 (Tabela 3, modelo 1). Embora as cargas fatoriais dos itens covariantes ao fator 3 tenham mostrado valores aceitáveis, a presença de um fator com apenas dois itens pode impactar na qualidade do ajuste do modelo proposto (Brown, 2006). Portanto, outro modelo foi testado pela AFE, sem os itens 4 e 6 (fator 3). O novo modelo encontrado apresentou todas as “comunalidades” >0,5 e a variância explicada foi de 59,2% (Tabelas 3 e 4). Logo, ao final dessa etapa, o modelo 2 (Tabela 3), contendo dois fatores e seis itens, apresentando cargas fatoriais aceitáveis e maior poder de explicação da variância, foi escolhido para novo reteste na AFC.

Tabela 3: Análise fatorial exploratória: autovalores e percentual da variância explicada nos modelos com 3 e 2 fatores.

Modelo	Autovalores	Variância explicada (%)	Acumulado da variância explicada (%)
Modelo 1			
F1	2,083	26,04	26,04
F2	1,588	19,09	45,89
F3	1,046	13,75	58,97
Modelo 2			
F1	1,784	29,73	29,73
F2	1,708	28,47	59,20

³As saídas (*outputs*) dos programas SPSS 17.0 e MPLUS 5.0 referentes às análises estatísticas estão disponíveis nos apêndices F e G respectivamente

Tabela 4: Análise fatorial exploratória; itens e respectivas cargas fatoriais

Itens	Modelos		
	Modelo 1 (3 fatores)		
	Fator 1	Fator 2	Fator 3
10.Os funcionários foram respeitosos	0,774	0,125	0,021
9. Meu fisioterapeuta me tratou com respeito	0,756	- 0,030	0,025
13.Meus fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	0,734	0,133	0,239
1.A(o) recepcionista do serviço foi atenciosa(o)	0,199	0,756	- 0,021
3.A sala de espera era confortável (iluminação, temperatura, os móveis)	- 0,011	0,731	0,293
15.O serviço utiliza equipamentos modernos	0,012	0,730	0,304
6.O horário de funcionamento era adequado para mim	0,162	0,124	0,795
4.A localização do serviço não era adequada	0,029	0,286	0,742
	Modelo 2 (2 fatores)		
	Fator 1	Fator 2	
3.A sala de espera era confortável (iluminação, temperatura, os móveis)	0,792	-0,154	
15.O serviço utiliza equipamentos modernos	0,778	0,019	
1. A(o) recepcionista do serviço foi atenciosa(o)	0,732	0,177	
10.Os funcionários foram respeitosos	0,019	0,812	
13.Meus fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	0,038	0,711	
9.Meus fisioterapeuta me tratou com respeito	0,020	0,709	

Considerando o arcabouço teórico que amalgama os itens do MedRisk, o fator 1 foi denominado de dimensão *Estrutura e apoio* ao serviço (itens 1,3,15); e o fator 2 denominado de dimensão *Interpessoal* (itens 9,10,13). Os valores médios, acima de 4,0, em ambas dimensões, mostraram altos índices de satisfação (Tabela 5). As correlações entre os itens e item-total apresentam baixos valores (Tabela 5). O instrumento final proposto com base nos resultados do presente estudo é apresentado no Apêndice H.

Este modelo, obtido no presente estudo (Tabela 4, modelo 2), quando submetido à análise fatorial confirmatória, apresentou resultados satisfatórios para os índices de ajuste do modelo, quais sejam: WRMR=0,58, CFI= 0,97, TLI = 0,96, RMSEA =0,05.

Tabela 5: Modelo proposto do Medrisk*: dimensão, valor médio da satisfação e correlação entre os itens

Dimensão	Valor médio	Itens	Correlação item-total	Correlação entre os itens
Interpessoal	4,47	I9	0,43	0,23 – 0,39
		I10	0,46	
		I13	0,42	
Estrutura e apoio ao serviço	4,13	I1	0,38	0,35 – 0,43
		I3	0,40	
		I13	0,42	

* Modelos com dois fatores e seis itens, obtidos pós análises fatorial exploratória e confirmatória.

A tabela 6 apresenta os resultados da análise fatorial confirmatória (cargas fatoriais, correlação entre os fatores/dimensões, variância residual) e ainda os resíduos/erro encontrados na AFE (*uniquenesses*). As cargas fatoriais confirmadas nesta AFC, embora inferiores às cargas extraídas pela análise fatorial exploratória (Tabela 4), apresentaram valores também satisfatórios (Tabela 6). A correlação entre os fatores foi estatisticamente significativa, porém de baixa magnitude ($r = 0,19$) (Tabela 6).

Tabela 6: Análise fatorial confirmatória da estrutura dimensional do instrumento MedRisk e resíduos da Análise fatorial exploratória

Fator	Item	Carga Fatorial	Correlação entre Fatores (ρ)[*]	Variância Residual	Uniquenesses da AFE^{**}
Interpessoal	9	0.651	0.19	0.576	0.497
	10	0.734		0.462	0.341
	13	0.711		0.495	0.494
Estrutura e Apoio	1	0.624		0.610	0.436
	3	0.620		0.651	0.346
	15	0.652		0.574	0.394

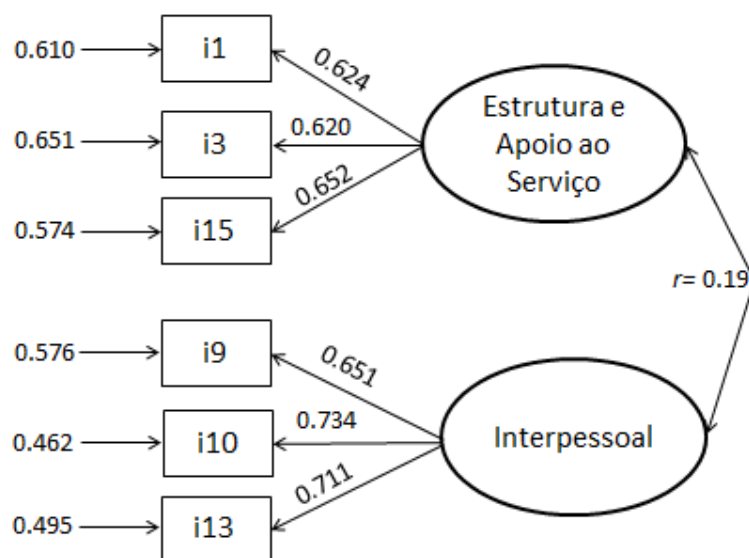
*Valores de correlação estatisticamente significativos, ($p < 0.05$);

** Valores apresentados são resultados da expressão $(1 - \text{comunalidade da AFE})$.

A Figura 4 esquematiza o modelo encontrado para o instrumento MedRisk neste estudo. As formas ovais representam as variáveis latentes (dimensões) e os quadrados representam as variáveis observadas (itens). As setas unidirecionais das variáveis latentes para os itens apresentam as respectivas cargas fatoriais; as demais setas unidirecionais aos itens apresentam o erro (variância

residual). A seta bidirecional entre as variáveis latentes aponta que existe correlação entre estas variáveis e o valor está expresso no centro desta seta (ρ).

Figura 4: Representação esquemática do modelo para o instrumento MedRisk obtido na Análise Fatorial Confirmatória



Ainda foram testados dois modelos alternativos para o instrumento MedRisk. O primeiro foi o modelo final obtido pelos autores do instrumento (Beattie et al., 2002), apresentando duas dimensões e 10 itens: (1) *Patients Interactions* – itens: 7,8,9,11,12,13,16) e (2) *Clinical Environment* (itens – 1,2,3). O segundo modelo foi o obtido por Oliveira et al. (2014), estudo desenvolvido no Brasil utilizando o MedRisk, que apresenta três dimensões e 11 itens: (1) interpessoal - itens 1,2,3,9,10,14; (2) conveniência - itens 4,6,12; (3) educação do pacientes - itens 13,16. Na comparação entre os dois modelos alternativos e o modelo final obtido pelo presente estudo (Tabela 6), os índices de ajuste de ambos modelos alternativos foram inferiores (Tabela 7).

Tabela 7: Comparação dos índices de ajuste do modelo* de três estruturas dimensionais propostas para o MedRisk

Autor/Ano	Modelo	Parâmetros estimados	CFI	TLI	RSMEA
Beattie et al (2002)	2 fatores 8 itens	49	0,46	0,41	0,21
Oliveira et al (2015)	3 fatores 11 itens	54	0,81	0,82	0,11
Almeida, 2015 (presente estudo)	2 fatores 6 itens	28	0,97	0,96	0,05

Legenda: CFI: comparative fit index; TLI: Tucker-Lewis index; RSMEA: Root mean square error of approximation; * Análise fatorial confirmatória por mínimos quadrados robustos ajustados (WLSMV)

4.2.3 Reprodutibilidade – confiabilidade e concordância

As análises de reprodutibilidade foram realizadas em 149 indivíduos, dado que 1 paciente não preencheu todo o instrumento, sendo, portanto, descartado. Os resultados da análise teste-reteste (ICC) se mostraram excelentes para ambas dimensões (Tabela 8). Assim como, para a concordância, por meio das medidas de erro. Contudo, a análise da consistência interna do instrumento, outro aspecto da confiabilidade, apresentou valores moderados. O valor do α de Cronbach foi de 0,63 para os itens do fator 1; 0,60 para os itens do fator 2 (Tabela 8).

Tabela 8: Confiabilidade e reprodutibilidade do instrumento MedRisk.

Dimensão	α Cronbach*	ICC (IC 95%)	SEM	MDC
Interpessoal	0.60 (0.37 – 0.56)**	0.96 (0.94-0.97)	0.13	0.36
Estrutura e apoio ao serviço	0.63 (0.46 – 0.60)	0.98 (0.98-0.99)	0.11	0.30

Legenda: ICC – *intraclass correlation* ; IC – intervalo de confiança; SEM – *standard error of measurement*; MDC – *minimal detectable change*.

* Análise da consistência interna baseou-se em todos os 304 questionários; o ICC e as medidas de erro foram calculadas com base nos dados de 149 questionários

** Resultados entre parênteses apresentam a variação do valor do α caso o item seja excluído.

5. DISCUSSÃO

O instrumento MedRisk, para mensuração da satisfação de pacientes com serviços de fisioterapia, ganhou destaque em decorrência de sua aplicação por diversos grupos (BEATTIE et al., 2011; OLIVEIRA et al., 2014; HUSH et al 2013). Atualmente uma colaboração internacional, apoiada pela confederação mundial de fisioterapia, vem fomentando a adaptação transcultural deste instrumento para diversos países e línguas (OLIVEIRA et al., 2014; HUSH et al., 2013; BEATTIE et al., 2007).

Desde o primeiro estudo de validação (BEATTIE et al., 2002), o histórico do uso do instrumento MedRisk permitiu produzir evidências que corroboram sua estrutura dimensional, reprodutibilidade e confiabilidade. Em 2005, os mesmos autores que o desenvolveram (BEATTIE et al., 2005), com base em uma grande amostra populacional (n= 1.449), confirmaram o modelo proposto pelo instrumento original com índices de ajuste excelentes (CFI = 0,98; RMR= 0,014; RSMEA = 0,05). Entretanto, dois itens apresentaram cargas fatoriais baixas (<0,40). Contudo, somente o presente estudo e o trabalho desenvolvido por Beattie et al. (2005) realizaram análise fatorial confirmatória do modelo original do MedRisk (duas dimensões, 10 itens e duas medidas globais de satisfação). Desta forma, o modelo intrínseco proposto no MedRisk não pode ser considerado definitivo e, conseqüentemente, pronto para uso indiscriminado em diferentes contextos (Brown, 2006).

Os achados do presente trabalho discordam ligeiramente dos obtidos por Beattie et al. (2005) no que tange à dimensionalidade do instrumento, pois eles confirmaram o modelo original com três itens na dimensão *Patients Interactions* e sete itens na dimensão *Clinical Environment*. Entretanto, os achados são convergentes quanto a reprodutibilidade do MedRisk apresentada por Beattie et al. (SEM = 0,24 e 0, 19 para cada dimensão). Contudo, vale ressaltar que a fórmula utilizada para cálculo do SEM no estudo de Beattie et al (2005) foi diferente da utilizada no presente estudo. Os autores não realizaram medidas repetidas (teste/reteste), logo esta medida foi calculada pela fórmula: $DP \times \sqrt{1 - \alpha}$; cujo DP é o desvio padrão da média do fator e α é o alpha de Cronbach. Existe suporte na literatura para utilização de ambas as fórmulas (VET et al., 2006), porém a fórmula dependente do alpha é sensível ao número de itens de uma mesma dimensão (NIMON et al., 2012). Assim, esta abordagem pode gerar valores distorcidos em instrumentos com poucos itens por dimensão, como o instrumento estudado na presente tese.

Em 2011, as propriedades psicométricas do MedRisk foram avaliadas em usuários de um serviço de quiropraxia nos Estados Unidos - esta é uma especialidade da fisioterapia (BEATTIE et al., 2011). Essa avaliação encontrou uma estrutura bidimensional, porém os itens retidos foram

diferentes daqueles incluídos no instrumento validado em 2002 (BEATTIE et al., 2011). Curiosamente, este foi o único estudo que apresentou os 18 itens específicos e as duas medidas globais de satisfação inicialmente examinados pelos autores do instrumento. No primeiro estudo de validação, Beattie et al. (2002) apenas citam que o modelo partiu de 18 itens, porém só apresentam os itens retidos no modelo final testado (com 10 itens e duas medidas globais de satisfação). A adaptação do Medrisk para uso em serviços de quiropraxia refez o percurso metodológico, partindo do instrumento originalmente examinado. A presente tese optou por, também, refazer o percurso metodológico do MedRisk; assim partiu da mesma versão utilizada por Beattie et al (2011) nos serviços de quiropraxia, adotando a formatação e opções de resposta empregadas neste estudo.

Embora o instrumento MedRisk tenha sido reavaliado quanto as suas propriedades métricas na população onde foi desenvolvido, (BEATTIE et al., 2002; BEATTIE et al., 2005; BEATTIE et al., 2011), os estudos de adaptação transcultural apresentaram resultados distintos. A adaptação transcultural para a língua espanhola (BEATTIE et al., 2007) corroborou a estrutura bidimensional do instrumento, contudo os itens apresentados na versão em espanhol não foram os mesmos do instrumento original. Partindo do modelo inicial com 18 itens, a avaliação dessa adaptação mostrou uma estrutura de duas dimensões, porém os itens 8,9,10,13,14,15,18 foram retidos na dimensão *Internal* e os itens 1,2,3 na dimensão *External*. Esta solução se mostrou idêntica ao modelo original para dimensão *Clinical environment* (itens 1,2,3), porém o estudo original de validação apresentou na outra dimensão (*Patients interactions*) os itens 7,8,9,11,12,13,16. Além disso, a nomeação utilizada para as dimensões parece denotar alguma distinção teórica-conceitual.

No estudo multicêntrico internacional, conduzido por Hush et al. (2013), que adaptou transculturalmente o Medrisk para população coreana e analisou as propriedades psicométricas do instrumento nas populações australiana e coreana, obteve-se resultados distintos. Os achados apontaram uma estrutura com quatro dimensões em ambas as populações. Contudo, o estudo carece de detalhes da análise dos dados, pois apresentaram apenas a redução de itens obtida na análise fatorial exploratória e a variância explicada pelos modelos (61% para a população australiana e 55% para coreana).

A mais recente avaliação psicométrica do MedRisk apresentou também o processo de adaptação para o português (Oliveira et al., 2014). Contemporâneo, mas independente do presente estudo, o MedRisk foi adaptado para o contexto brasileiro e aplicado em pacientes de serviços de fisioterapia de um município do estado de Minas Gerais; os resultados indicaram uma estrutura com três dimensões e excelente grau de reprodutibilidade (OLIVEIRA et al., 2014).

Comparação das duas versões brasileiras do instrumento MedRisk

Tanto este estudo como o desenvolvido por Oliveira et al. (2014) adotaram a perspectiva universalista (HERDMAN et al., 1998) e se detiveram sobre as mesmas etapas metodológicas para avaliar a equivalência transcultural.

O Quadro 7 apresenta as similaridades e diferenças semânticas entre as duas versões do MedRisk para a língua portuguesa. Nesse quadro, a numeração dos itens segue a do instrumento com 18 itens específicos e mais duas medidas globais de satisfação ao final - dois últimos itens (BEATTIE et al., 2011). Contudo, nem todos os itens são apresentados, pois Oliveira et al. (2014) descreveram apenas os itens retidos após avaliação das cargas fatoriais. As construções gramaticais das duas versões se mostraram diferentes para os itens: 1, 2, 4, 5, 6, 7, 8, 10, 11 e 13. Das diferenças listadas (Quadro 7), vale destacar a referente ao item 2. Na versão de Oliveira et al. (2014) a sentença para este item é “O processo de registro foi adequado”. No presente estudo, uma tradução idêntica foi inicialmente proposta pelo painel de especialistas. Entretanto, durante a realização do pré-teste, 50% dos indivíduos relataram não compreender essa afirmativa, levando a modificação da sentença para melhor compreensão, como apresentado anteriormente no Quadro 6. A dificuldade de compreensão de itens pode ocorrer em função do conteúdo dos itens e das opções de resposta (MEYER et al., 2007). No caso específico, o problema de compreensão nesse item (2) foi relacionado à expressão “processo de registro”.

Tanto o artigo publicado (OLIVEIRA et al., 2014), como a dissertação de mestrado que o originou (OLIVEIRA, 2013), não detalharam o processo de discussão no painel de especialistas nem o pré-teste. Assim, não foi possível avaliar se houve dificuldade na interpretação de algum outro item e/ou algum ajuste. No presente estudo, vale reforçar a importância do pré-teste que identificou problemas na compreensão dos itens do instrumento, como aliás reportado na literatura (REIS, LAGUARDIA e MARTINS, 2012; REICHENHEIM e MORAES, 2007; PAIXÃO et al., 2007). A questão que se coloca é que, possivelmente, se a versão proposta por Oliveira et al. (2014) fosse aplicada no cenário deste estudo, a avaliação da satisfação dos pacientes poderia ser distorcida por conta da dificuldade no entendimento do item.

Uma questão que desponta do estudo de Oliveira et al. (2014) diz respeito à versão do MedRisk utilizada. No instrumento utilizado por esses autores, há um item extra apresentado ao final, relacionado à mensuração de resultado do tratamento (com utilização de uma medida de auto percepção de mudança geral do estado de saúde). A versão apresentada por esses autores como sendo a original e base da adaptação realizada, mostrava ainda diferenças no cabeçalho instrutivo e

mudança nas opções de respostas, com a supressão da opção “não aplicável” (NA) nos itens 2,4,6,7,8,9,10, 11,12,13,14,15,16,17,18,19,20. Contudo, esta versão não havia sido apresentada em publicações anteriores do MedRisk (BEATTIE et al., 2002; BEATTIE et al., 2005; BEATTIE et al., 2007; BEATTIE et al., 2011, HUSH et al., 2012). Embora as versões tomadas como ponto de partida para adaptação para o português não sejam idênticas, os 18 itens específicos e as duas medidas globais para satisfação que compõem o MedRisk eram idênticos. Tal fato viabiliza alguma comparação entre as duas versões.

Oliveira et al. (2014) não indicaram correlação entre satisfação e a mudança do estado de saúde geral. Neste sentido, a inserção de uma medida global do estado de saúde não mostrou relevância para a mensuração da satisfação com serviço e, além disso, parecem ser constructos distintos. Em geral, medidas do estado de saúde global são utilizadas em inquéritos populacionais ou na avaliação da satisfação com um tratamento específico, diferentemente da pesquisa de satisfação com um serviço de saúde (HUDAK e WRIGHT, 2000). Beattie et al. (2011) aplicaram dois distintos questionários para mensurar satisfação e outro para mensurar a auto percepção da mudança global no estado de saúde (categorizado em “melhorei” ou “não melhorei”). Os achados desses autores (BEATTIE et al., 2011) evidenciaram que, independentemente da mudança no estado de saúde, a frequência observada de satisfeitos ou muito satisfeitos foi elevada (valores acima de 4,0 com a utilização do MedRisk); além de não ter sido encontrada associação entre satisfação e resultado positivo com o tratamento. Quanto à supressão das opções de resposta NA, não foram encontrados estudos que discutissem tal estratégia pelos autores do instrumento original.

Quadro 7: Comparação semântica entre as duas versões do instrumento MedRisk para a língua portuguesa

Construção Gramatical	
Presente estudo	Estudo de Oliveira et al (2014)
1-A(o) recepcionista do serviço foi atenciosa(o)	1-A recepcionista foi cortês
2-A matrícula no serviço de fisioterapia foi adequada	2-O processo de registro foi adequado
3-A sala de espera era confortável (iluminação, temperatura, os móveis)	3-A sala de espera era confortável (iluminação, temperatura, móveis)
4-A localização do serviço não era adequada	ND
5-Eu esperei demais para ser atendido pelo meu fisioterapeuta	ND
6-O horário de funcionamento era adequado para mim	6-Os horários de atendimento desta clínica foram convenientes para mim
7-Meu fisioterapeuta não dedicou tempo suficiente para mim	ND
8-Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi	8-Meu fisioterapeuta me explicou cuidadosamente os tratamentos que eu recebi
9-Meu fisioterapeuta me tratou com respeito	9-Meu fisioterapeuta me tratou respeitosamente
10-Os funcionários foram respeitosos	10-Os funcionários da clínica foram respeitosos
11-Meu fisioterapeuta não escutou as minhas preocupações	ND
12-Meu fisioterapeuta respondeu a todas as minhas perguntas	12-Meu fisioterapeuta respondeu a todas as minhas questões
13-Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	13-Meu fisioterapeuta aconselhou-me sobre formas de evitar futuros problemas
14-O serviço e suas instalações eram limpos	14-A clínica e suas dependências estavam limpas
15-O serviço utiliza equipamentos modernos	ND
16-Meu fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar	16-Meu fisioterapeuta forneceu-me instruções detalhadas sobre meu programa de exercícios para casa
17-Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta	17- De uma forma geral, estou completamente satisfeito (a) com os serviços que eu recebi do meu fisioterapeuta
18-Eu voltaria a este serviço para futuros atendimentos ou cuidados	18-Eu retornaria a esta clínica para futuros serviços ou tratamento

Legenda: ND-Não disponível a tradução do item para a língua portuguesa (Oliveira et al., 2014)

Quanto à configuração da estrutura dimensional, há diferenças entre o resultado deste estudo com o de Oliveira et al. (2014). Esses autores apresentaram uma dimensão denominada de “Interpessoal”; entretanto dois itens parecem não estar conceitualmente conectados a esta dimensão: (i) “Item 3: A sala de espera era confortável (iluminação, temperatura, móveis) e (ii) “Item 14: A clínica e suas dependências estavam limpas”. Nos resultados aqui obtidos, o item 3 foi retido na dimensão denominada de “Estrutura e Apoio ao Serviço”. Esta decisão buscou coerência entre o conteúdo do item e o conceito imbricado na dimensão. O item 14 não foi retido em nenhuma dimensão do presente estudo. A outra dimensão, definida por Oliveira et al. (2014), foi “Conveniência e eficiência”, porém dois itens nessa dimensão parecem incoerentes com o conceito que define a dimensão, quais seja: (i) “Item 12: Meu fisioterapeuta respondeu a todas as minhas questões” e (ii) “Item 8: Meu fisioterapeuta me explicou cuidadosamente os tratamentos que eu recebi”. Dado que o conceito de eficiência, no contexto da avaliação de serviços de saúde, está associado a evitar desperdício de insumos e equipamentos ou maior produtividade (IOM, 2001), unificar conveniência e eficiência numa única dimensão, não parece conceitualmente adequado. Assim, o conteúdo desses dois itens (8 e 12) parece intrinsicamente relacionados à relação terapeuta-paciente, que é um dos principais pressupostos teóricos do conceito de satisfação do paciente (ALMEIDA et al., 2015; ALMEIDA et al, 2013; SCHOLTE et al, 2014).

Por fim, a análise de reprodutibilidade do instrumento foi outro aspecto que apresentou discordância. No presente estudo, as medidas de reprodutibilidade e confiabilidade teste-reteste foram excelentes. Oliveira et al. (2014) encontraram reprodutibilidade e confiabilidade moderadas para o instrumento, com valores do coeficiente de correlação intra-classe variando entre 0,51 e 0,85 e medidas de concordância também divergentes. Contudo, outros estudos sobre o MedRisk também mostraram valores semelhantes aos aqui encontrados, com SEM variando entre 0,19-0,25 (BEATTIE et al., 2005); 0,16-0,25 (BEATTIE et al.; 2007); 0,17-0,20 (BEATTIE et al., 2011). Entretanto, os outros estudos, excluindo Oliveira et al (2014) e o presente estudo, não avaliaram a confiabilidade do MedRisk utilizando teste-reteste e o coeficiente de correlação intra-classe.

Em suma, as diferenças e discordância nos resultados entre os dois únicos estudos brasileiros indicam, por um lado, a necessidade de um esforço compartilhado de maior vulto, e por outro lado, levanta o questionamento sobre a viabilidade de ajustar um único instrumento validado para todo o contexto brasileiro.

Avaliando a estrutura dimensional do MedRisk

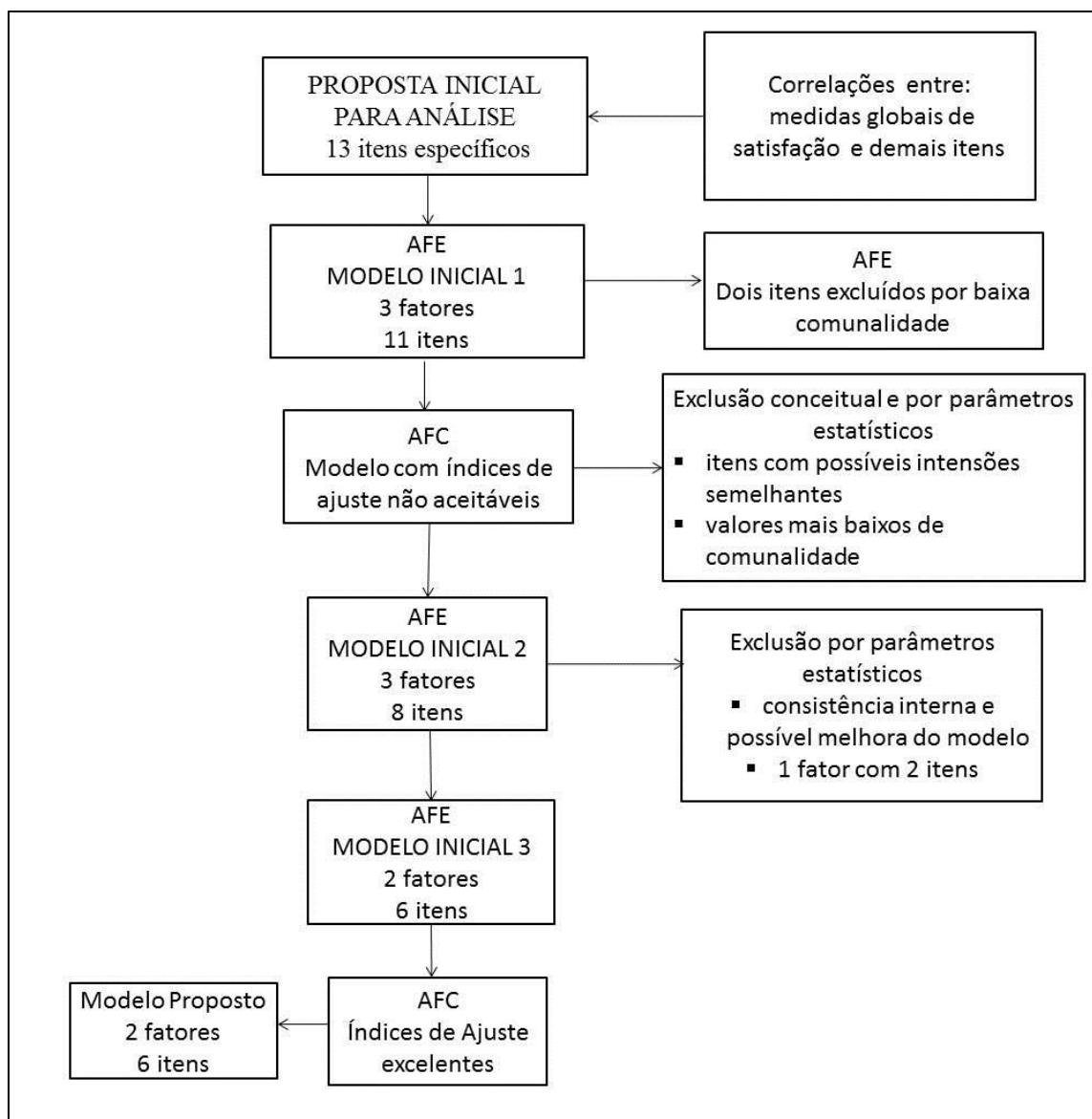
A análise fatorial confirmatória (AFC), embora tenha apontado excelentes índices de ajuste do modelo, não confirmou o modelo proposto por Oliveira et al. (2014) nem o originalmente proposto por Beattie et al. (2002). Contudo, a não confirmação de um modelo não indica sua invalidação, apenas chama atenção para eventuais espaços de melhoria (SCHREIBER et al., 2006). Uma possível explicação para as diferenças encontradas pode estar no número de parâmetros estimados na análise estatística e de observações. Schreiber et al. (2006) apontam que, apesar de não existir uma regra quanto ao número mínimo de observações inseridas na AFC, parece existir um consenso sobre o mínimo de 10 casos para cada parâmetro estimado, dos quais inclui-se as variâncias e covariâncias entre os itens e os fatores. Assim, a população de estudo do presente trabalho (N=304) possivelmente permitiu análises mais estáveis para confirmação do modelo aqui proposto, dado que a estimativa de um modelo com 28 parâmetros minimamente necessitaria de 280 indivíduos. Por sua vez, o teste dos modelos obtidos por Oliveira et al. (2014) e Beattie et al. (2002), devido a quantidade de itens, ao estimarem 54 e 49 parâmetros respectivamente, necessitariam de uma população maior que a do presente estudo (304 casos) para assegurar a robustez das análises.

A Figura 5 apresenta as etapas da tomada da decisão que levaram ao modelo proposto pelo presente estudo. A proposição de um instrumento com seis itens e duas medidas globais de satisfação se apoiou tanto nos parâmetros estatísticos do modelo encontrado quanto nos pressupostos teóricos que sustentam o conceito de satisfação. O modelo proposto é composto por duas dimensões: *Estrutura e apoio ao serviço* (itens 1,3,15) e *Interpessoal* (itens 9,10,13).

Por sua vez, o item 3 “A sala de espera era confortável (iluminação, temperatura, os móveis)” e o item 15 “O serviço utiliza equipamentos modernos”, claramente estão relacionados aos aspectos da estrutura do serviço. Embora o item 1 “A(o) recepcionista do serviço foi atenciosa(o)” gere um tipo de informação evidentemente distinta, optou-se por manter este item nesta dimensão. Esta decisão foi baseada na carga fatorial do item, e na inexistência de carga cruzada, além disso, houve correlação com os itens desta dimensão, relevância para consistência interna desta dimensão e confirmação do modelo pela AFC (BROWN, 2006; SCHREIBER, 2006; REICHENHEIM e MORAES, 2012). Vale ainda destacar que a recepção/recepcionista do serviço é geralmente o primeiro contato do paciente com o cuidado procurado, podendo ser fundamental para o tipo de “experiência do paciente” e também para o “funcionamento dos serviços” – dois elementos essenciais para a qualidade do cuidado e do serviço de saúde orientada para qualidade

(BERWICK, 2002). Argumento esse que auxilia a justificar teoricamente e estatisticamente a manutenção deste item nesta dimensão denominada Estrutura e Apoio ao Serviço.

Figura 5: Modelo proposto para o MedRisk - etapas da tomada de decisão no presente estudo



Fato semelhante ao item 3 aconteceu na dimensão Interpessoal do instrumento. Nesta, os itens 9 -“Meu fisioterapeuta me tratou com respeito”- e 13 -“Meu fisioterapeuta me aconselhou sobre as maneiras de evitar problemas futuros”- são claramente ligados à relação interpessoal entre paciente e fisioterapeuta. Contudo, o item 10 -“Os funcionários foram respeitosos”- poderia também ser visto como um possível elemento constitutivo da dimensão Estrutura e Apoio ao Serviço. Mas considerando os mesmos critérios estatísticos apontados anteriormente (BROWN, 2006;

SCHREIBER, 2006; REICHENHEIM e MORAES, 2012) justificou-se sua manutenção na dimensão Interpessoal. Além disto, a reflexão sobre o significado de “funcionários respeitosos” nos faz pensar que esta construção parece trazer uma informação, não de apoio ao serviço, mas sim da representação mais abrangente de um serviço voltado para o paciente, fato de grande importância para o constructo satisfação do paciente (HUSH et al., 2011; TZLEPIS et al., 2014)

Os resultados quanto à consistência interna das dimensões do instrumento são preocupantes, pois os valores do *alpha* de Cronbach são ligeiramente abaixo do aceitável (BROWN, 2006). Sobretudo, quando a revisão sistemática realizada mostrou que o valor do *alpha* é em média maior em outros instrumentos que mensuram a satisfação do paciente (ALMEIDA et al., 2015). Ao compararmos com os outros estudos do MedRisk, nossos achados também foram inferiores: Oliveira et al. (2014) reportaram valores entre 0,63 – 0,77; Beattie et al. (2002) valores entre 0,73 e 0,91; Beattie et al. (2005) valores entre 0,87 e 0,90; Beattie et al. (2007), na adaptação transcultural para o espanhol, encontraram valores entre 0,82 e 0,90; e Beattie et al. (2011) encontraram valores entre 0,90 e 0,91. A adaptação transcultural, realizada por Hush et al. (2013), não reportou os valores da consistência interna. Entretanto, vale ressaltar novamente que, como o *alpha* de Cronbach é sensível ao número de itens de uma escala, instrumentos que apresentam poucos itens em uma mesma dimensão podem sofrer distorção e apresentar valores mais baixos (HENSON, 2001; NIMON et al., 2012). Tal explicação pode responder, em parte, os baixos valores na consistência interna. Outra questão que pode também explicar esses achados foi a correlação baixa ou moderada entre os itens de cada dimensão. Esta magnitude da associação entre os itens pode também influenciar o cálculo do *alpha* de Cronbach (HENSON, 2001; NIMON et al., 2012). Mesmo com estas conhecidas limitações do método, o uso do *alpha* de Cronbach foi mantido para comparação com os estudos anteriores, já que todos utilizaram esse teste para avaliar a consistência interna (BEATTIE et al., 2002; BEATTIE et al., 2005; BEATTIE et al., 2007; BEATTIE et al., 2011).

O presente trabalho corroborou a estrutura bidimensional, embora com itens diferentes em cada dimensão quando comparado ao modelo inicialmente proposto por Beattie et al. (2002) e Beattie et al. (2005). Todavia, contrariou os modelos propostos por Oliveira et al. (2014) e Hush et al. (2012). Vale ressaltar que a não confirmação do modelo original de instrumentos para mensuração da satisfação de pacientes, após validação transcultural, é um achado relativamente comum (VANTI et al., 2013; LAM et al., 2005; MULTINOVIC' et al., 2012).

Nexos conceituais das dimensões de instrumentos voltados para mensuração da satisfação de pacientes

O debate acerca dos nexos conceituais que apoiam os instrumentos para mensuração da satisfação dos pacientes, na fisioterapia em particular, apresenta lacunas. A literatura aponta no sentido da satisfação do paciente ser um constructo multidimensional, entretanto, a definição de quais aspectos devem estar imbricados em cada uma destas dimensões ainda é fonte de debate intenso e longe de consenso (ROUSH e STROEM, 1999; GOLDSTAIN et al., 2000; MONNIN e PERNEGER, 2002; BEATTIE et al., 2002; MENDONÇA e GUERRA, 2007; MOREIRA et al., 2007; OLIVEIRA et al., 2014).

A distinção entre a forma com que a aferição da satisfação dos pacientes é conduzida e também sua finalidade devem ser prioritárias para avanço de tal discussão (HUDAK et al, 2000). A revisão sistemática conduzida por Hush et al. (2011) descreveu 15 diferentes instrumentos para mensuração de satisfação de pacientes na área de fisioterapia. Alguns destinavam-se à investigação da satisfação com determinado cuidado recebido, outros, com o serviço de saúde. A metodologia empregada nestes estudos também foi variada; alguns estudos utilizaram medidas únicas globais (com apenas uma pergunta geral sobre satisfação), outros, questionários multidimensionais ou até mesmo abordagens qualitativas. Discussões em torno de instrumentos tão distintos são mais difíceis, assim, o debate sobre aspectos teóricos constituintes dos instrumentos de satisfação deve perpassar inicialmente pela pergunta: “o que se pretende medir dentro do nexo conceitual da satisfação do paciente – satisfação com o serviço ou com um tratamento específico? Esta decisão é fundamental para escolha do instrumento a utilizar (HUDAK et al., 2000).

Tzelepis et al. (2014) encontraram, na revisão sistemática sobre os instrumentos voltados para a avaliação do cuidado centrado no paciente, que os eixos relacionados à comunicação, informação e educação do paciente são os mais comumente incorporados nas ferramentas disponíveis. Dentro do constructo da satisfação do paciente, existem pressupostos teóricos que podem ajudar a avaliar elementos desses eixos. Esta afirmação considera as dimensões usualmente inseridas nos instrumentos para mensuração da satisfação dos pacientes com serviços de saúde (IOM, 2001; ALMEIDA et al., 2015).

Além disso, as discussões teóricas, que pressupõem os instrumentos, são tão importantes quanto “a pescaria por resultados estatisticamente significativos” (SCHREIBER et al., 2006). Na fisioterapia, poucos são os estudos recentes que utilizaram análises confirmatórias dos instrumentos que medem satisfação dos pacientes (ROUSH e STROEM, 1999; BEATTIE et al., 2005), havendo

assim dificuldade de encontrar instrumentos robustos, tanto conceitualmente quanto estatisticamente.

Neste sentido, o estudo de um mesmo instrumento, por diversos grupos de autores e em diferentes países, pode contribuir para definição de uma ferramenta mais robusta e ainda alimentar o debate teórico sobre a satisfação do paciente em fisioterapia. Os achados, a respeito do instrumento MedRisk, exemplificam isso. Os resultados do presente trabalho corroboram com os estudos anteriores (Beattie et al. 2002; Beattie et al. 2005; Battie et al., 2011; Oliveira et al. 2014) no que tange ao quadro conceitual que amálgama cada dimensão constituinte desta ferramenta. Todos os estudos identificaram as relações interpessoais (no tripé paciente-fisioterapeuta-profissionais de apoio) e a estrutura do serviço como parte constituinte de um instrumento de satisfação. Os estudos desenvolvidos por Beattie et al. (2002; 2005; 2007; 2011) destacam as dimensões *Cinical Environment* (denominadas por eles de variáveis “externas” ao cuidado) e *Patient Interactions* (variáveis “internas” ao cuidado).

O Quadro 8 mostra aspectos conceituais, dimensões e itens de todos os estudos que realizaram adaptação transcultural do instrumento MedRisk. Embora o estudo de Oliveira et al. (2014) aponte para 3 dimensões (Interpessoal, Conveniência e eficiência e Educação do Paciente), as dimensões Conveniência e eficiência e Educação do paciente incorporam aspectos conceitualmente referentes à estrutura do serviço e relação paciente-profissional, respectivamente. Deve-se destacar que, apesar da estrutura e nomenclatura diferentes, os instrumentos citados possuem dimensões explicativas da satisfação que guardam alguma semelhança. Assim, embora os modelos dimensionais propostos sejam distintos, há uma tendência ao agrupamento dos itens ao redor destas dimensões citadas acima. Apenas a adaptação transcultural conduzida por Hush et al. (2013) apresentou resultados discrepantes do ponto de vista dos conceitos imbricados em cada dimensão.

Quadro 8: Comparação dos aspectos conceituais que moldaram as dimensões apresentadas nos estudos de adaptação transcultural do MedRisk

Estudo	País/Língua de adaptação	Dimensões e respectivos itens	Aspectos conceituais imbricados nas dimensões
Beattie et al. (2007)	Estados Unidos/Espanhola	1-Variáveis internas (9,8,10,13,14,15,18) 2-Variáveis externas (1,2,3)	- Relação paciente-profissional da saúde - Estrutura física do serviço e processos de apoio
Hush et al. (2013)	Austrália/Inglesa	1-Comunicação e respeito (1,9,10,11,12,14,15,18) 2-Conveniência (2,3*,4,5,6*,7) 3-Qualidade temporal (6*,8,13) 4-Cuidado focado no indivíduo** (3*,16,17)	- Relação paciente-profissional da saúde - Estrutura física do serviço e processos de apoio - Tempo destinado ao cuidado
Hush et al. (2013)	Coréia/Coreana	1-Comunicação e respeito (1,9,10,11,12,14,15,18) 2-Conveniência e qualidade temporal (6,8,4,5,13) 3-Gentileza (1,2,3*,7) 4-Cuidado focado no indivíduo** (3*,16,17)	-Relação paciente-profissional da saúde - Estrutura física do serviço e processos de apoio - Tempo destinado ao cuidado
Oliveira et al. (2014)	Brasil/Portuguesa	1-Interpessoal (1,2,3,10,11,16) 2-Conveniência e eficiência (7,9,14) 3-Educação do paciente (15,18)	Relação paciente-profissional da saúde - Estrutura física do serviço e processos de apoio - Orientações aos pacientes
Almeida (2015) - Tese de doutorado	Brasil/Portuguesa	1 – Interpessoal (10,11,15) 2 - Estrutura e apoio ao serviço (1,3,17)	- Relação paciente-profissional da saúde - Estrutura física do serviço e processos de apoio

Legenda: * Os autores apresentaram o mesmo item em duas dimensões. ** O estudo mostrou cargas fatoriais negativas nesta dimensão.

Abordagem processual para análise psicométrica do instrumento MedRisk

As questões conceituais que permeiam as dimensões propostas pelos estudos sobre o MedRisk apresentam similaridades com outros estudos de satisfação (HUSH et al, 2011; ALMEIDA et al, 2015). Contudo, a condução da análise psicométrica nos estudos sobre o MedRisk possui particularidades e lacunas processuais que certamente diferem daquelas realizadas em instrumentos mais consolidados e revisitados (REICHENHEIM et al., 2009; REICHENHEIM et al., 2011).

Apesar das lacunas acima destacadas, com o intuito de comparação dos resultados, o presente estudo optou pela mesma abordagem processual utilizada pelos demais estudos que realizaram a análise psicométrica do MedRisk (BEATTIE et al, 2002, 2005, 2007; OLIVEIRA et al, 2014). Primeiramente, o critério aqui adotado, de inserir no modelo de análise da estrutura dimensional (AFE/AFC) somente os itens com correlações estatisticamente significativas com as duas medidas globais de satisfação, possivelmente enviesará o resultado final. Esse viés está relacionado ao risco de “hiper exclusão” de itens, dado que itens não correlacionados às medidas globais de satisfação deveriam também ser testados na análise fatorial confirmatória/exploratória (REICHENREIM et al., 2014). Aliado a este viés processual, a análise da validade de constructo foi circunscrita. Tal análise repousa tanto sobre a estrutura dimensional do instrumento, como sobre aspectos referentes ao padrão das cargas fatoriais, medidas de erro, resíduos, intensidade e discriminabilidade dos itens na estrutura dimensional postulada. Nesse sentido, analisar todos os itens do instrumento original torna-se imprescindível (REICHENREIM et al., 2014).

Embora o *checklist* COSMIN, aqui aplicado, seja uma importante ferramenta para avaliação qualitativa do desenho dos estudos voltados para desenvolvimento e adaptação de instrumentos de medida em saúde, a área ainda carece de discussões referentes aos aspectos processuais que devem nortear a análise configuracional, métrica e escalar de tais instrumentos. As análises confirmatórias de modelos previamente postulados, referentes à estrutura dimensional, devem sempre ser consideradas na análise inicial dos instrumentos e após identificação da necessidade de ajustes no modelo, uma exploração dos dados deve ser conduzida (REICHENREIM et al, 2014). Nestas perspectivas, a tomada de decisão ao longo da análise psicométrica (Figura 5), semelhante aos estudos prévios voltados para o MedRisk, torna-se pragmática e com viés processual.

Assim, o estado da arte, referente ao processo de análise da estrutura dimensional e métrica (das cargas fatoriais) do instrumento MedRisk, carece de aprofundamento e merece ainda contribuições futuras. A análise apenas dos itens que apresentam correlações com as medidas

globais de satisfação é uma prática processual que deve ser confrontada com a análise do comportamento de todos os itens, independente se estes se correlacionem ou não com as medidas globais de satisfação. Nesta perspectiva, a análise fatorial confirmatória, com dados de indivíduos inseridos em diferentes contextos, partindo do modelo previamente postulado com todos os 18 itens, faz-se necessária. Assim, a rejeição ou confirmação deste modelo original pode contribuir para o histórico psicométrico do instrumento e, conseqüentemente, para o delineamento de um instrumento mais robusto.

As limitações inerentes ao instrumento MedRisk se apresentam, não só quanto aos aspectos metodológicos das estratégias empregadas na sua validação, mas também quanto aos conceituais. Pode-se citar como exemplo a presença da opção de resposta “não aplicável” (NA) no instrumento, que merece maior debate. Esta opção pode, por vezes, mascarar uma dificuldade de entendimento do item. No mesmo sentido, destaca-se o problema dos dados faltantes, já que nenhum estudo prévio do instrumento MedRisk (BEATTIE et al., 2002, 2005, 2007, 2011; HUSH et al., 2011; OLIVEIRA et al, 2013) apresentou a estratégia processual utilizada para tratamento destes tipos de dados. Esse fato dificulta discussões voltadas ao percentual de itens não respondidos ou apenas faltantes entre os estudos voltados para o MedRisk.

Portanto, outros estudos devem ser encorajados tomando o MedRisk como objeto, tendo como base outros perfis populacionais, maior quantidade de indivíduos e contextos distintos. Só assim, o modelo proposto para o instrumento poderá ser posto à prova quanto a sua estabilidade e também quanto à aferição da compreensão dos itens propostos.

Particularmente em países como o Brasil, com grande extensão geográfica, diferenças culturais regionais e um sistema de saúde complexo, um mesmo instrumento pode apresentar dificuldades quanto à compreensão por parte população alvo, e, assim, enviesar os resultados mensurados por este. Os achados deste estudo ilustram a importância da adaptação transcultural de instrumentos, sobretudo daqueles empregados no campo da qualidade do cuidado. Especialmente aqueles que se propõem a mensurar a satisfação de pacientes devem ter seu grau de compreensão, adequação e equivalência transcultural avaliados e testados em diferentes cenários antes de serem considerados prontos para uso em toda população.

Limitações do Estudo

O presente estudo não teve por pretensão apresentar um instrumento validado em todos os aspectos descritos na literatura (TERWEE et al., 2007), muito menos postular um modelo “ideal”

para o instrumento MedRisk. Mesmo avançando na investigação do comportamento transcultural desta ferramenta, algumas dúvidas quanto ao melhor modelo possível ainda persiste.

O estudo apresenta limitações quanto ao desenho, particularmente com relação ao número de observações e aplicação em um único serviço de uma única capital de estado brasileiro. A estratégia adotada para a análise psicométrica também traz importantes limitações, das quais a exclusão de itens com base na correlação com as duas medidas globais de satisfação é o principal problema. Esta abordagem assume que essas medidas seriam semelhantes a um padrão ouro e capazes de aferir completamente a satisfação do paciente; fato questionável. Essa escolha, semelhante ao percurso metodológico de estudo anteriores (BEATTIE et al., 2002, 2007, 2011; OLIVEIRA et al., 2015), levou a inclusão na análise fatorial confirmatória de um instrumento com número de itens reduzido. Outro ponto problemático diz respeito aos achados da consistência interna de ambas as dimensões. Os valores apresentados são levemente abaixo do recomendado, embora os índices de ajuste do modelo encontrados tenham sido satisfatórios. Outras abordagens de aferição da consistência interna devem também ser conjecturadas, como, por exemplo, a confiabilidade composta (BROWN, 2006).

Apesar dos limites, o presente trabalho se propôs a contribuir para a aplicabilidade do instrumento MedRisk no contexto brasileiro e os pressupostos teóricos que abarcam o constructo satisfação do paciente em fisioterapia. Buscou-se ainda destacar a relevância contemporânea do estudo de satisfação para a melhoria da qualidade do cuidado prestado em todos os níveis e especialidades. Por fim, os achados ilustram a importância da adaptação transcultural de instrumentos, sobretudo daqueles empregados no campo da qualidade do cuidado, majoritariamente utilizados a despeito da ausência de informações sobre suas propriedades métricas e sensibilidade ao contexto.

6. CONSIDERAÇÕES FINAIS

Em termo geral, é possível afirmar que o campo da validação de instrumentos utilizados para mensuração da qualidade do cuidado em saúde, embora tenha evoluído muito nos últimos anos, incluindo a disseminação das etapas necessárias e padronização deste processo, ainda carece de esforço conjunto para responder a importantes lacunas. Nesse cenário, instrumentos-chave ou estratégicos poderiam ser privilegiados por distintos grupos de pesquisa e, assim, aprimorados. Esse esforço representaria ainda uma iniciativa para ajustar ferramentas que viabilizassem estudos comparativos e, portanto, contribuir para melhor compreensão de fenômenos tão complexos e interdependentes, como qualidade do cuidado, centralidade do cuidado no paciente e satisfação.

O presente estudo realizou a adaptação transcultural do instrumento MedRisk para o contexto brasileiro e assim buscou contribuir para o fortalecimento e histórico psicométrico deste instrumento. Os principais resultados voltaram-se para a avaliação da equivalência conceitual, de itens e semântica, além da equivalência de mensuração. Mesmo considerando os efeitos indesejados da estratégia de análise definida, a versão proposta, ao final deste processo, apontou um instrumento com duas dimensões, denominadas: “Interpessoal” e “Estrutura e apoio ao serviço”, contendo três itens cada. As análises apresentaram bom ajuste deste modelo, com confiabilidade excelente após teste-reteste, porém a consistência interna mostrou valores moderados. Entretanto, decisões calcadas numa população e contexto de aplicação circunscrito, conjugadas com as etapas de decisões tomadas na análise psicométrica, devem ser examinadas com precaução.

Portanto, a versão aqui proposta deve ser aplicada em outros grupos populacionais e em distintos serviços das diferentes regiões brasileiras, pois, só assim, poderíamos consolidar evidências, pró ou contra, acerca da convicção sobre a qualidade deste instrumento para aferir todos os aspectos relevantes para satisfação do paciente com o serviço de saúde. Essa preocupação advém das diferenças encontradas entre os dois únicos estudos que propuseram versões brasileiras do instrumento, mas também das abordagens para validação desse instrumento realizadas em estudo anteriores. Embora ambos os estudos brasileiros tenham obtido valores semelhantes para satisfação nos respectivos grupos populacionais, há importantes diferenças quanto à equivalência semântica, psicomетria do instrumento e características do delineamento do estudo.

Este estudo buscou também contribuir para as discussões sobre a satisfação do paciente e preocupação com a qualidade do cuidado especificamente no campo da fisioterapia. Particularmente, quanto ao MedRisk, os achados deste e de estudo anteriores, sublinham que um instrumento “validado” para diferentes culturas não suprime a necessidade de ajustes. Esta questão

é central, devendo sempre ser considerada por clínicos e outros usuários dos instrumentos na gestão cotidiana dos serviços, face à decisão de escolha do instrumento a utilizar. A identificação das possíveis limitações de aferição dos instrumentos é essencial para sua utilização, minimizando assim o emprego de medidas carregadas de “inverdades”.

Especificamente no campo da satisfação do paciente, é evidente a relação deste constructo com o conceito de cuidado centrado no paciente, embora satisfação não abarque todas as dimensões do cuidado centrado no usuário – aspecto chave para qualidade do cuidado (IOM, 2001).

Por fim, apesar dos limites deste trabalho, é importante destacar que os achados e a adaptação transcultural do MedRisk para o contexto brasileiro contribuem para o debate voltado para validação de instrumentos de aferição na área da fisioterapia; e espera-se também que o presente estudo ajude a fomentar a discussão sobre a qualidade dos instrumentos de aferição da satisfação do paciente nas demais especialidades da saúde. Dessa forma, espera-se que os achados contribuam para diminuir uma lacuna importante no campo da avaliação e melhoria da qualidade do cuidado, sobretudo no Brasil onde essas iniciativas são comparativamente mais circunscritas e descontínuas.

7 REFERÊNCIAS

- ALMEIDA, C. et al. Health conditions and health-policy innovations in Brazil: the way forward. **Lancet**, v. 377, p. 2042–53, 2011.
- ALMEIDA, R.S.; NOGUEIRA, L.A.C.; BOURLIATAUX-LAJOINE, S. Analysis of the user satisfaction level in a public physical therapy service. **Braz J Phys Ther.**, v. 17, n.4, p. 328-335, 2013.
- ALMEIDA, R.S.; BOURLIATAUX-LAJOINE, S.; MARTINS, M. Instrumentos para mensuração da satisfação dos usuários de serviços de saúde: revisão sistemática. **Cad. Saúde Pública.**, v. 31, n. 1, p.11-25, 2015.
- ARRARAS, J.I. et al. The EORTC cancer in-patient satisfaction with care questionnaire: EORTC IN-PATSAT32. **Clin Trans Onc.**, v. 11, n. 4, p. 237-42, 2009.
- BAKER, T.L.; CRONIN, J.J.; HOPKINS, C.D. The impact of involvement on key service relationships. **J Serv Mark.**, v. 23, n. 2, p. 115–124, 2009.
- BARDES, C.L. Defining “Patient-Centered Medicine”. **N Engl J Med.**, v. 366, n. 9, p. 782-783, 2012.
- BEATTIE, M . et al. Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol. **Syst Rev.**, v. 3, p. 4, 2014.
- BEATTIE, P.F.; NESLSON, R.M.; HEINTZELMAN, M. The relationship between patient satisfaction with physical therapy care and global rating of change reported by patients receiving worker’s compensation. **Phy Theory Pract.**, v. 27, n. 4, p. 310–318, 2011.
- BEATTIE, P.F. et al. Patient Satisfaction With Outpatient Physical Therapy: Instrument Validation. **Phys Ther.**, v. 82, p. 557–565, 2002.
- BEATTIE, P.F. et al. The MedRisk Instrument for Measuring Patient Satisfaction With Physical Therapy Care: A Psychometric Analysis. **J Orthop Sports Phys Ther**, v. 35, p. 24-32, 2005.
- BEATTIE, P.F.; NELSON, R.; MURPHY, D.R. Development and preliminary validation of the medrisk instrument to measure patient satisfaction with chiropractic care. **J Manipulative Physiol Ther.**, v. 34, p. 23-29, 2011.

BEATTIE, P.F.; NELSON, R.M.; Lis, A. Spanish-Language Version of the MedRisk Instrument for Measuring Patient Satisfaction With Physical Therapy Care (MRPS): Preliminary Validation. **Phys Ther**, v. 87, n. 6, p.793-800, 2007.

BEATON, D.E. Et al. Guidelines for the process of cross-cultural adaptation of self-report measures. **Spine**, v. 25, n. 24, p. 3186-9, 2000.

BERWICK, D.M. A User's Manual For The IOM's 'Quality Chasm' Report. **Health Affairs**, v. 2, p. 380-90, 2002.

BLEICH, A.S.; ÖZALTIN, E.; MURRAY, C.J.L. How does satisfaction with the health-care system relate to patients experience. **Bull World Health Organ.**, v. 87, P. 271–278, 2009.

BORGES, J.B.; CARVALHO, S.M.; SILVA, M.A. Quality of service provided to heart surgery patients of the Unified Health System-SUS. Rev Bras Cir Cardiovasc, v. 25, n. 2, p.172-82, 2010.

BOYER, L. et al. Assessment of psychiatric inpatient satisfaction: A systematic review of self-reported instruments. **European Psychiatry**, v. 24, p. 540–549, 2009.

BRÉDART, A. Interviewing to develop Patient-Reported Outcome (PRO) measures for clinical research: eliciting patients' experience. **Health and Quality of Life Outcomes**, v. 12, p. 15, 2014.

CAPPELLERI, J.C. et al. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient reported outcomes measures. **Clin. Ther.**, v. 36, n.5, p. 648-662, 2014.

CHANG J. T. et al. Patients' Global Ratings of Their Health Care Are Not Associated with the Technical Quality of Their Care. **Ann Intern Med.**, v. 144, p. 665-672, 2006.

COUTLER, A. Patient engagement-what works? **J Ambul Care Manage.**, v. 35, n. 2, p. 80-89, 2012.

CURRY, A. ; Sinclair, M. Assessing the quality of physiotherapy service using Servqual. **Int J Health Care Qual Assu.**, Reino Unido, v.15, n.5, p.197-205, 2002.

DANIELSEN, K.S. et al. The association between demographic factors, user reported experiences and user satisfaction: Results from three casualty clinics in Norway. **BMC Family Practice**, Londres, v.11, p. 73, 2010.

DIERCKX, K. et al. Implementation of Shared Decision Making in Physical Therapy: Observed Level of Involvement and Patient Preference. **Phys Ther.**, v. 93, p. 1321-1330, 2013.

DONABEDIAN, A. The definition of quality and approaches to its assessment. In:_____. **Explorations in quality assessment and monitoring**. Vol.I. Ann Arbor, MI: Health Administration Press, 1980.

DONABEDIAN, A The quality of care: How can it be assessed? **JAMA**, v. 260, p.1743–1748, 1988.

DOYLE, C.; LENNOX, L.; Bell, D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. **BMJ Open**, v. 3, 2013.

ELLEUCH, A. Patient satisfaction in Japan. **Inter J Health Care Qual Assur.**, v.21, n. 7, p. 692-705, 2008.

EPSTEIN, R.M.; STREET, J.R. The Values and Value of Patient-Centered Care. **Ann Fam Med**, Cleveland, v.9, p.100-103, 2009.

EVANS, J. et al. VOICE: developing a new measure of service users' perceptions of inpatient care, using a participatory methodology. **J Ment Health** , v.21, n. 1, p. 57-71, 2012.

FEENEY, S.N.C. Patient satisfaction with private physiotherapy for musculoskeletal pain. **BMC Musculoskeletal Disorders**, v.9, p. 50-62, 2008.

FENTON, J.J. The Cost of Satisfaction: A National Study of Patient Satisfaction, Health Care Utilization, Expenditures, and Mortality. **Arch Intern Med.**, v.172, n. 5, p. :405-41, 2012.

FELDMAN, L.B. et al. História da Evolução da Qualidade Hospitalar: dos padrões a acreditação. **Acta Paul Enferm**, São Paulo, v. 18, n.2, p. 213- 219, 2005.

FITZPATRICK, R. Satisfaction with health care. In: **The Experience of Illness**, ed. R. Fitzpatrick, 1984: 154-175; Tavistock, London.

SISTEMA DE INFORMAÇÃO AMBULATORIAL OU HOSPITALAR. SURCA/SMSDC RIO, 2014. Disponível em: www0.rio.rj.gov.br/ufs/tabnet. Acesso em 23/01/2015.

GERKENSMEYER, J.E., AUSTIN, J.K. Development and testing of a scale measuring parent satisfaction with staff interactions. **J Behav Health Serv Res.**, v.32, n.1. p.61-73, 2005.

GOLDSTEIN, M.S., ELLIOTT, S.D., GUCCIONE, A.A. The development of an instrument to measure satisfaction with physical therapy. **Phys Ther**, v. 80, n.9, p.853–863, 2000.

GÜLER, N., UYANIK, G.K., TEKER, G.T. Comparison of classical test theory and item response theory in terms of item parameters. **Eur J Res Edu**, Oxford, v. 2, n.1, p.1-6, 2014.

GROENE, O. Patient centredness and quality improvement efforts in hospitals: rationale, measurement, implementation. **Int J Qual Health Care.**, v. 23, n.5, p.531–537, 2011.

HAIR, J.F. et al., **Multivariate data analysis: with readings**. 3rd ed. Nove Iorque. Ed. Macmillan Publishing Co.; 1987.

HALL, J.A.; DORNAN, M.C. Meta-analysis of satisfaction with medical care: description of research domain and analysis of overall satisfaction levels. **Soc Sci Med.**, Londres, v.27, p. 637–44, 1988.

HEKKERT, K.D. et al. Patient satisfaction revisited: A multilevel approach. **Soc Sci Med.**, Londres, v. 69, p. 68–75, 2009.

HENSON, R.K. Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha. **Meas eval couns devel**, Los Angeles, v. 34, p. 177-189, 2001.

HERCOS, S.B.V., BEREZOVSKY, A. Ophthalmological service quality offered to outpatients of the Public Healthcare System. Arq Bras Oftalmol, São Paulo, v. 69, n. 2, p. 213-9, 2006.

HERDMAN, M.; FOX-RUSHBY, J.; BADIA, X. “Equivalence” and the translation and adaptation of health-related quality of life questionnaires. **Int J Qual Health Care** v. 6; n, 3; p. 237-47, 1997.

HERDMAN M, FOX-RUSHBY J, BADIA X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. **Qual Life Res.**, v. 7, n. 4, p. 323-35, 1998.

HILLS, R.; KITCHEN, S. Satisfaction with outpatient physiotherapy: Focus groups to explore the views of patients with acute and chronic musculoskeletal conditions. **Phys Theory Pract**, v. 23, n. 1, p.1-20, 2007.

HOLMSTROM,I.; ROING, M. The relation between patient-centeredness and patient empowerment: A discussion on concepts. **Patient Edu Couns**, v.79, p.167–172, 2010.

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

HUDAK, P.L. et al. The metaphor of “patient as customer”: implications for measuring patient satisfaction. **J Clin Epidemiol**, v. 52, p.485, 1999.

HUDAK, P.L., WRIGHT, J.G. The Characteristics of patient satisfaction Measures. **Spine**, v.25, n. 24, p. 3167-3177, 2000.

HUDON, C.; Measuring Patients’ Perceptions of Patient-Centered Care: A Systematic Review of Tools for Family Medicine. **Ann Fam Med**, v.9, p.155-164, 2011.

HUSH, J.M. et al. Intercultural comparison of patient satisfaction with physiotherapy care in Australia and Korea: an exploratory factor analysis. **J Man Manip Ther.**, v. 21, p.103-112, 2013.

HUSH, J.M. et al. Patient satisfaction with musculoskeletal physiotherapy care in Australia: an international comparison. **J Man Manipul Ther**, v. 20, n.4, p. 201-208, 2012.

HUSH, J.M.; CAMERON, K.; MACKEY, M. Patient satisfaction with musculoskeletal physical therapy care: a systematic review. **Phys Ther**, v.91, n.1, p. 25-36, 2011.

INSTITUTE OF MEDICINE (IOM). **Committee on quality of health care in America**. Crossing the quality chasm: a new health system for the 21st century. National Academy press. Washington, 2001.

KAMPER, S.J. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. **J Clin Epidemiol**, v. 63, n.7, p.760-766, 2010.

KENNEDY, G.D. et al. Is There a Relationship Between Patient Satisfaction and Favorable Outcomes? **Ann Surg.**, Philadelphia, V.260, P. 592–600, 2014.

KHADKA, J. The Importance of Rating Scale Design in the Measurement of Patient-Reported Outcomes Using Questionnaires or Item Banks. **Inv Ophthal Vis Sci**, v.53, n.7, p. 4041-4054, 2012.

KNIGHT, P.K. et al. Results of a survey of client satisfaction with outpatient physiotherapy care. **Phys Theory Pract**, v. 26, n.5, p.297–307, 2010.

MACHADO, N.P.; NOGUEIRA, L.P. Evaluation of physical therapy service user satisfaction. **Braz J Phys.**, v.12, n.5, p.401-8, 2008.

MACHADO, J.P., MARTINS, A.C.M., MARTINS, M. Avaliação da qualidade do cuidado hospitalar no Brasil: uma revisão sistemática. **Cad. Saúde Pública** V. 29, n.6: 1063-1082, 2013.

MENDES W.; NORONHA, J.C; TRAVASSOS, C.; MARTINS, M.; LAJOLO, C.; CALDAS, B.; PORTELA, M.C.; GRABOIS; V., GOMES, A.Q.; PEREIRA, T.R. Qualidade de Serviços de Saúde no SUS (Qualisus): Uma contribuição para a gestão da qualidade da atenção à saúde do SUS. 2013. (Relatório de pesquisa), apoio Ministério da Saúde, disponível em: <http://proqualis.net/>.

MENDONÇA, K.M.P.; GUERRA, R.O. Desenvolvimento e validação de um instrumento de medida da satisfação do paciente com a fisioterapia. **Braz J Phys** v.11, n.5, p.369-376, 2007.

MENDONÇA, K.M.P.P. Satisfação do paciente com a fisioterapia: tradução, adaptação cultural e validação de um instrumento de medida [dissertação]. Natal (RN): UFRN; 2004.

MEYER, T. et al. Problems completing questionnaires on health status in medical rehabilitation patients. **J Rehabil Med**, v. 39, p. 633–639, 2007.

MILUTINOVIC', D. et al. The patient satisfaction with nursing care quality: The psychometric study of the Serbian version of PSNCQ questionnaire. **Scan J Caring Sci.**, v. 26, n. 3, p.598-606, 2012.

MOKKINK, L.B. et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. **BMC Med Res Method**, v.10, p.22, 2010.

MONNIN, D.; PERNEGER, G.T. Scale to measure patient satisfaction with physical therapy. **Phys Ther**, v.82, n.7, p.682-691, 2002.

MOREIRA, C.F. et al. Instrumento para aferir a satisfação do paciente com a assistência fisioterapêutica na rede pública de saúde. **Fisiot Pesq.**, v.14, n.3, p. 37-43, 2007.

NORONHA, J.C.; ROSA, M.L.G. Quality of health care – growing awareness in Brazil. **Int J Qual Health Care**, v.11, n.5, p.437-441, 1999.

OLIVEIRA, N.F.C. et al. Measurement properties of the brazilian portuguese version of the MedRisk instrument for measuring patient satisfaction with physical therapy care. **J Orthop Sports Phys Ther.**, v. 44, n.11, p. 879-889, 2014.

OLIVER, R. L. A conceptual model of service quality and service satisfaction: compatible goals, different concepts. **Adv Serv Mark Manag**, v.2, p.65-85, 1993

OTTONELLO, M. et al. Patient satisfaction with hospital rehabilitation: Validation of the SAT-16 questionnaire through Rasch analysis. **Min Med.**, v.103, n.1,p.1-11, 2012.

PAIM, J. et al. The Brazilian health system: history, advances, and challenges. **Lancet**, v.377, p.1778–97, 2011.

PAIXÃO, C.M. et al. Adaptação transcultural para o Brasil do instrumento Caregiver Abuse Screen (CASE) para detecção de violência de cuidadores contra idosos. **Cad. Saúde Pública.**, v. 23, n. 9, p. 2013-2022, 2007.

PARASURAMAN, A. et al. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. **J Retailing**, v. 64, n.1, 1988.

PASCOE, G.C. Patient satisfaction in primary health care: a literature review and analysis. **Eval Prog Plan.**, v. 6, p.185-210, 1983.

POINSOT. R. et al. Multisite validation study of questionnaire assessing out-patient satisfaction with care questionnaire in ambulatory chemotherapy or radiotherapy treatment. **Bull Cancer**, v.93, n.3, p. 315-27, 2006.

QUIGLEY, H.L.et al. Assuring the quality of health care in the European Union: a case for action. **World Health Organization - Observatory Studies Series**, n.12, 2008.

RAHMQVIST, M.; BARA, A.C. Patient characteristics and quality dimensions related to patient satisfaction. **Int J Qual Health Care**, v. 22, n. 2, p. 86–92, 2010.

RASHID, W.E.; JUSOFF, K. Service quality in health care setting. **Int J Health Care Qual Assur.**, v.22, n.5, p. 471-482, 2009.

REICHENHEIM, M.E.; MORAES, C.L. Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. **Rev Saúde Pública.**, v. 41, n.4, p. 665-73, 2007.

REICHENHEIM, M.E. et al. Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): empirical evidence for a general factor. **BMC Med Res Method.**, v. 11, p. 93, 2011.

REICHENHEIM, M.E. ; PAIXÃO, C.M.; MORAES, C.L. Reassessing the construct validity of a Brazilian version of the instrument Caregiver Abuse Screen (CASE) used to identify risk of domestic violence against the elderly. **J Epidemiol Community Health.**, v.63, p.878–883, 2009.

REICHENHEIM, M.E.; HÖKERBERG, Y.H.M.; MORAES, C.L. Assessing construct structural validity of epidemiological measurement tools: a seven-step roadmap. **Cad. Saúde Pública**, v. 30, n.5, p. 927-939, 2014.

REIS, C.T.; LAGUARDIA, J.; MARTINS, M. Adaptação transcultural da versão brasileira do Hospital Survey on Patient Safety Culture: etapa inicial. **Cad Saúde Pública**, v. 28, n.11, p. 2199-2210, 2012.

ROCHA, L.R.M. et al. Health service quality scale: Brazilian Portuguese translation, reliability and validity. **BMC Health Serv Res.**, v.13, p. 24, 2013.

ROSENZVEIG, A. Toward Patient-Centered Care: A Systematic Review of How to Ask Questions That Matter to Patients. **Medicine**, v. 93, n. 22, p. 120, 2014.

ROUSH, S.E.; SONSTROEM, R.J. Development of the Physical Therapy Outpatient Satisfaction Survey (PTOPS). **Phys Ther**, v.79, p.159-170, 1999.

ROZENBLUM, R. et al. Uncovering the blind spot of patient satisfaction: an international survey. **BMJ Qual Saf.**, v. 20, p.959-965, 2011.

ROZENBLUM, R. et al. The patient satisfaction chasm: the gap between hospital management and frontline clinicians. **Qual Saf Health Care.**, v.22, p. 242–250, 2013.

SCHRAMM, J.M. et al. Transição epidemiológica e o estudo de carga de doença no Brasil. **Ciência Saúde Col.**, v. 9, n. 4, p. 897-908, 2004.

SCHREIBER, J.B. et al. Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. **The Journal Edu Res.**, v.99, n.6, p.323-337, 2006.

SCHOENFELDER, T. et al. Determinants of patient satisfaction: a study among 39 hospitals in an in-patient setting in Germany. **Int J Qual Health Care.**, v. 23. n. 5, p. 503–509, 2011.

SHAW, C. et al. Accreditation and ISO certification: do they explain differences in quality management in European hospitals? **Int J Qual Health Care.**, v.22, n. 6, p. 445–45, 2010.

SHEA, J.A. et al. Adapting a patient satisfaction instrument for low literate and Spanish-speaking populations: Comparison of three formats. **Patient Edu Coun.**, v. 73, n.1,p.132-40, 2008.

SITZIA, J.; WOOD, N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.*, v. 45, n.12, p.1829-1843, 1997.

SOFAR, S. ; FIRMINGER, K. Patient perceptions of the quality of health services. **Annu Rev Public Health.**, v. 26, p. 513–59, 2005.

TENNANT, A.; CONAGHAN, P.G. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? **Arthritis Rheum.**, v.57, p.1358-1362, 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. **Qual Life Res.** v. 21, p. 651–657, 2011.

TERWEE, C.B. et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. **Qual Life Res.**, v.21, n 4, p.651–657, 2012.

TERWEE CB. et al. Quality criteria were proposed for measurement properties of health status questionnaires. **J Clin Epidemiol.**, v. 60, p. 34-42, 2007.

THOMAS, ML. The value of item response theory in clinical assessment: a review. **Assessment.**, v.18, n.3, p. 291-307, 2011.

TRAVASSOS, C; CALDAS, B. A qualidade do cuidado e a segurança do paciente: histórico e conceitos. In: (Ed.). **Assistência segura: uma reflexão teórica aplicada à prática.** Brasília: ANVISA, 2013.

TZEPELIS, F. et al. Are we missing the Institute of Medicine's mark? A systematic review of patient-reported outcome measures assessing quality of patient-centred cancer care. **BMC Cancer**, v. 15, p. 41, 2014.

VANTI, C. et al. Italian version of the physical therapy patient satisfaction questionnaire: cross-cultural adaptation and psychometric properties. **Phys Ther.**, v.93, n.7, p. 911-22, 2013.

VET, H.C. et al. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. **Health Qual Life Out.**, v. 4, n. 54, 2006.

VET, H.C. et al. When to use agreement versus reliability measures. **J Clinical Epidemiol.**, v.59, p.1033–1039, 2006.

VICTORA, C.G. et al. Condições de saúde e inovações nas políticas de saúde no Brasil: o caminho a percorrer. **Lancet** (Série Brasil), v.90, p.102, 2011.

VINAGRE, M.H.; NEVES, J. The influence of service quality and patients' emotions on satisfaction. *Int J Health Care Qual Assur.*, v.21. n.1, p. 87-103, 2008.

WARE, J.E. et al. The measurement and meaning of patient Satisfaction: a review of the literature. **The Rand paper series**, p. 6036, 1977.

WILDE, E.F.; HENDRIKS, V.M. The Client Satisfaction Questionnaire: psychometric properties in a Dutch addict population. **Eur Addict Res.**, v.11, n.4, p.157-62, 2005.

WILLIAMS, B. Patient satisfaction: a valid concept? **Soc Sci Med.**, v. 38, n.4, p. 509-516, 1994.

WILLIS, G.B.; ARTINO, A.R. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. **J Graduate Med Educ.**, v.5, n.3, p. 353-356, 2013.

ZILL, J.M. et al. " Measurement of Physician-Patient Communication - A Systematic Review. *PLoS ONE*, v.9, n.12, p. 126-37, 2014.

ANEXO A – Critérios para análise da qualidade metodológica dos instrumentos proposta pelo Grupo COSMIN

COSMIN checklist with 4-point scale

Contact

CB Terwee, PhD
VU University Medical Center
Department of Epidemiology and Biostatistics
EMGO Institute for Health and Care Research
1081 BT Amsterdam
The Netherlands
Website: www.cosmin.nl, www.emgo.nl
E-mail: cb.terwee@vumc.nl



Instructions

This version of the COSMIN checklist is recommended for use in systematic reviews of measurement properties. With this version it is possible to calculate overall methodological quality scores per study on a measurement property. A methodological quality score per box is obtained by taking the lowest rating of any item in a box ('worse score counts'). For example, if for a reliability study one item in the box 'Reliability' is scored poor, the methodological quality of that reliability study is rated as poor. The Interpretability box and the Generalizability box are mainly used as data extraction forms. We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box (e.g. norm scores, floor-ceiling effects, minimal important change) of the instruments under study from the included articles. Similar, we recommend to use the Generalizability box to extract data on the characteristics of the study population and sampling procedure. Therefore no scoring system was developed for these boxes.

Step 1. Evaluated measurement properties in the article

	Internal consistency	Box A
	Reliability	Box B
	Measurement error	Box C
	Content validity	Box D
	Structural validity	Box E
	Hypotheses testing	Box F
	Cross-cultural validity	Box G
	Criterion validity	Box H
	Responsiveness	Box I

Box A. Internal consistency				
	excellent	good	fair	poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model? <i>Design requirements</i>				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another study
6 Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥ 100	5* #items and ≥ 100 OR 6-7* #items but < 100	5* #items but < 100	< 5 * #items

7	Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	Internal consistency statistic calculated for each subscale separately			Internal consistency statistic NOT calculated for each subscale separately
8	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
9	for Classical Test Theory (CTT), continuous scores: Was Cronbach's alpha calculated?	Cronbach's alpha calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated
10	for CTT, dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item-total correlations calculated
11	for IRT: Was a goodness of fit statistic at a global level calculated? E.g. χ^2 , reliability coefficient of estimated latent trait value (index of (subject or item) separation)	Goodness of fit statistic at a global level calculated			Goodness of fit statistic at a global level NOT calculated

Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)				
	excellent	good	fair	poor
<i>Design requirements</i>				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
4 Were at least two measurements available?	At least two measurements			Only one measurement
5 Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6 Was the time interval stated?	Time interval stated		Time interval NOT stated	
7 Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8 Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
11	for continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated
12	for dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			Only percentage agreement calculated
13	for ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated	Only percentage agreement calculated
14	for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described		

Box C. Measurement error: absolute measures				
	excellent	good	fair	poor
<i>Design requirements</i>				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
4 Were at least two measurements available?	At least two measurements			Only one measurement
5 Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6 Was the time interval stated?	Time interval stated		Time interval NOT stated	
7 Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8 Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
11	for CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population

Box D. Content validity (including face validity)					
		excellent	good	fair	poor
General requirements					
1	Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Assessed if all items refer to relevant aspects of the construct to be measured		Aspects of the construct to be measured poorly described AND this was not taken into consideration	NOT assessed if all items refer to relevant aspects of the construct to be measured

2	Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	Assessed if all items are relevant for the study population in adequate sample size (≥ 10)	Assessed if all items are relevant for the study population in moderate sample size (5-9)	Assessed if all items are relevant for the study population in small sample size (< 5)	NOT assessed if all items are relevant for the study population OR target population not involved
3	Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Assessed if all items are relevant for the purpose of the application	Purpose of the instrument was not described but assumed	NOT assessed if all items are relevant for the purpose of the application	
4	Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Assessed if all items together comprehensively reflect the construct to be measured		No theoretical foundation of the construct and this was not taken into consideration	NOT assessed if all items together comprehensively reflect the construct to be measured
5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Box E. Structural validity				
	excellent	good	fair	poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the analysis adequate?	7* #items and ≥ 100	5* #items and ≥ 100 OR 5-7* #items but < 100	5* #items but < 100	$< 5^*$ #items
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. rotation method not described)	Other important methodological flaws in the design or execution of the study (e.g. inappropriate rotation method)
Statistical methods				
6 for CTT: Was exploratory or confirmatory factor analysis performed?	Exploratory or confirmatory factor analysis performed and type of factor analysis appropriate in view of existing information	Exploratory factor analysis performed while confirmatory would have been more appropriate		No exploratory or confirmatory factor analysis performed
7 for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	IRT test for determining (uni)dimensionality performed		IRT test for determining (uni)dimensionality NOT performed	

Box F. Hypotheses testing				
	excellent	good	fair	Poor
<i>Design requirements</i>				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥ 100 per analysis)	Good sample size (50-99 per analysis)	Moderate sample size (30-49 per analysis)	Small sample size (< 30 per analysis)

4	Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?	Multiple hypotheses formulated a priori	Minimal number of hypotheses formulate a priori	Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
5	Was the expected <i>direction</i> of correlations or mean differences included in the hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
6	Was the expected absolute or relative <i>magnitude</i> of correlations or mean differences included in the hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		
7	for convergent validity: Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Adequate description of most of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
8	for convergent validity: Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population	No information on the measurement properties of the comparator instrument(s)

9 Were there any important flaws in the design or methods of the study? <i>Statistical methods</i>	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	
10 Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Assumable that statistical methods were appropriate, e.g. Pearson correlations applied, but distribution of scores or mean (SD) not presented	Statistical methods applied NOT optimal Statistical methods applied NOT appropriate

Box G. Cross-cultural validity				
	excellent	good	fair	poor
<i>Design requirements</i>				
1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	

3	Was the sample size included in the analysis adequate?	CTT: 7* #items and ≥ 100 IRT: ≥ 200 per group	CTT: 5* #items and ≥ 100 OR 5-7* #items but < 100 IRT: ≥ 200 in 1 group and 100-199 in 1 group	CTT: 5* #items but < 100 IRT: 100-199 per group	CTT: < 5 * #items IRT: (< 100 in 1 or both groups
4	Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known
5	Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, expertise in the construct to be measured, expertise in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described	
6	Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent
7	Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation
8	Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were resolved	Poorly or NOT described how differences between translators were resolved		

9	Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a committee		
10	Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension?	Translated instrument pre-tested in the target population	Translated instrument pre-tested, but unclear if this was done in the target population	Translated instrument pre-tested, but NOT in the target population	Translated instrument NOT pre-tested
11	Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described	
12	Were the samples similar for all characteristics except language and/or cultural background?	Shown that samples were similar for all characteristics except language /culture	Stated (but not shown) that samples were similar for all characteristics except language /culture	Unclear whether samples were similar for all characteristics except language /culture	Samples were NOT similar for all characteristics except language /culture
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Statistical methods		
14 for CTT: Was confirmatory factor analysis performed?	Multiple-group confirmatory factor analysis performed	Multiple-group confirmatory factor analysis NOT performed
15 for IRT: Was differential item function (DIF) between language groups assessed?	DIF between language groups assessed	DIF between language groups NOT assessed

Box H. Criterion validity					
		excellent	good	fair	poor
<i>Design requirements</i>					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
4	Can the criterion used or employed be considered as a reasonable 'gold standard'?	Criterion used can be considered an adequate 'gold standard' (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'

5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>				
6	for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated		Correlations or AUC NOT calculated
7	for dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated		Sensitivity and specificity NOT calculated

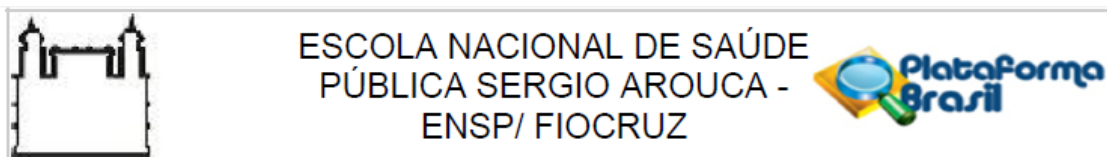
Box I. Responsiveness					
		excellent	good	fair	poor
<i>Design requirements</i>					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Was a longitudinal design with at least two measurement used?	Longitudinal design used			No longitudinal design used
5	Was the time interval stated?	Time interval adequately described			Time interval NOT described

6	If anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Anything that occurred during the interim period (e.g. treatment) adequately described	Assumable what occurred during the interim period	Unclear or NOT described what occurred during the interim period
7	Was a proportion of the patients changed (i.e. improvement or deterioration)?	Part of the patients were changed (evidence provided)	NO evidence provided, but assumable that part of the patients were changed	Unclear if part of the patients were changed Patients were NOT changed
<i>Design requirements for hypotheses testing</i>				
For constructs for which a gold standard was not available:				
8	Were hypotheses about changes in scores formulated a priori (i.e. before data collection)?	Hypotheses formulated a priori		Hypotheses vague or not formulated but possible to deduce what was expected Unclear what was expected
9	Was the expected <i>direction</i> of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated	
10	Were the expected absolute or relative <i>magnitude</i> of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated	

11	Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
12	Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>				
14	Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

<i>Design requirement for comparison to a gold standard</i>					
For constructs for which a gold standard was available:					
15	Can the criterion for change be considered as a reasonable gold standard?	Criterion used can be considered an adequate 'gold standard' (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'
16	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
17	for continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?	Correlations or Area under the ROC Curve (AUC) calculated			Correlations or AUC NOT calculated
18	for dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated

ANEXO B – Parecer de aprovação do estudo pelo Comitê de Ética em Pesquisa FIOCRUZ



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Satisfação de pacientes: validação de instrumento de mensuração e análise das variáveis explicativas do construto

Pesquisador: Renato Santos de Almeida

Área Temática:

Versão: 2

CAAE: 18565913.6.0000.5240

Instituição Proponente: FUNDACAO OSWALDO CRUZ

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 427.839

Data da Relatoria: 23/10/2013

Apresentação do Projeto:

Descrito no parecer anterior.

Objetivo da Pesquisa:

Descrito no parecer anterior.

Avaliação dos Riscos e Benefícios:

Descritos no parecer anterior.

Comentários e Considerações sobre a Pesquisa:

Descritos no parecer anterior.

Considerações sobre os Termos de apresentação obrigatória:

Descritos no parecer anterior.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

Pendências atendidas.

Situação do Parecer:

Aprovado

ANEXO C – Autorização para uso do instrumento MedRisk



Thank you for your interest in MedRisk Instrument for Measuring Physical Therapy Patient Satisfaction. As you know, this 12-question survey has been validated in both English and Spanish and is currently being used in the Netherlands. Obviously this tool holds value from a research stand-point; it also offers value for facility managers and/or owners interested in the ongoing development and improvement of a private practice.

For Research Purposes:

If your interest is in research, the MedRisk Instrument for Measuring Physical Therapy Patient Satisfaction can be licensed without charge under the following conditions:

- 1.) If the instrument will be used for research purposes
- 2.) And if the outcomes data will be reported to Expert Clinical Benchmarks, LLC

For Facility Benchmarking:

If you are interested in using this tool to benchmark satisfaction at your facility we would very much like to contact you for further details. We are currently in the research and development phase of creating an online survey tool with automated reporting to help facility owners and therapists benchmark and improve patient satisfaction.

Please let us know which area you are interested in so we may better serve you:

☒ **I am interested in the tool for Research Purposes;** please forward me a licensing contract and contact me with any further requirements:

Name: Renato Santos de Almeida
 Address: Barão do Bom Retiro St., 2720/203 - Grajaú, Rio de Janeiro, Brazil
 Phone: 552132174435
 Fax:
 Email: renato.fisio@gmail.com

You may fax or email this information to:

Katie O'Brien, Marketing Assistant
 Expert Clinical Benchmarks, LLC
 2701 Renaissance Blvd., Suite 200
 King of Prussia, PA 19406
 Fax: 610-768-5812
kobrien@medrisknet.com

Sincerely,
 Roger Nelson PT, PhD, FAPTA
 Vice President
 Expert Clinical Benchmarks, LLC

APÊNDICE A - Artigo publicado: revisão sistemática sobre instrumentos para mensuração da satisfação do paciente

SATISFACTION MEASUREMENT INSTRUMENTS FOR HEALTHCARE SERVICE USERS: A SYSTEMATIC REVIEW.

Renato Santos de Almeida; Stephane Bourliataux-Lajoie; Mônica Martins.
Cadernos de Saúde Pública, Rio de Janeiro, 31(1):11-25, jan, 2015.

ABSTRACT

Patient satisfaction surveys can be an interesting way to improve quality and discuss the concept of patient-centered care. This study aimed to conduct a systematic review of the validated patient satisfaction measurement instruments applied in healthcare. The systematic review searched the MEDLINE/PubMed, LILACS, SciELO, Scopus and Web of Knowledge. The search strategy used the terms: “Patient Satisfaction” AND “Patient centered care” AND “Healthcare survey OR Satisfaction questionnaire” AND “Psychometric properties”. 37 studies were included and almost all studies showed that satisfaction is a multidimensional construct. In these studies, 34 different instruments were used and most surveys contained the dimension patient-healthcare professional interactions, physical environment and management process. The COSMIN score for methodological quality showed that most of them scored a good or fair average. We can conclude that there is not a gold standard instrument for patient satisfaction assessment but some dimensions are essential for this construct.

Background

In the past decade, several programs for quality improvement were developed in many health services and systems^{1,2}. Leatherman *et al*² reported that there are publications on this area in various world regions, such as India, South America, Asia and North America. More specifically, on the topic of how to increase quality, many institutions have been trying to establish guidelines and parameters for quality assessment and improvement.

The patient-centered care has been accepted as one of the pillars for health services achieve quality on care³. This issue must involve the patients' needs and preferences, as well as previous experience within these services, thus, the literature shows a trend towards studying what actually is relevant for patient's experience within healthcare^{3,4}.

Considering this goal, patient satisfaction studies can be an useful way of measuring a part of this complex construct, because it is an important concern related to quality of healthcare and patient centeredness, both in public as private sectors⁵. The concept of satisfaction is well known and applied in management and especially in marketing. Evrard⁶ defines satisfaction as a psychological state, appearing after consumption (of a service) and compared to an initial baseline (services' expectation). Olivier⁷, Fournier and Mick⁸ expressed satisfaction as the subjective comparison between expectations and perceptions of service performance. The satisfaction is subjective, yet measurable, provided that the consumer has an initial point of reference. Satisfaction on healthcare must consider the "emotional" aspect related to a medical action that may have a significant impact on patients' own health. Thereby, understanding all dimensions and factors related to patient satisfaction is quite important to improve quality of healthcare^{6,7,8}.

Recent publications on quality of care have discussed the challenge to improve quality⁹ and most of them have stressed the patient-centered care as one of the core goals^{10,11}. In this context, the user satisfaction construct has already been scrutinized by many different study designs in different areas^{12,13,14}. Surveys' validation and cross-cultural adaptation are the most common study designs, but, until these days, this area needs more investigation^{15,16}.

Choosing the appropriate instrument for patient-related outcome measurement is difficult; it depends on the proposed use and probably can impact the accuracy of the results. The discussion about which are the essentials psychometric properties for healthcare instruments validation has still a lack in the literature^{14,15,16,17}. So, is quite important to know if is there some reference instrument to measure patient satisfaction with healthcare service and if is there a similar scope in the structure of these instruments after the validation process.

This study aimed to conduct a systematic review on the validated patients satisfaction measurement instruments applied in healthcare services; identifying the dimensions analyzed by satisfaction surveys, the quality and psychometric properties of the instruments found.

Methods

Study design and information sources

A systematic review on patient satisfaction surveys was undertaken. Electronic databases were chosen based on their indexed journals profile. Then, the Medline/Pubmed, Lilacs (Virtual Health Library – Bireme PAHO), Scielo ([Scientific Electronic Library Online](#)), Scopus and Web of Knowledge were used as reference material in this study. The Lilacs and Scielo were included due the fact that both have Latin-American studies in their collections. All databases were screened in December, 2013. The search strategy was oriented only for validation studies of patient satisfaction instruments and for this reason other terms that could be relevant for quality of care or other pillars of this complex construct were not included. So terms that could have other concepts inputted were not included as: patient preference, patient' needs and consumer satisfaction. So, the search strategy used the terms: “Patient Satisfaction[MeSH Terms]” AND “Patient centered care[MeSH Terms]” AND “Healthcare survey OR Satisfaction questionnaire” AND “Psychometric properties”.

Inclusion and Exclusion Criteria

The research assumed the following eligibility criteria: (a) only studies published in the English, Spanish, French and Portuguese languages; (b) studies published between 2002 and 2013; (c) studies which have used a patient satisfaction instrument to investigate only the opinion of health service users ; (d) studies whose measurement properties were evaluated and described. No types of health service or medical specialty were excluded to guarantee different patient profiles. The definition of health service used was “any center that provides diagnosis, treatment or prevention for more than one kind of disease, illness or injury in the human body”. However, studies that used instruments to measure patients' satisfaction regarding one specific treatment were excluded, as diabetes or chemotherapy, for instance. Systematic review and metanalysis studies were excluded.

Study Selection

Two independent revisers applied the inclusion and exclusion criterion on the articles found. It was screened all titles and abstracts after the duplicated removed and based on this examination, a list of references for full text reading was prepared. After this, both revisers selected the papers to be included in the next step (qualitative synthesis of the studies). The search results and papers screened were discussed by e-mail and also during some meetings. Any disagreements on the records included in this review were submitted to a third reviser.

Data Extraction

A form was prepared to collect wished information; the following items were extracted from each study: year, authors, study design (Instrument validation, adaptation - changes suggested from the original, and cross-cultural adaptation), healthcare setting (hospital, clinic), satisfaction instrument utilized, dimensions of the instrument and psychometric properties.

Study Evaluation of the Methodological quality

It is necessary to emphasize that the methodological quality of the study is as important as the quality of a measurement instrument. The psychometric properties of an instrument are merely one part of the validation process and, thus, other steps must be demonstrated during this process^{14,15}. The COSMIM checklist is one of the guidelines to evaluate the methodological quality of measurement instruments in different areas, and many systematic reviews have already proved its usefulness^{17,18,19}. Considering this, the present study used the COSMIM checklist to evaluate the methodological quality of the studies included in the reviewing process. This checklist is a standardized tool to assess studies on measurement properties. It contains 9 boxes and each one has a different measurement property, with 5–18 items per box on design aspects and statistical methods. The boxes concern: internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, and responsiveness¹⁷.

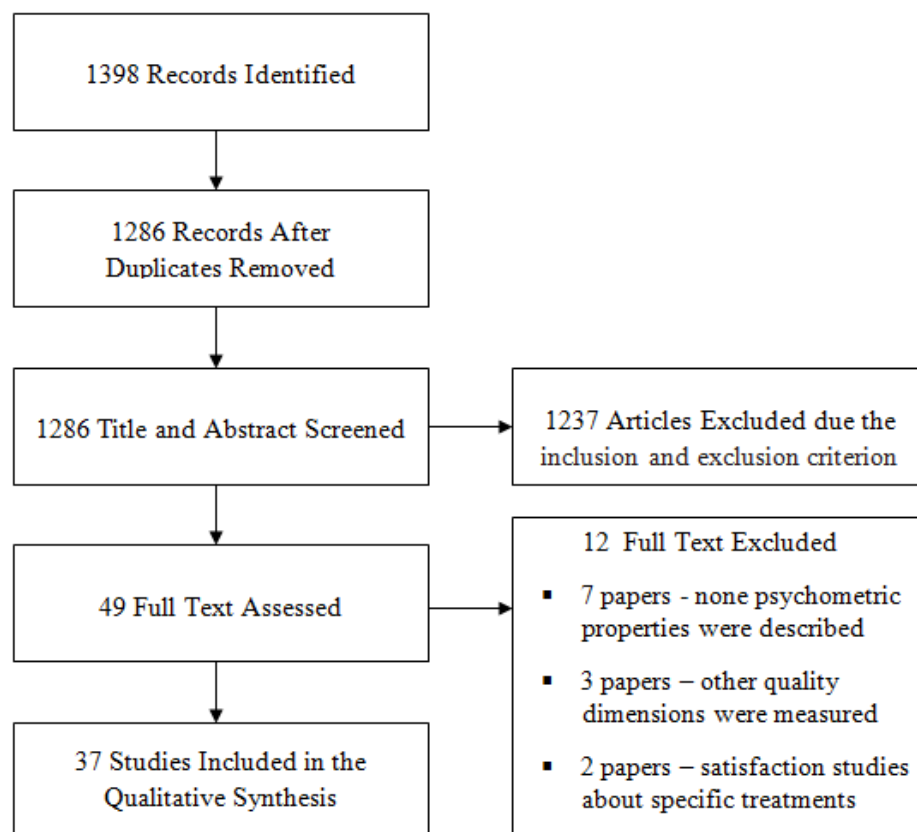
Each measurement property is evaluated separately. This means that if multiple measurement properties are assessed in one study, several COSMIN boxes need to be completed. The content evaluated in each box can be judged as: “excellent”; “good”; “fair”; “poor”. The methodological quality score per box is obtained by taking the lowest rating of any item in the box (‘worse score counts’)¹⁷. The checklist includes one additional box for

studies that used item response theory (IRT). Two additional boxes are recommended (interpretability and generalizability of results), but, as they do not provide a score, they were not used here.

Results

The initial search strategy found 1,398 articles, and after the duplicated ones excluded, the total was 1,286. After the inclusion and exclusion criteria were applied, there were only 49 articles remaining. All these studies were retrieved for reading and twelve were excluded in this phase. Seven studies were excluded because only a simple questionnaire application was done, without description of any measurement properties of these instruments. Other three studies were rejected for being mainly aimed at other quality dimension and measures. Finally, two articles were excluded because the setting was an extremely specialized health service (supplying just one kind of treatment). By the end of this phase, 37 articles were included in the qualitative synthesis of the reviewing process (Figure 1).

Figure 1: Flow of information through the different phases of the systematic review



General data

The results showed that Satisfaction measurement instruments have been applied all around the world (Europe, Asia, North America, Australia, and Africa). The sample size in the studies screened showed a wide variation. The number of people included ranged from 80²⁰ to 7.093²¹; almost all studies collected some kind of cultural and socio demographic variables from the patients (age, gender, family income, educational level and health area of treatment).

It was found that there is not a gold standard or a standardized instrument for satisfaction measurements in health services. There are many different instruments being used, whether in hospitals as in clinics. It was difficult to identify which health services were public or private, most studies did not clearly state this^{20,21}. Only five studies^{22,23,24,25,26} clearly stated the ownership of the service studied.

Thirty-four different instruments for satisfaction measurement were found and few of them were utilized by more than one study: Consumer Emergency Care Satisfaction Scale (CECSS)^{26,27} – twice; *Echelle de Qualité des Soins en Hospitalisation* (EQS-H scale)^{28,29} – twice; Primary Care Satisfaction Survey for Women (PCSSW)^{30,31} – twice.

The most frequent study design observed was instrument validation and only six cross-cultural adaptations were found^{26,30,32,33,34,35}. Service complexity level showed major differences due to the fact that 3 studies utilized university hospitals to measure patient satisfaction, 10 used only clinics and other 20 studies measured patient satisfaction in general hospitals (Table 1).

Tabel 1: Characteristics of the studies

Study	N	Health care setting	Study design	Instrument used	Construct dimensions	Measuremet proprieties**	Psychometric values	Cosmin scores**
Meehan <i>et al</i> (2002) [46]	356	Hospital	Validation	Inpatient Evaluation of Service Questionnaire (IESQ)	Multidim. (3)*	1-Internal consistency 2-Structural validity 3-Hipoheses testing 4-Content validity	$\alpha > 0.86$; correlation with overall satisfaction measure ($r=0.78$)	1-Fair 2-Fair 3-Fair 4-Exellent
Hendriks <i>et al</i> (2004) [49]	728	Hospital	Adaptation	Satisfaction with Hospital Care Questionnaire	Multidim. (12)*	1-Internal consistency 2-Structural validity	$\alpha > 0.6$; inter-scales correlations ($r = 0.3-0.7$)	1-Fair 2-Fair
Scholle <i>et al</i> (2004) [31]	1202	Clinic	Validation	Primary Care Satisfaction Survey for Women (PCSSW)	Multidim. (3)*	1-Structural validity 2-Internal consistency 3-Hipoheses testing 4-Content validity	$\alpha > 0.80$; correlation between scales ($r=0.65 - 0.69$); Correlation with other instruments ($r > 0.42$)	1-Good 2-Good 3-Good 4-Exellent
Lam <i>et al</i> (2005) [34]	711	Hospital	Trans cultural Adaptation	Medical Interview Satisfaction Scale (C-MISS)	Bidim.	1-Internal consistency 2-Structural validity 3-Cross-cultural validity 4-Hipoheses testing 5-Reliability	$\alpha > 0.84$; inter-item and item-total correlations ($r > 0.3$) ; sub-scales to total correlation ($r > 0.87$); inter-scales ($r = 0.58$)	1-Good 2-Good 3-Fair 4-Fair 5-Fair
Gonzales <i>et al</i> (2005) [43]	1910	Hospital	Validation	In-patient Satisfaction Questionnaire	Multidim. (6)*	1-Internal consistency 2-Structural validity 3-Hipoheses testing 4-Content validity	$\alpha > 0.60$; inter-scales correlations ($r=0.16 - 0.62$)	1-Exellent 2-Exellent 3-Fair 4-Exellent
Wilde and Hendriks (2005) [36]	262	Clinic	Adaptation	Client Satisfaction Questionnaire (CSQ)	Unidim.	1-Structural validity 2-Internal consistency 3-Hypotheses testing	$\alpha = 0.92$; correlations between other instruments ($r > 0.46$)	1-Good 2-Exellent 3-Good
Gerkenmey and Austinr (2005) [22]	278	Clinic Children Mental Health Service	Validation	Parent Satisfaction Scale (PSS)	Unidim.	1-Internal consistency 2-Structural validity 3-Hipoheses testing	$\alpha > 0.70$; correlation with other instruments ($r = 0.86$)	1-Fair 2-Fair 3-Fair

Study	N	Healthcare setting	Study design	Instrument used	Construct dimensions	Measurement properties**	Psychometric values	Cosmin scores**
Otani <i>et al</i> (2005) [52]	8465	Hospital	Adaptation	Patient satisfaction survey	Multidim. (3)*	1-Internal consistency 2-Structural validity 3-Measurement error	$\alpha > 0.8$; RMSA=0.12	1-Fair 2-Fair 3-Fair
Poinsot <i>et al</i> (2006) [41]	416	Hospital (Oncology Center)	Validation	Out-Patsat35	Bidim.	1-Structural validity 2-Internal consistency	$\alpha > 0.90$; inter-tem correlations ($r = 0.33 - 0.91$); correlation with other instrument ($r = 0.38$)	1-Poor 2-Poor
Tso <i>et al</i> (2006) [25]	344	Hospital	Validation	Concise Outpatient Department User Satisfaction Scale	Bidim.	1-Internal consistency 2-Structural validity 3-Hypotheses testing	$\alpha > 0.90$; Guttman split-half reliability (0.93); Spearman-Brown reliability (0.94);	1-Fair 2-Fair 3-Fair
Gagnon <i>et al</i> (2006) [48]	873	Clinic	Validation	Health Care Satisfaction Questionnaire (HCSQ)	Multidim. (3)*	1-Structural validity 2-Internal consistency 3-Content validity 4-Reliability	$\alpha > 0.74$; KMO (0.95); CFI=0.93; NFI=88; AGFI=0.87; RMSEA=0.6; correlations between factor ($r > 0.6$); ICC (> 0.45)	1-Excellent 2-Excellent 3-Excellent 4-Fair
Campbell <i>et al</i> (2007) [37]	1250	Out-of-our health service	Validation	Out-of-hours Patient Questionnaire (OPQ)	Multidim. (8)*	1-Internal consistency 2-Structural validity 3-Reliability 4-Content validity	$\alpha = 0.70$, inter-item and total item correlation ($r > 0.20$)	1-Fair 2-Fair 3-Fair 4-Fair
Moret <i>et al</i> (2007) [28]	2684	Hospital	Adptation	Echelle de Qualité des Soins en Hospitalisation (EQS-H scale)	Bidim.	1-Internal consistency 2-Structural validity 3-Content validity 4-Measurement error	$\alpha > 0.80$; correlation inter-item ($r > 0.40$) and item-total ($r < 0.40$); RMSA=0.063	1-Good 2-Good 3-Fair 4-Good
Orden <i>et al</i> (2007) [45]	651	University Hospital	Validation	SUCE (Satisfacción de los Usuarios de las Consultas Externas)	Bidim.	1-Internal consistency 2-Structural validity 3-Hypotheses testing 4-Content validity	$\alpha > 0.90$; correlation with other instrument ($r = 0.85$); item-total correlation ($r = 0.52 - 0.73$)	1-Good; 2-Good 3-Fair 4-Excellent
Shea <i>et al</i> (2008) [38]	3215	Private Clinic	Adaptation	The Consumer Assessment of Healthcare Providers and systems surveys (CAHPS)	Multidim. (7)*	1-Internal consistency 2-Hypotheses testing 3-Content validity	$\alpha = 0.39 - 0.90$	1-Fair 2-Poor 3-Fair

Study	N	Healthcare setting	Study design	Instrument used	Construct dimensions	Measuremet proprieties**	Psychometric values	Cosmin scores**
Kristensson and Ekwall (2008) [27]	128	Hospital	Adaptation	Consumer Emergency Care Satisfaction Scale (CECSS)	Multidim. (3)*	1-Structural validity 2-Internal consistency	$\alpha > 0.85$	1-Fair 2-Fair
Clark <i>et al</i> (2008) [50]	1415	Clinic and Hospital (Mental Health Services)	Validation	Consumer Perceptions of Care (CPC)	Multidim. (4)*	1-Internal consistency 2-Structural validity 3-Reliability 4-Hipoheses testing 5-Content validity	$\alpha > 0.70$; correlations between instruments and different moments ($r = 0.34 - 0.56$)	1-Good 2-Good 3-Fair 4-Fair 5-Good
Coluccia <i>et al</i> (2009) [40]	1804	Hospital	Adaptation	Questionnaire of Quality Perception (QQP)	Multidim. (4)*	1-Structural Validity	Correlations between factors 0.56 – 0.87; GFI=0.92; RMSEA=0.18; CFI=0.96; NNFI=0.96	1-Exellent
Arraras <i>et al</i> (2009) [20]	80	Hospital	Adaptation	European Organization for Research and Treatment of Cancer in-patient questionnaire (EORTC IN-PATSAT 32)	Multidim. (13)*	1-Internal consistency 2-Hipohese testing	$\alpha > 0.70$; correlations with other instruments > 0.40	1-Fair 2-Fair
Orden <i>et al</i> (2010) [24]	296	Hospital Emergency	Validation	Satisfacción del usuario con el área de urgencia hospitalaria (SUAUH)	Bidim.	1-Internal consistency 2--Structural validity 3-Hipoheses testing 4-Content validity	$\alpha > 0.83$; correlation with other instrument and global satisfaction measures ($r = 0.37-0.51$); item-total correlation ($r = 0.56-0.71$)	1-Fair 2-Fair 3-Fair 4-Exellent
Danielsen <i>et al</i> (2010) [39]	225	Clinic	Validation	Users' experiences and satisfaction with casualty clinics	Multidim. (4)*	1-Internal consistency	$\alpha > 0.82$	1-Fair

Study	N	Healthcare setting	Study design	Instrument used	Construct dimensions	Measurement properties**	Psychometric values	Cosmin scores**
Ekwall and Davis (2010) [26]	157	Hospital	Trans cultural Adaptation	Consumer Emergency Care Satisfaction Scale (CECSS)	Multidim. (3)*	1-Structural validity 2-Internal consistency 3-Cross-cultural validity	$\alpha > 0.77$	1-Fair 2-Fair 3-Fair
Charalambous (2010) [33]	194	Hospital (Oncology Center)	Trans cultural Adaptation	Risser Patient Satisfaction Scale (RPSS)	Multidim. (3)*	1-Structural validity 2-Internal consistency 3-Reliability 4- Cross-cultural validity	$\alpha = 0.89$; Kappa =0.84; inter-item (r=0.45) and item-total correlations (r=0.51); sub scale correlation (r=0.55-0.70)	1-Fair 2-Fair 3-Fair 4-Fair
Erci and Çiftçioğlu (2010) [30]	381	Clinic (Primary Care)	Transcultural Adaptation	Primary Care Satisfaction Survey for Women (PCSSW)	Multidim. (3)*	1-Internal consistency 2-Structural validity 3-Content validity 4-Cross-cultural validity 5-Reliability	$\alpha > 0.80$; item-total correlations (=0.43-0.73); test-retest correlation (r=0.85)	1-Fair 2-Fair 3-Fair 4-Fair 5-Fair
Soufi <i>et al</i> (2010) [29]	214	University Hospital	Adaptation	Echelle de Qualité des Soins en Hospitalisation (EQS-H scale)	Bidim.	1-Internal consistency 2-Structural validity 3-Content validity 4-Cross-cultural validity	$\alpha > 0.88$; GFI = 0.96; RMSEA = 0.090; CFI = 0.93	1-Fair 2-Fair 3-Fair 4-Fair
Schröder <i>et al</i> (2011) [47]	1340	Hospital	Validation	Quality in Psychiatric Care – Outpatient (QPC–OP) instrument	Multidim. (7)*	1-Internal consistency 2-Structural validity	$\alpha = 0.66$; CFI = 1.0, RMSEA = 0.036; SRMR = 0.053; inter-scales correlations (r =0.35 - 0.67)	1-Excellent 2-Excellent
Webster <i>et al</i> (2011) [42]	486	Hospital	Validation	Patient assessment of healthcare for inpatient care (I-PAHC) Patient assessment of healthcare for outpatient care (O-PAHC)	Multidim. (5)*	1-Internal consistency 2-Structural validity 3-Content validity	$\alpha > 0.70$; inter-scales and item correlations (r= 0.2 - 0.6)	1- Good 2-Good 3-Poor

Study	N	Healthcare setting	Study design	Instrument used	Construct dimensions	Measurement properties**	Psychometric values	Cosmin scores**
Sjetne <i>et al</i> (2011) [35]	1324	Hospital	Validation	Generic Short Patient Experiences Questionnaire (GS-PEQ)	Multidim. (7)*	1-Content validity	Theoretical approach	---
Latour <i>et al</i> [51]	441	University Hospital (Neonatal Intensive Care)	Validation	EMPATHIC-N (Empowerment of Parents in The Intensive Care- Neonatology)	Multidim. (5)*	1-Internal consistency 2-Structural validity 3-Reliability 4-Hypotheses testing 5-Content validity	$\alpha > 0.82$; correlation with other instrument and global satisfaction measures ($r = 0.37-0.51$)	1-Good 2-Good 3-Good 4-Fair 5-Good
Evans <i>et al</i> (2012) [23]	397	Clinic (Mental Health)	Validation	Views on Inpatient Care (VOICE)	Multidim. (7)*	1-Structural validity 2-Internal consistency 3-Content validity 4-Hypotheses testing 5-Reliability validity	$\alpha > 0.92$; correlation with other instrument ($r = 0.82$); Kappa ($p = 0.88$)	1-Poor 2-Poor 3-Fair 4-Fair 5-Fair
Manaf <i>et al</i> (2012) [44]	424	Clinic	Validation	Outpatient Satisfaction Instrument	Bidim.	1-Internal consistency 2-Structural validity 3-Hypotheses testing 4-Content validity	$\alpha > 0.94$; correlation with other instruments and overall satisfaction measure ($r = 0.59; 0.62$)	1-Excellent 2-Excellent 3-Good 4-Excellent
Milutinovic <i>et al</i> (2012) [32]	235	Hospital	Transcultural Adaptation	Patient Satisfaction Nursing Care Quality Questionnaire	Unidim.	1-Structural validity 2-Internal consistency 3-Cross-cultural validity	$\alpha > 0.90$; inter-item correlation and item total correlation ($r = 0.52 - 0.92$); GFI (Area under Roc curve - AUC = 0.82)	1-Good 2-Good 3-Good
Ottonello <i>et al</i> (2012) [21]	7093	Hospital (Emergency Care)	Adaptation	SAT-16	Bidim.	1-Structural validity 2-Internal consistency	$\alpha > 0.99$; correlation between item ($r < 0.30$); Rasch reliability values (Item separation index:18.1 and 20.7 ; item separation reliability: 1; person separation index:0.70	1-Excellent 2-Excellent

Study	N	Healthcare setting	Study design	Instrument used	Construct dimensions	Measurement properties**	Psychometric values	Cosmin scores**
Cstillo <i>et al</i> (2012) [53]	483	Hospital	Adaptation	Satisfaction Scale with Primary Health Care Services	Multidim. (4)*	1-Structural validity 2-Internal consistency	$\alpha > 0.75$; RMSEA=0,09; GFI =0.92; AGFI=0.87; NFI=0.94; CFI = 0.95	1-Fair 2-Fair
Dyer <i>et al</i> 2012 [54]	21318	Hospital/ Clinic	Adaptation	Consumer Assessment of Healthcare Providers and Systems (CAHPS-CG)	Multidimen. (3)*	1-Structural validity 2-Internal consistency 3-Reliability	$\alpha > 0.75$; RMSEA=0,05; SRMR=0,04; CFI = 0.97; ICC= 0,01 – 0,11	1-Good 2-Good 3-Good
Vanti <i>et al</i> 2013 [35]	317	Hospital	Transcultural Adaptation	Physical Therapy Patient Satisfaction Questionnaire (PTPSQ-I)	Multidimen. (3)*	1- Structural validity 2-Internal consistency 3-Cross-cultural validity 4-Hypotheses testing	$\alpha > 0.90$; correlations between item and global perceive measures (r=0.1 – 0.42)	1-Fair 2-Fair 3-Good 4-Fair
Vanti <i>et al</i> 2013[56]	354	Hospital	Transcultural Adaptation	Physical Therapy Patient Satisfaction Questionnaire (PTPSQ-I)	Multidimen. (4)*	1- Structural validity 2-Internal consistency 3-Cross-cultural validity 4-Reliability 5-Hypotheses testing	$\alpha > 0.75$; ICC=0.62 – 0.93; correlations between item and global perceive measures (r=0.28 – 0.42)	1-Fair 2-Fair 3-Good 4-Fair 5-Fair

Satisfaction Construct

One can argue that there is almost a consensus between studies on the multidimensional characteristic of the satisfaction construct (Table 1). Only three studies showed a unidimensional structure for the tested instrument^{22,32,36}. However, there is not a clear pattern of what the essential dimensions (factors) that must be included in a satisfaction measurement are. The number of dimension/factors in the instruments scanned ranged from 1 to 13.

It was also possible to observe a trend showing that more than dimension should be used to develop or adapt an instrument. The main dimensions are: patient - healthcare professional interactions, physical environment and internal management process^{22,23,32,36}. Only services with some different characteristics (home care, out-of-hours, psychiatric and pediatric services, and more multidisciplinary complex services) usually utilized other specific dimensions^{22,23,38}.

Methodological quality

The analysis of the psychometric properties indicates that most studies used the classical test theory (CTT) to evaluate the instruments' structure. Only one paper used the item response theory (IRT)¹⁹. The main properties assessed were structural validity and internal consistency. It was observed that just five papers did not approach at least these two proprieties together^{18,37,39,40,41}. Content validity (face validity), hypotheses testing, reliability (test-retest) and cross-cultural validity were other properties analyzed in many studies.

The COSMIN checklist scores demonstrated that most papers achieved a medium score ("Good" or "Fair") on the methodology quality of the psychometric analysis. Only four studies demonstrated low scores ("poor")^{23,39,42,43} and eleven studies had at least one property categorized as excellent^{21,24,31,37,41,44,45,46,47,48,49}. Thirty-seven studies²⁰⁻⁵⁶ presented the Cronbach α value as an internal consistence measure (ranged from 0.39 to 0.99); most of them showed correlations between items, sub-scales and other satisfaction measures (Table 1). The main statistic methods utilized were the exploratory factorial analysis (EFA) and confirmatory factorial analysis (CFA) to evaluate the instruments' structure.

Discussion

Considering the focus on patient-centered care has been proposed all around the world, the development or adaption of satisfaction surveys has been the most common method to evaluate patient's satisfaction. However, the very high specificity of the different types of care makes the development of a single measurement scale transposed into all services more complex. The findings of this revision have shown that most surveys demonstrated that "user satisfaction in health care services" is a multifactorial construct, but there is not a reference or gold standard instrument for the measurement of such concept.

Previously, Boyer *et al*¹¹ showed, in their systematic review about patient satisfaction surveys in psychiatric services, a wide variation in the number of dimensions of the instruments scanned. Even though our review had screened instruments applied on different healthcare settings, the results support these findings, the number of dimensions in the instruments scanned ranged from 1 to 13.

A good theory-based approach is the first step to define any construct. When it comes to users' satisfaction, it has already been shown that the essential aspects are: patient-professional interactions, physical environment and internal management process^{17,18,22,23}. However, services with some different characteristics (home care, out-of-hours, psychiatric and pediatric services, and more multidisciplinary complex services) usually utilized other specific elements, such as phone contact or clear information to the caregiver^{18,24,25,26}.

The complexity of this construct can difficult the development of measurement instruments and the definition of their real contribution for healthcare quality and management. It must be understood that the concept of satisfaction can change through many cultures and different variables may be related depending on patient profile or service^{57,58,59}.

Although almost all clinicians and patients agree that clinical skill, rapport and health-related communication behavior are key element for quality and satisfaction^{59,60}. Our findings demonstrated that all multidimensional instruments utilized sub-scales that were linked with these key elements.

The methodological quality of the questionnaires and their validation process are important topics that should be highlighted⁶¹. All instruments found were developed in

the English language and there were only six cross-cultural adaptation papers^{26,30,32,33,34,35}. The adaptation process to other language and cultures is not simple and the scrutiny made by different researches on distinct sociocultural contexts would help to develop a reliable instrument^{61,62,63}.

The psychometric properties of the questionnaires are just one point of the validation or adaptation process and other steps should be taken^{62,63}. Our results showed that the Structural Validity and Internal Consistency were the main measurement properties described by the studies, but none of them had already undergone all other steps proposed by the COSMIN checklist for validation: reliability, measurement error, content validity, hypotheses testing, criterion validity, cross-cultural validity and responsiveness^{14,15,16,17}.

Discussion about what drives users' satisfaction is necessary to guide the development of new surveys and classic confirmatory analysis. Assessing this issue is quite important to choose an instrument that has already been scrutinized by several authors and crossed many important validation steps^{15,61}.

Unfortunately, until now there are few guidelines to evaluate instruments' validation processes and help to define the most important steps for the development of a new instrument^{14,15,16}, which led to the large quantity of instruments found. This study used the COSMIN guide to check the quality of each measurement property described on instruments validation process and psychometric properties¹⁷. This analysis demonstrated that most of the instrument achieved a medium score ("good" or "fair"). The comparison with other studies is difficult because no other systematic review with similar characteristics was found in the literature.

Even though patient satisfaction has been widely studied in healthcare, no gold standard instrument or even a validated instrument was found for many different contexts²⁰⁻⁵⁶. The option to develop new instruments by the healthcare organizations can reflect the specificity of each culture or health settings, but, at the same time, this phenomenon can show a gap in state-of-the-art approaches in this area.

It should be highlighted that users' satisfaction is just one facet of quality of care and the dimension that support this topic is the concept of patient centered care. Satisfaction studies can help some local health policies because it can show a tendency of presence of quality on care. It had already been shown that improvement in actions to

understand patients' preferences or patients-physicians communication can impact outcomes⁶⁴.

Rathert⁶⁵ reported in their systematic review that patient centered care actions may affect patients over time in their ambulatory settings. Moreover, to understand of how these concepts is related to patients health outcomes can help determine the effectiveness of patient centered care.

This study can present some limitations, above all because the search strategy faced some difficulties to find studies from different areas of healthcare as physiotherapy, odontology or nutrition. It can generate some difficulty in the interpretation of the satisfaction dimensions for the entire health field and the structure of main instruments. Another point is that because the search strategy assumed some restricted headings and only papers written from 2002 on.

Conclusion

The findings demonstrated that there is not a consensus about a reference instrument (gold standard) for patient satisfaction measurement in healthcare. However, the main dimensions of many instruments utilized are patient-professional interactions, physical environment and internal management process. Besides, the measurement instruments analyzed have shown a medium score of methodological quality by the COSMIN checklist. Considering the volume of studies, few papers focused on the psychometric proprieties, validation quality and cross-cultural adaptation of satisfaction surveys. Finally, it seems clear that the "patient satisfaction construct" should be measured considering a multidimensional approach, because even the clinical patients' profile, such as health service environment or professional behavior, can be determinants for patient satisfaction.

References

- 1- Shaw C, Groene O, Mora N, Sunol R. **Accreditation and ISO certification: do they explain differences in quality management in European hospitals?** *International Journal for Quality in Health Care* 2010; 22(6): 445–45.
- 2- Leatherman S, Ferri TG, Berwick D, Omaswa F, Crisp N. **The role of quality improvement in strengthening health systems in developing countries.** *International Journal for Quality in Health Care* 2010; 22(4):237–243.

- 3- Institute of Medicine (Committee on quality of healthcare in America - IOM). **Crossing the quality chasm: a new health system for the 21st century.** National Academy press. Washington, 2001.
- 4- Doyle C, Lennox L, Bell D. **A systematic review of evidence on the links between patient experience and clinical safety and effectiveness.** *BMJ Open* 2013;3.
- 5- Machado NP, Nogueira LP. **Evaluation of Physical Therapy service user satisfaction.** *Brazilian Journal of Physical Therapy* 2008; 12 (5):401-8.
- 6- Evrard Y. **La satisfaction des consommateurs : état des recherches.** *Revue Française du Marketing* 1993; 144: 53-65.
- 7- Oliver RL. **A cognitive model of the antecedents and consequences of satisfaction decision.** *Journal of Marketing Research* 1980; 44(4):460-467. 151
- 8- Fournier S, Mick DG. **Rediscovering satisfaction.** *Journal of Marketing* 1999; 63(4): 5-23.
- 9- Paim J, Travassos C, Almeida C, Bahia L, Macinko J. **The Brazilian health system: history, advances, and challenges.** *Lancet* 2011; 377: 1778–97.
- 10- Hibbard JH, Greene J and Daniel D. **What is Quality Anyway? Performance Reports That Clearly Communicate to Consumers the Meaning of Quality of Care.** *Med Care Res Rev* 2010; 67: 275-293.
- 11- Boyer L, Baumstarck-Barrau K, Cano N *et al.* **Assessment of psychiatric inpatient satisfaction: A systematic review of self-reported instruments.** *European Psychiatry* 2009; 24:540–549.
- 12- Fenton JJ, Jerant AF, Bertakis KD, Franks P. **The Cost of Satisfaction: A National Study of Patient Satisfaction, Health Care Utilization, Expenditures, and Mortality.** *Arch Intern Med* 2012;172(5):405-41.
- 13- Hekkert KD, Cihangir S, Kleefstra SM *et al.* **Patient satisfaction revisited: a multilevel approach.** *Social Science and Medicine* 2009;69: 68–75.
- 14- Mokkink LB, Terwee CB, Patrick DL *et al.* **The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study.** *Qual Life Res* 2010;19:539–549.

- 15- Mokkink LB, Terwee CB, Patrick DL *et al.* **The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes** *Journal of Clinical Epidemiology* 2010; 63:737-745.
- 16- Mokkink¹ LB, Terwee¹ CB, Gibbons E *et al.* **Inter-rater agreement and reliability of the COSMIN (Consensus-based Standards for the selection of health status Measurement Instruments) Checklist.** *BMC Medical Research Methodology* 2010; 10:82.
- 17- Terwee CB, Mokkink LB, Knol DL *et al.* **Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist.** *Qual Life Res* 2011; 21:651–657.
- 18- Mutsaers JH A M, Peters R, Pool-Goudzwaard AL *et al.* **Psychometric properties of the Pain Attitudes and Beliefs Scale for Physiotherapists: a systematic review.** *Manual Therapy* 2012; 17:213-218. 152
- 19- Janaudis-Ferreira T, Beauchamp MK, Goldstein RS, Brooks D. **How should we measure arm exercise capacity in COPD? A systematic review.** *Chest* 2012; 141:111-120.
- 20- Arraras JI, Vera R, Martánez M, Hernández B, Lãnez N, Rico M, et al. **The EORTC cancer in-patient satisfaction with care questionnaire: EORTC IN-PATSAT32.** *Clinical and Translational Oncology* 2009;11(4):237-42.
- 21- Ottonello M, Franchignoni F, Giordano A, Benevolo E. **Patient satisfaction with hospital rehabilitation: Validation of the SAT-16 questionnaire through Rasch analysis.** *Minerva Medica* 2012; 103(1):1-11.
- 22- Gerkenmeyer JE, Austin JK. **Development and testing of a scale measuring parent satisfaction with staff interactions.** *J Behav Health Serv Res* 2005; 32(1):61-73.
- 23- Evans J, Rose D, Flach C, Csipke E, Glossop H, McCrone P, et al. **VOICE: developing a new measure of service users' perceptions of inpatient care, using a participatory methodology.** *J Ment Health* 2012;21(1):57-71.
- 24- Orden SG, Garcia AC, Flavia L, Gijón R, Rieiro CR, De Blas CS, et al. **Development and validation of a questionnaire to assess satisfaction with hospital emergency care.** *Emergency Medicine Journal* 2010;28(9):770-4.

- 25- Tso IF, Ng SM, Chan CLW. **The development and validation of the concise outpatient department user satisfaction scale.** *International Journal for Quality in Health Care* 2006; 18(4):275-80.
- 26- Ekwall A, Davis BA. **Testing a Swedish version of the consumer emergency care satisfaction scale in an emergency department and 2 observation wards.** *Journal of Nursing Care Quality* 2010; 25(3):266-73.
- 27- Kristensson J, Ekwall A. **Psychometric properties of the consumer emergency care satisfaction scale: tested on persons accompanying patients in emergency department.** *J Nurs Care Qual* 2008; 23(3):277-82.
- 28- Moret L, Nguyen JM, Pillet N, Falissard B, Lombrail P, Gasquet I. **Improvement of psychometric properties of a scale measuring inpatient satisfaction with care: a better response rate and a reduction of the ceiling effect.** *BMC Health Serv Res* 2007; (7):197.
- 29- Soufi G, Belayachi J, Himmich S, Ahid S, Soufi M, Zekraoui A, et al. **Patient satisfaction in an acute medicine department in Morocco.** *BMC Health Services Research* 2010; 10:149. 153
- 30- Erci B, Ciftcioglu S. **Psychometric evaluation of the primary health-care satisfaction scale in Turkish women.** *Int J Qual Health Care* 2010; 22(6):500-6.
- 31- Scholle SH, Weisman CS, Anderson RT, Camacho F. **The development and validation of the primary care satisfaction survey for women.** *Women's Health Issues* 2004; 14(2):35-50.
- 32- Milutinovic' D, Simin D, Brkic' N, Brkic'S. **The patient satisfaction with nursing care quality: The psychometric study of the Serbian version of PSNCQ questionnaire.** *Scandinavian Journal of Caring Sciences* 2012; 26(3):598-606.
- 33- Charalambous A. **Validation and test-retest reliability of the Risser patient satisfaction scale in Cyprus.** *J Nurs Manag* 2010; 18(1):61-9.
- 34- Lam WW, Fielding R, Chow L, Chan M, Leung GM, Ho EY. **The Chinese Medical Interview Satisfaction Scale-revised (C-MISS-R): development and validation.** *Qual Life Res* 2005;14(4):1187-92.
- 35- Vanti C, Monticone M, Ceron D, Bonetti F, Piccarreta R, Guccione AA, et al. **Italian version of the physical therapy patient satisfaction questionnaire: cross-cultural adaptation and psychometric properties.** *Phys Ther* 2013;93(7):911-22.

- 36- Sjetne IS, Bjertnaes OA, Olsen RV, Iversen HH, Bukholm G. **The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway.** *BMC Health Serv Res* 2011; 11:88.
- 37- Wilde EF, Hendriks VM. **The Client Satisfaction Questionnaire: psychometric properties in a Dutch addict population.** *Eur Addict Res* 2005; 11(4):157-62.
- 38- Campbell JL, Dickens A, Richards SH, Pound P, Greco M, Bower P. **Capturing users' experience of UK out-of-hours primary medical care: piloting and psychometric properties of the Out-of-hours Patient Questionnaire.** *Qual Saf Health Care* 2007; 16(6):462-8.
- 39- Shea JA, Guerra CE, Weiner J, Aguirre AC, Ravenell KL, Asch DA. **Adapting a patient satisfaction instrument for low literate and Spanish-speaking populations: Comparison of three formats.** *Patient Education and Counseling* 2008;73(1):132-40.
- 40- Danielsen K, Bjertnaes OA, Garratt A, Forland O, Iversen HH, Hunskaar S. **The association between demographic factors, user reported experiences and user satisfaction: Results from three casualty clinics in Norway.** *BMC Family Practice* 2010; 11:73. 154
- 41- Coluccia A, Cioffi R, Ferretti F, Lorini F. **Patient perception of the quality of healthcare services: a psychometric study on the Questionnaire of Quality Perception.** *Gazz Med Ital - Arch Sci Med* 2009, 168(3):179-87.
- 42- Poinot R, Altmeyer A, Conroy T, Savignoni A, Asselain B, Leonard I, et al. **Multisite validation study of questionnaire assessing out-patient satisfaction with care questionnaire in ambulatory chemotherapy or radiotherapy treatment.** *Bull Cancer* 2006; 93(3):315-27.
- 43- Webster TR, Mantopoulos J, Jackson E, Cole-Lewis H, Kidane L, Kebede S, et al. **A brief questionnaire for assessing patient healthcare experiences in low-income settings.** *International Journal for Quality in Health Care* 2011; 23(3):258-68.
- 44- Gonzalez N, Quintana JM, Bilbao A, Escobar A, Aizpuru F, Thompson A, et al. **Development and validation of an in-patient satisfaction questionnaire.** *International Journal for Quality in Health Care* 2005;17(6):465-72.
- 45- Manaf NHA, Mohd D, Abdullah K. **Development and validation of patient satisfaction instrument.** *Leadership in Health Services* 2012; 25(1):27-38.

- 46- Orden S, Rieiro RC, Luceron OCM, Chacon Garia A *et al.* **Design and validation of a questionnaire of the hospital outpatient clinics in Madrid, Spain.** *Rev Esp Salud Publica* 2007; 81(6):637-45.
- 47- Meehan T, Bergen H, Stedman T. **Monitoring consumer satisfaction with inpatient service delivery: the Inpatient Evaluation of Service Questionnaire.** *Aust N Z J Psychiatry* 2002; 36(6):807-11.
- 48- Schroder A, Ahlström G, Wilde-Larsson B, Lundqvist LO. **Psychometric properties of the Quality in Psychiatric Care - Outpatient (QPC-OP) instrument.** *International Journal of Mental Health Nursing* 2011; 20(6):445-53.
- 49- Gagnon M, Hébert R, Dubé M, Dubois MF. **Development and validation of the Health Care Satisfaction Questionnaire (HCSQ) in elders.** *Journal of Nursing Measurement* 2006;14(3):190-204.
- 50- Hendriks AA, Vrieling MR, van Es SQ, De Haes HJ, Smets EM. **Assessing inpatients' satisfaction with hospital care: should we prefer evaluation or satisfaction ratings?** *Patient Educ Couns* 2004; 55(1):142-6.
- 51- Clark C, Young MS, Jackson E, Graeber C, Mazelis R, Kammerer N, et al. **Consumer perceptions of integrated trauma-informed services among women with co-occurring disorders.** *J Behav Health Serv Res* 2008; 35(1):71-90.
- 52- Latour JM, Duivenvoorden HJ, Hazelzet JA, van Goudoever JB. **Development and validation of a neonatal intensive care parent satisfaction instrument.** *Pediatr Crit Care Med* 2012; 13:554-559.
- 53- Otani K, Kurz RS, Harris LE. **Managing primary care using patient satisfaction measures.** *J Healthc Manag* 2005; 50(5):311-24.
- 54- Castillo AF, Vilchez-Lara MJ, Sada-lázaro E. **Satisfaction with hospitalization scale: Adaption and psychometrics properties.** *Behavioral Psychology* 2012; 20(2):383-400.
- 55- Dyer N, Sorra JS, Smith SA, Cleary PD, Hays RD. Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician and Group Adult Visit Survey. *Med Care* 2012; 50 (Suppl):28-34.
- 56- Vanti C, Bonetti F, Ceron D, Piccarreta R, Violante FS, Guccione A, et al. Cross-cultural adaptation and validation of the Physical Therapy Outpatient

- Satisfaction Survey in an Italian musculoskeletal population. *BMC Musculoskelet Disord* 2013;14:125.
- 57- Rozenblum R , Lisby M, Hockey PM, et al. **The patient satisfaction chasm: the gap between hospital management and frontline clinicians.** *Quality and Safety in Health Care* 2013; 22:242–250.
- 58- Rozenblum, Lisby M, Hockey PM, *et al.* **Uncovering the blind spot of patient satisfaction: an international survey.** *BMJ Qual Saf* 2011; 20:959-965.
- 59- Levine R, Shore K, Lubalin J, Garfinkel S, Hurtado M, Craman K. **Comparing physician and patient perceptions of quality in ambulatory care.** *International Journal for Quality in Health Care* 2012; 24(4):348–356.
- 60- Rashid WE; Jusoff K. **Service quality in health care setting.** *International Journal of Health Care Quality Assurance* 2009; 22(5):471-482.
- 61- Reichenheim ME, Moraes CL, Oliveira ASD, Lobato G. **Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): empirical evidence for a general factor.** *BMC Medical Research Methodology* 2011, 11:93.
- 62- Herdman M, Fox-Rushby J, Badia X. **A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach.** *Qual Life Res* 1998;7(4):323-35.
- 63- Herdman M, Fox-Rushby J, Badia X. **“Equivalence” and the translation and adaptation of health-related quality of life questionnaires.** *Qual Life Res* 1997; 6(3):237-47.
- 64- Lee YY, Lin JL. **Do patient autonomy preferences matter? Linking patient-centered care to patient-physician relationship and health outcomes.** *Social Science and Medicine* 2010; 71:1811-1818.
- 65- Rathert C, Wyrwich MD, Boren SA. **Patient-centered care and outcomes: a systematic review of the literature.** *Medical Care Research and Review* 2012; 70(4):351-379.

APÊNDICE B – Versão do MedRisk utilizada na fase de pré-teste⁴

Por favor responda as seguintes questões:

() Idade () Masculino () Feminino

Área geral de tratamento (marque todas as que se aplicam) :

()PESCOÇO ()COSTAS ()BRAÇO ()PERNA ()PÉ/tornozelo ()MÃO/punho
()Outros (especifique) _____

Você completou o tratamento?

() Sim () Não Data do último tratamento: _____

Você retornou ao trabalho?

() Não () Sim / Data de retorno ao trabalho: _____
() Não se aplica/Não trabalhava antes da lesão
() A lesão não me tirou do trabalho

Como você se lesionou?

() No trabalho () Atividades pessoais () Outros _____

Escolaridade:

() **Elementar Incompleto (até 3ª série do 1º grau)**
() **Elementar Completo (até 4ª série do 1º grau)**
() **Básico Incompleto (até 7ª série do 1º grau)**
() **Básico Completo (até 8ª série do 1º grau)**
() **Médio Incompleto (até 2º ano do 2º grau)**
() **Médio Completo (até 3º ano do 2º grau)**
() **Ensino Superior Incompleto**
() **Ensino Superior Completo**

Renda familiar:

() **Até 1 salário mínimo** () **1 à 3 salários mínimos**
() **3 à 5 salários mínimos** () **Mais de 5 salários mínimos**

Você possui alguma doença crônica?

() Não () Sim/Qual _____

⁴ Em negrito estão destacadas as questões utilizadas para coleta das informações socioeconômicas (não presentes no instrumento original).

Você acabou de receber tratamento fisioterápico para sua lesão. Nós gostaríamos de conhecer sua opinião a respeito do tratamento que recebeu. Por favor, responda as questões a baixo marcando com um círculo a resposta que melhor descreve sua opinião sobre o tratamento fisioterápico.

	Discordo Fortemente	Discordo	Neutro	Concordo	Concordo Fortemente	Não se Aplica
1. A(o) recepcionista do serviço foi atenciosa(o)	1	2	3	4	5	NA
2. O processo de registro foi adequado	1	2	3	4	5	NA
3. A sala de espera era confortável (iluminação, temperatura, os móveis)	1	2	3	4	5	NA
4. A localização do serviço não era adequada	1	2	3	4	5	NA
5. Eu esperei demais para ser atendido pelo meu fisioterapeuta	1	2	3	4	5	NA
6. O horário de funcionamento era adequado para mim	1	2	3	4	5	NA
7. Meu fisioterapeuta não dedicou tempo suficiente para mim	1	2	3	4	5	NA
8. Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi	1	2	3	4	5	NA
9. Meu fisioterapeuta me tratou com respeito	1	2	3	4	5	NA
10. Os funcionários foram respeitosos	1	2	3	4	5	NA
11. Meu fisioterapeuta não escutou as minhas preocupações	1	2	3	4	5	NA
12. Meu fisioterapeuta respondeu a todas as minhas perguntas	1	2	3	4	5	NA
13. Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	1	2	3	4	5	NA
14. O serviço e suas instalações eram limpos	1	2	3	4	5	NA
15. O serviço utiliza equipamentos modernos	1	2	3	4	5	NA
16. Meu fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar	1	2	3	4	5	NA
17. Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta	1	2	3	4	5	NA
18. Eu voltaria a este serviço para futuros atendimentos ou cuidados	1	2	3	4	5	NA

APÊNDICE C – Perguntas anexadas ao instrumento apenas na fase do pré-teste para análise da compreensão do instrumento

Itens não compreendidos (Por quê / Como você perguntaria):

- CONFIRMAÇÃO SENTIDO DENOTATIVO / CONOTATIVO

Item 1 - Para você, o que é uma pessoa **atenciosa**?

Item 2 - O que você entende por “**processo de registro**” ?

Item 4 - E se eu te perguntasse: “A localização da clínica não era **conveniente**” . Você entenderia melhor?

Item 5 - Você tem a sensação de que possui um fisioterapeuta exclusivo (**MEU**) voltado para o seu tratamento ?

Item 6 - E se eu te perguntasse: “O horário de funcionamento era **conveniente** para mim” . Você entenderia melhor?

Item 7 - O que você entende por “**tempo suficiente de tratamento**”?

Item 8 - E se eu te perguntasse: “meu fisioterapeuta me **explicou todos os detalhes do tratamento que recebi** . Você entenderia melhor?

Item 9 - O que você entende por “**me tratou com respeito**”?

Item 10 - E se eu te perguntasse: ” Meu fisioterapeuta não deu **atenção** às minhas preocupações”. Você entenderia melhor?

Item 14 - O que você entende por “**instalações de um serviço de saúde**”?

Item 16 – O que você entende por “**programa domiciliar**”?

Para avaliar a sua satisfação, você acha relevante ter alguma pergunta referente ao local de estacionamento disponibilizado pelo serviço?

Durante os atendimentos, o fisioterapeuta possuía algum assistente?

Para avaliação da sua satisfação com o serviço, você acha importante a inclusão de algum outro ponto, não abordado por este questionário?

APÊNDICE D – Versão MedRisk pré-testada e aplicada para análise psicométrica (N 325)

POR FAVOR RESPONDA AS SEGUINTEs QUESTÕES:

() Idade ☐ Masculino ☐ Feminino

ÁREA GERAL DE TRATAMENTO (MARQUE TODAS AS QUE SE APLICAM):

☐ Pescoço ☐ Costas ☐ Braço ☐ Perna ☐ Pé/tornozelo ☐ Mão/punho
☐ Outros (especifique) _____

VOCÊ COMPLETOU O TRATAMENTO?

☐ Sim ☐ Não Data do último tratamento: _____

VOCÊ RETORNOU AO TRABALHO?

☐ Não ☐ Sim / Data de retorno ao trabalho: _____☐ Não se aplica/Não trabalhava antes da lesão☐ A lesão não me tirou do trabalho

COMO VOCÊ SE LESIONOU?

☐ No trabalho ☐ Atividades pessoais ☐ Outros _____ESCOLARIDADE:☐ Elementar Incompleto (até 3ª série do 1º grau)
4ª série do 1º grau)☐ Elementar Completo (até☐ Básico Incompleto (até 7ª série do 1º grau)
8ª série do 1º grau)☐ Básico Completo (até☐ Médio Incompleto (até 2º ano do 2º grau)
3º ano do 2º grau)☐ Médio Completo (até☐ Ensino Superior Incompleto☐ Ensino Superior CompletoRENDA FAMILIAR:☐ Até 1 salário mínimo☐ 1 à 3 salários mínimos☐ 3 à 5 salários mínimos☐ Mais de 5 salários mínimos

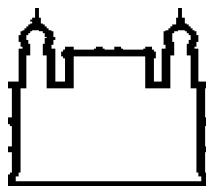
VOCÊ POSSUI ALGUMA DOENÇA CRÔNICA?

☐ Não ☐ Sim/Qual _____

Você acabou de receber tratamento fisioterápico para sua lesão. Nós gostaríamos de conhecer sua opinião a respeito do tratamento que recebeu. Por favor, responda as questões a baixo marcando com um círculo a resposta que melhor descreve sua opinião sobre o tratamento fisioterápico.

	Discordo Fortemente	Discordo	Neutro	Concordo	Concordo Fortemente	Não se Aplica
1. A(o) recepcionista do serviço foi atenciosa(o)	1	2	3	4	5	NA
2. A matrícula no serviço de fisioterapia foi adequada	1	2	3	4	5	NA
3. A sala de espera era confortável (iluminação, temperatura, os móveis)	1	2	3	4	5	NA
4. A localização do serviço não era adequada	1	2	3	4	5	NA
5. Eu esperei demais para ser atendido pelo meu fisioterapeuta	1	2	3	4	5	NA
6. O horário de funcionamento era adequado para mim	1	2	3	4	5	NA
7. Meu fisioterapeuta não dedicou tempo suficiente para mim	1	2	3	4	5	NA
8. Meu fisioterapeuta me deu explicação completa sobre o tratamento que eu recebi	1	2	3	4	5	NA
9. Meu fisioterapeuta me tratou com respeito	1	2	3	4	5	NA
10. Os funcionários foram respeitosos	1	2	3	4	5	NA
11. Meu fisioterapeuta não escutou as minhas preocupações	1	2	3	4	5	NA
12. Meu fisioterapeuta respondeu a todas as minhas perguntas	1	2	3	4	5	NA
13. Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	1	2	3	4	5	NA
14. O serviço e suas instalações eram limpos	1	2	3	4	5	NA
15. O serviço utiliza equipamentos modernos	1	2	3	4	5	NA
16. Meu fisioterapeuta me deu orientações detalhadas a respeito do meu programa domiciliar	1	2	3	4	5	NA
17. Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta	1	2	3	4	5	NA
18. Eu voltaria a este serviço para futuros atendimentos ou cuidados	1	2	3	4	5	NA

APÊNDICE E – Termo de consentimento livre e esclarecido



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Escola Nacional de Saúde Pública Sergio Arouca

Comitê de Ética em Pesquisa



SATISFAÇÃO DE PACIENTES: VALIDAÇÃO DE INSTRUMENTO DE MENSURAÇÃO E ANÁLISE DAS VARIÁVEIS EXPLICATIVAS DO CONSTRUTO

Prezado participante,

Você está sendo convidado(a) a participar da pesquisa “Estudo da satisfação de pacientes em serviços de fisioterapia ambulatorial na rede do sistema único de saúde (SUS).

Esta, será desenvolvida por Renato Santos de Almeida, discente do curso de Doutorado em Saúde Pública, da Escola Nacional de Saúde Pública Sergio Arouca da Fundação Oswaldo Cruz (ENSP/FIOCRUZ), sob orientação da Professora Dr. Mônica Martins.

Este termo está redigido em duas vias, das quais uma permanecerá com o participante e a outra cópia com o pesquisador responsável.

Objetivo central

O objetivo central do estudo é: utilizar o questionário Medrisk para identificar o grau de satisfação do paciente com o serviço de saúde utilizado e ainda determinar as questões que podem influenciar seu julgamento.

Sua participação no estudo

O convite a sua participação se deve ao fato do senhor ter realizado fisioterapia neste centro que oferece atendimento pelo SUS e sua queixa estar dentro do perfil de pacientes pretendidos.

Sua participação é voluntária, isto é, ela não é obrigatória e você tem plena autonomia para decidir se quer ou não participar, bem como retirar sua participação a qualquer momento. Você não será penalizado de nenhuma maneira caso decida não consentir sua participação, ou desistir da mesma. Contudo, ela é muito importante para a execução da pesquisa.

Mecanismos para garantir o sigilo e privacidade

Serão garantidas a confidencialidade e a privacidade das informações por você prestadas.

Qualquer dado que possa identificá-lo será omitido na divulgação dos resultados da pesquisa e o material armazenado em local seguro.

A qualquer momento, durante a pesquisa, ou posteriormente, você poderá solicitar do pesquisador informações sobre sua participação e/ou sobre a pesquisa, o que poderá ser feito através dos meios de contato explicitados neste Termo.

Procedimentos detalhados que serão utilizados na pesquisa

A sua participação consistirá em responder às perguntas de um questionário fornecido pelo pesquisador. São questões referentes ao seu grau de satisfação com o serviço oferecido por este centro de fisioterapia e também referente ao seu perfil social e econômico. Responda da maneira mais realista e verdadeira possível.

Tempo de duração

O tempo de duração para preenchimento do questionário é de aproximadamente 15 minutos.

Guarda dos dados e material coletados na pesquisa

Ao final da pesquisa, todo material será mantido em arquivo, por pelo menos 5 anos, conforme Resolução 466/12 e orientações do CEP/ENSP.

Benefícios

O benefício relacionado com a sua colaboração nesta pesquisa é o de implementação de melhorias na qualidade do atendimento ao paciente em serviços de saúde que ofereçam este tipo de tratamento.

Previsão de riscos ou desconfortos

O presente estudo não possui riscos a sua saúde. Se o senhor se sentir constrangido em responder alguma pergunta do questionário não é necessário respondê-la.

Sobre divulgação dos resultados da pesquisa

Os resultados serão divulgados para comunidade científica por meio do desenvolvimento de uma tese de doutorado, artigos científicos e palestras na área da saúde.

Contato com o(a) pesquisador(a) responsável

Endereço: Escola Nacional de Saúde Pública Sergio Arouca/ FIOCRUZ, Rua Leopoldo Bulhões, 1480 – Andar Térreo - Manguinhos - Rio de Janeiro – RJ - CEP: 21041-210

Telefone: (21) 8187-5601

Rio de Janeiro, ____/____/____

Declaro que entendi os objetivos e condições de minha participação na pesquisa e concordo em participar.

Assinatura do sujeito da pesquisa

Renato Almeida
(Doutorando em Saúde pública) ENSP/FIOCRUZ

Em caso de dúvida quanto à condução ética do estudo, entre em contato com o

Comitê de Ética em Pesquisa da ENSP: Tel:(21) 2598-2863 ;

E-mail:cep@ensp.fiocruz.br

APÊNDICE F– Análise fatorial exploratória realizada com o programa SPSS 17.0

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,628
Bartlett's Test of Sphericity	Approx. Chi-Square	115,667
	df	15
	Sig.	,000

Communalities

	Initial	Extraction
i1	1,000	,564
i3	1,000	,654
i15	1,000	,606
i9	1,000	,503
i10	1,000	,659
i13	1,000	,506

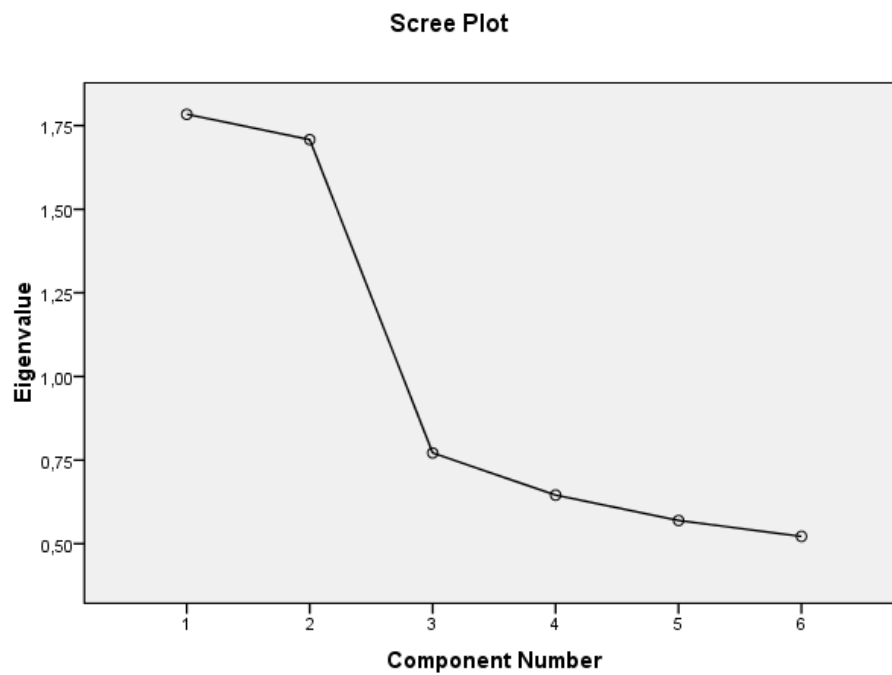
Extraction Method: Principal
Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings ^a
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	1,784	29,730	29,730	1,784	29,730	29,730	1,770
2	1,708	28,468	58,199	1,708	28,468	58,199	1,722
3	,771	12,857	71,056				
4	,645	10,756	81,812				
5	,569	9,491	91,302				
6	,522	8,698	100,000				

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.



Pattern Matrix^a

	Component	
	1	2
i3	,794	-,164
i15	,778	,009
i1	,730	,168
i10	,009	,812
i13	,030	,710
i9	-,029	,709

Extraction Method: Principal

Component Analysis.

Rotation Method: Promax with
Kaiser Normalization.

a. Rotation converged in 3
iterations.

APÊNDICE G – Análise fatorial confirmatória realizadas com o programa Mplus 5.0

Mplus VERSION 5.1
MUTHEN & MUTHEN
04/09/2015 12:10 AM

INPUT INSTRUCTIONS

TITLE: CFA MEDRISK DOUTORADO

DATA:
FILE IS planilha medrisk.dat;

VARIABLE:
NAMES ARE i1-i18;
USEVARIABLES ARE i1,i3,i9,i10,i13,i15;
CATEGORICAL ARE i1,i3,i9,i10,i13,i15;

MODEL:
int by i9,i10,i13;
esteapo by i1,i3,i15;

ANALYSIS:
TYPE IS GENERAL;
ESTIMATOR IS WLSMV;
ITERATIONS = 1000;
CONVERGENCE = 0.00005;

OUTPUT: SAMPSTAT MODINDICES STANDARDIZED FSDETERMINACY;

*** WARNING in OUTPUT command
FSDETERMINACY is not available for analysis with at least one
categorical
dependent variable. Request for FSDETERMINACY is ignored.
1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

CFA MEDRISK DOUTORADO

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	304
Number of dependent variables	6
Number of independent variables	0
Number of continuous latent variables	2

Observed dependent variables

Binary and ordered categorical (ordinal)					
I1	I3	I9	I10	I13	I15

CFI/TLI

CFI	0.975
TLI	0.964

Number of Free Parameters 28

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.059
-----------------	--------------

WRMR (Weighted Root Mean Square Residual)

Value	0.583
--------------	--------------

MODEL RESULTS

STANDARDIZED MODEL RESULTS

STDYX Standardization

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT	BY				
	I9	0.651	0.059	10.974	0.000
	I10	0.734	0.050	14.643	0.000
	I13	0.711	0.059	12.041	0.000
ESTEAPO	BY				
	I1	0.624	0.073	8.608	0.000
	I3	0.591	0.072	8.264	0.000
	I15	0.652	0.078	8.359	0.000
ESTEAPO	WITH				
	INT	0.196	0.089	2.202	0.028

					Two-Tailed
Variances					
	INT	0.424	0.077	5.487	0.000
	ESTEAPO	0.390	0.091	4.304	0.000

STANDARDIZED MODEL RESULTS

STD Standardization

R-SQUARE

Variable Observed	Two-Tailed			
Variance Residual	Estimate	S.E.	Est./S.E.	P-Value
I1 0.610	0.390	0.091	4.304	0.000
I3 0.651	0.349	0.085	4.132	0.000
I9 0.576	0.424	0.077	5.487	0.000
I10 0.462	0.538	0.074	7.322	0.000
I13 0.495	0.505	0.084	6.021	0.000
I15 0.574	0.426	0.102	4.179	0.000

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix
 0.721E-02
 (ratio of smallest to largest eigenvalue)

MODEL MODIFICATION INDICES

Minimum M.I. value for printing the modification index 10.000

	M.I.	E.P.C.	Std E.P.C.	StdYX
E.P.C.				

No modification indices above the minimum value.

Beginning Time: 00:10:23
 Ending Time: 00:10:23
 Elapsed Time: 00:00:00

Saídas do software Mplus 5.0 da análise fatorial confirmatória do modelo proposto por Oliveira et al (2014)

Mplus VERSION 5.1
 MUTHEN & MUTHEN
 04/08/2015 2:34 PM

INPUT INSTRUCTIONS

TITLE: CFA MEDRISK DOUTORADO

```

DATA:
  FILE IS planilha medrisk.dat;

VARIABLE:
  NAMES ARE i1-i18;
  USEVARIABLES ARE i1,i2,i3,i4,i6,i9,i10,i12,i13,i14,i16;
  CATEGORICAL ARE i1,i2,i3,i4,i6,i9,i10,i12,i13,i14,i16;

MODEL:
  int                by i1,i2,i3,i9,i10,i14;
  conv              by i4,i6,i12;
  Edu                by i13,i16;

ANALYSIS:
  TYPE IS GENERAL;
  ESTIMATOR IS WLSMV;
  ITERATIONS = 1000;
  CONVERGENCE = 0.00005;

OUTPUT:  SAMPSTAT MODINDICES STANDARDIZED FSDETERMINACY;

```

```

*** WARNING in OUTPUT command
FSDETERMINACY is not available for analysis with at least one
categorical
dependent variable. Request for FSDETERMINACY is ignored.
1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

```

CFA MEDRISK DOUTORADO

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	304
Number of dependent variables	11
Number of independent variables	0
Number of continuous latent variables	3

Observed dependent variables

Binary and ordered categorical (ordinal)					
I1	I2	I3	I4	I6	I9
I10	I12	I13	I14	I16	

Continuous latent variables

INT	CONV	EDU
-----	------	-----

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Parameterization	DELTA

Input data file(s)
planilha medrisk.dat

Input data format FREE

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)				
	I9	I10	I12	I13
I14				
	_____	_____	_____	_____
I10	0.487			
I12	0.356	0.340		
I13	0.474	0.496	0.516	
I14	0.217	0.251	0.230	0.341
I16	0.154	0.198	0.138	0.193
0.273				

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)
I16

TESTS OF MODEL FIT

Chi-Square Test of Model Fit

Value	143.978*
Degrees of Freedom	30**
P-Value	0.0000

Chi-Square Test of Model Fit for the Baseline Model

Value	645.080
Degrees of Freedom	28
P-Value	0.0000

CFI/TLI

CFI	0.815
TLI	0.828

Number of Free Parameters	54
---------------------------	----

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.112
----------	-------

WRMR (Weighted Root Mean Square Residual)

Value	1.228
-------	-------

MODEL RESULTS

STANDARDIZED MODEL RESULTS

STDYX Standardization

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT	BY				
	I1	0.580	0.058	10.086	0.000
	I2	0.651	0.048	13.539	0.000
	I3	0.313	0.069	4.561	0.000
	I9	0.556	0.055	10.186	0.000
	I10	0.656	0.043	15.233	0.000
	I14	0.508	0.049	10.333	0.000
CONV	BY				
	I4	0.342	0.070	4.884	0.000
	I6	0.468	0.060	7.748	0.000
	I12	0.568	0.060	9.410	0.000
EDU	BY				
	I13	0.581	0.107	5.455	0.000
	I16	0.332	0.073	4.556	0.000
CONV	WITH				
	INT	0.851	0.083	10.259	0.000
EDU	WITH				
	INT	1.131	0.194	5.844	0.000
	CONV	1.393	0.248	5.614	0.000

STDY Standardization

R-SQUARE

Observed Variable Residual				Two-Tailed P-Value
Variance	Estimate	S.E.	Est./S.E.	P-Value
I1	0.336	0.067	5.043	0.000
0.664				
I2	0.423	0.063	6.769	0.000
0.577				
I3	0.098	0.043	2.280	0.023
0.902				
I4	0.117	0.048	2.442	0.015
0.883				

I6	0.219	0.056	3.874	0.000
0.781				
I9	0.309	0.061	5.093	0.000
0.691				
I10	0.430	0.056	7.617	0.000v
0.570				
I12	0.323	0.069	4.705	0.000
0.677				
I13	0.338	0.124	2.728	0.006
0.662				
I14	0.258	0.050	5.167	0.000
0.742				
I16	0.110	0.048	2.278	0.023
0.890				

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix
 0.160E-02
 (ratio of smallest to largest eigenvalue)

MODEL MODIFICATION INDICES

Minimum M.I. value for printing the modification index 10.000

	M.I.	E.P.C.	Std E.P.C.	StdYX
E.P.C.				

No modification indices above the minimum value.

Beginning Time: 14:34:21
 Ending Time: 14:34:21
 Elapsed Time: 00:00:00

Saídas do *software* Mplus 5.0 da análise fatorial confirmatória do modelo proposto por Beattie et al (2002)

Mplus VERSION 5.1
 MUTHEN & MUTHEN
 04/08/2015 2:43 PM

INPUT INSTRUCTIONS

TITLE: CFA MEDRISK DOUTORADO

DATA:
 FILE IS planilha medrisk.dat;

VARIABLE:
 NAMES ARE i1-i18;
 USEVARIABLES ARE i1,i2,i3,i7,i8,i9,i11,i12,i13,i16;
 CATEGORICAL ARE i1,i2,i3,i7,i8,i9,i11,i12,i13,i16;

```

MODEL:
    int                                by i1,i12,i13;
    est                                by i7,i8,i9,i11,i12,i13,i16;

```

```

ANALYSIS:
    TYPE IS GENERAL;
    ESTIMATOR IS WLSMV;
    ITERATIONS = 1000;
    CONVERGENCE = 0.00005;

```

```

OUTPUT:  SAMPSTAT MODINDICES STANDARDIZED FSDETERMINACY;

```

```

*** WARNING in MODEL command
Variable is uncorrelated with all other variables:  I2
*** WARNING in MODEL command
Variable is uncorrelated with all other variables:  I3
*** WARNING in MODEL command
All least one variable is uncorrelated with all other variables
in the model.
Check that this is what is intended.
*** WARNING in OUTPUT command
FSDETERMINACY is not available for analysis with at least one
categorical
dependent variable. Request for FSDETERMINACY is ignored.
4 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

```

CFA MEDRISK DOUTORADO

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	304
Number of dependent variables	10
Number of independent variables	0
Number of continuous latent variables	2

Observed dependent variables

Binary and ordered categorical (ordinal)					
I1	I2	I3	I7	I8	I9
I11	I12	I13	I16		

Continuous latent variables

INT	EST
-----	-----

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Parameterization	DELTA

Input data file(s)
planilha medrisk.dat

Input data format FREE

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)				
	I1	I2	I3	I7
I8				
I1				
I2	0.554			
I3	0.345	0.179		
I7	0.080	0.139	0.141	
I8	0.288	0.282	0.051	0.128
I9	0.036	0.221	-0.035	-0.031
0.564				
I11	0.098	0.275	0.005	0.453
0.204				
I12	0.250	0.283	0.087	-0.067
0.410				
I13	0.229	0.333	0.129	0.010
0.354				
I16	0.251	0.239	0.154	-0.039
0.199				

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)				
	I9	I11	I12	I13
I16				
I11	0.055			
I12	0.356	0.251		
I13	0.474	0.098	0.516	
I16	0.154	0.078	0.138	0.193

TESTS OF MODEL FIT

Chi-Square Test of Model Fit

Value	296.250*
Degrees of Freedom	22**
P-Value	0.0000

Chi-Square Test of Model Fit for the Baseline Model

Value	537.401
Degrees of Freedom	24
P-Value	0.0000

CFI/TLI

CFI	0.466
TLI	0.417

Number of Free Parameters	49
---------------------------	----

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.202
----------	-------

WRMR (Weighted Root Mean Square Residual)

Value	2.203
-------	-------

MODEL RESULTS

				Two-Tailed
Variances				
INT	0.119	0.049	2.437	0.015
EST	0.019	0.017	1.092	0.275

STANDARDIZED MODEL RESULTS

STDYX Standardization

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT	BY				
I1		0.345	0.071	4.875	0.000
I12		1.401	1.668	0.840	0.401
I13		0.907	0.673	1.347	0.178
EST	BY				
I7		0.136	0.062	2.184	0.029
I8		0.780	0.047	16.470	0.000
I9		0.679	0.052	13.158	0.000
I11		0.259	0.053	4.912	0.000
I12		-0.743	1.692	-0.439	0.661
I13		-0.268	0.699	-0.383	0.702
I16		0.288	0.067	4.271	0.000
EST	WITH				
INT		0.910	0.144	6.318	0.000

STDY Standardization

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
R-SQUARE					
Observed	Variable				Two-Tailed
Residual					
	Estimate	S.E.	Est./S.E.	P-Value	Variance

I1	0.119	0.049	2.437	0.015
0.881				
I7	0.019	0.017	1.092	0.275
0.981				
I8	0.609	0.074	8.235	0.000
0.391				
I9	0.461	0.070	6.579	0.000
0.539				
I11	0.067	0.027	2.456	0.014
0.933				
I12	0.622	0.439	1.416	0.157
0.378				
I13	0.452	0.190	2.382	0.017
0.548				
I16	0.083	0.039	2.135	0.033
0.917				

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix
0.964E-08
(ratio of smallest to largest eigenvalue)

MODEL MODIFICATION INDICES

Minimum M.I. value for printing the modification index 10.000

		M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
BY Statements					
INT	BY I2	126.574	1.677	0.578	0.578
EST	BY I2	122.360	4.335	0.590	0.590

APÊNDICE H – Versão do MedRisk final obtida no presente estudo

POR FAVOR, RESPONDA AS SEGUINTE QUESTÕES:

() Idade ☐ Masculino ☐ Feminino

ÁREA GERAL DE TRATAMENTO (MARQUE TODAS AS QUE SE APLICAM):

☐ Pescoço ☐ Costas ☐ Braço ☐ Perna ☐ Pé/tornozelo ☐ Mão/punho
☐ Outros (especifique) _____

VOCÊ COMPLETOU O TRATAMENTO?

☐ Sim ☐ Não Data do último tratamento: _____

VOCÊ RETORNOU AO TRABALHO?

☐ Não ☐ Sim / Data de retorno ao trabalho: _____
☐ Não se aplica/Não trabalhava antes da lesão
☐ A lesão não me tirou do trabalho

COMO VOCÊ SE LESIONOU?

☐ No trabalho ☐ Atividades pessoais ☐ Outros _____

Você acabou de receber tratamento fisioterápico para sua lesão. Nós gostaríamos de conhecer sua opinião a respeito do tratamento que recebeu. Por favor, responda as questões a baixo marcando com um círculo a resposta que melhor descreve sua opinião sobre o tratamento fisioterápico.

	Discordo Fortemente	Discordo	Neutro	Concordo	Concordo Fortemente	Não se Aplica
1. A(o) recepcionista do serviço foi atenciosa(o)	1	2	3	4	5	NA
2. A sala de espera era confortável (iluminação, temperatura, os móveis)	1	2	3	4	5	NA
3. Meu fisioterapeuta me tratou com respeito	1	2	3	4	5	NA
4. Os funcionários foram respeitosos	1	2	3	4	5	NA
5. Meu fisioterapeuta me aconselhou sobre maneiras de evitar problemas futuros	1	2	3	4	5	NA
6. O serviço utiliza equipamentos modernos	1	2	3	4	5	NA
7. Em geral, eu estou completamente satisfeito com os serviços que recebi do meu fisioterapeuta	1	2	3	4	5	NA
8. Eu voltaria a este serviço para futuros atendimentos ou cuidados	1	2	3	4	5	NA