



# Adaptação transcultural e validação de questionários na área da saúde

*Cross-cultural adaptation and validation of health questionnaires*

Sérgio Duarte Dortas Junior, MD, MSc<sup>1</sup>; Omar Lupi, MD, PhD<sup>2</sup>; Gabriela Andrade Coelho Dias, MD, MSc<sup>3</sup>; Manuela Boleira Sieiro Guimarães, MD, MSc<sup>4</sup>; Solange Oliveira Rodrigues Valle, MD, PhD<sup>2</sup>

## RESUMO

Um bom questionário para avaliação da evolução de uma enfermidade deve gerar respostas válidas, ser bem aceito pelos entrevistados e motivar a participação e o fornecimento das informações desejadas. Em virtude do crescente número de ensaios clínicos multicêntricos, há necessidade de se desenvolver medidas específicas para utilização em países cujo idioma não é o inglês. Como diferenças culturais importantes podem estar presentes, essas medidas podem ser elaboradas de duas formas: desenvolver uma nova medida, ou traduzir e adaptar culturalmente uma medida previamente validada em outro idioma. A adaptação cultural de um instrumento é menos onerosa e requer menor tempo para obtenção de uma medida comum para pesquisa do status da doença, além de proporcionar comparações entre grupos culturais diferentes. A adaptação de um questionário para uso em uma nova população culturalmente distinta requer tempo e é dispendiosa. Entretanto, acreditamos que esta é a melhor e mais adequada maneira de garantir uma equivalência. Desta maneira, permite a coleta de dados em diversos países ou num mesmo país de indivíduos de culturas diferentes, evitando, assim, um viés de seleção.

**Descritores:** Inquéritos e questionários, qualidade de vida, atenção à saúde.

## ABSTRACT

A good questionnaire for assessing the progression of a disease should generate valid answers, be well accepted by responders and encourage their participation and the provision of the desired information. Because of the increasing number of multicenter clinical trials being conducted, there is a need to develop specific measures to be used in countries where languages other than English are spoken. Considering that important cultural differences may be present, there are two ways to develop such specific measures: developing a new tool, or translating and cross-culturally adapting an existing tool previously validated in another language. Cross-cultural adaptation of an instrument is less costly and requires less time to obtain a common measure for the evaluation of disease status. In addition, it allows comparisons among different cultural groups. Adapting a questionnaire for use in a new, culturally distinct population is time-consuming and expensive. However, we believe that this is the best and most appropriate way to ensure equivalence. It will allow the collection of data in different countries or among individuals of different cultures within the same country, and avoid a selection bias.

**Keywords:** Surveys and questionnaires, quality of life, health care.

## INTRODUÇÃO

Um bom questionário para avaliação da evolução de uma enfermidade deve gerar respostas válidas, ser bem aceito pelos entrevistados e motivar a participação e o fornecimento das informações desejadas. Em virtude do crescente número de ensaios clínicos multicêntricos, há necessidade de se desenvolver medidas específicas

para utilização em países cujo idioma não é o inglês. Como diferenças culturais importantes podem estar presentes, essas medidas podem ser elaboradas de duas formas: desenvolver uma nova medida ou traduzir e adaptar culturalmente uma medida previamente validada em outro idioma<sup>1</sup>.

<sup>1</sup> Universidade Iguaçu UNIG, Nova Iguaçu, RJ.

<sup>2</sup> Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro HUCFF-UFRJ, Rio de Janeiro, RJ.

<sup>3</sup> Hospital Universitário Pedro Ernesto, Universidade Estadual do Rio de Janeiro HUPE-UERJ, Rio de Janeiro, RJ.

<sup>4</sup> Policlínica do Rio de Janeiro, Rio de Janeiro, RJ.

### Correspondência para:

Solange Oliveira Rodrigues Valle  
E-mail: solangervalle@gmail.com

Não foram declarados conflitos de interesse associados à publicação deste artigo.

Submetido em 01/05/2016,  
aceito em 2/12/2016.

A primeira opção demanda mais tempo e investimento, tanto pessoal como financeiro, para criação de um novo instrumento. Devem-se avaliar os conceitos, selecionar seus itens e testar sua validade. Além disso, é importante avaliar a real necessidade de se criar um novo instrumento quando já existe um desenvolvido anteriormente com o mesmo propósito e de boa qualidade. Segundo Casas-Anguita e cols., o desenvolvimento de um instrumento original deve ser sempre a última opção, dando-se prioridade aos já existentes<sup>2</sup>.

A adaptação cultural de um instrumento é menos onerosa e requer menor tempo para obtenção de uma medida comum para pesquisa do *status* da doença, além de proporcionar comparações entre grupos culturais diferentes.

A adaptação transcultural (ATC) possui dois componentes básicos: a tradução do instrumento e sua adaptação, ou seja, a combinação da tradução literal de palavras e sentenças de um idioma para o outro, e a adaptação ao contexto cultural e estilo de vida encontrados na cultura-alvo<sup>2,3</sup>.

### TRADUÇÃO E SÍNTESE DA TRADUÇÃO

Em relação à metodologia, a ATC se restringe à tradução e retradução (retrotradução) do instrumento. A perspectiva “universalista” não assume que os construtos são os mesmos em diferentes culturas, assim é necessário primeiro investigar se um conceito efetivamente existe ou se é interpretado da mesma forma na nova cultura, para depois se estabelecer sua equivalência transcultural por meio de metodologia própria. Com a necessidade de chegar a um método mais apurado, Herdman e cols. então propuseram um modelo para ATC que leva em conta não apenas a tradução do instrumento, mas também sua adaptação ao contexto cultural da população-alvo<sup>4</sup>.

O processo de ATC engloba a avaliação de seis tipos de equivalência: conceitual, de itens, semântica, operacional, de mensuração e funcional. Na equivalência conceitual, é avaliado se os conceitos mensurados através do instrumento que está sendo traduzido são semelhantes em ambas as culturas. O pesquisador deve ter conhecimento do conceito de saúde e doença na população a ser estudada e na população do questionário original, além de discussão com especialistas e com a população-alvo<sup>5</sup>.

A equivalência de itens ocorre quando os itens são capazes de aferir o mesmo parâmetro em ambas as culturas. A equivalência conceitual consiste na definição dos conceitos a serem estudados na população do país de origem e na população-alvo onde o instrumento será utilizado. A importância dos itens abordados varia entre as populações, e mesmo

perguntas sobre atividades consideradas universais podem estar abordando aspectos muito diferentes em diversas culturas. De acordo com Herdman e cols., para que se obtenha equivalência conceitual, os domínios do instrumento a ser adaptado devem ser avaliados quanto à sua pertinência na população-alvo e os itens devem ser investigados quanto à sua capacidade de refletir estes domínios nesta população. Em certas situações, questões ou partes delas podem necessitar substituições, entretanto estas medidas devem assegurar que os mesmos construtos estejam sendo mensurados<sup>4</sup>. Para realizar a equivalência dos itens, também é necessária a discussão com especialistas e com a população-alvo feita com base em publicações sobre as etapas envolvidas na elaboração do instrumento original<sup>5</sup>.

A equivalência semântica deve ser realizada em seis etapas: tradução, retrotradução, avaliação da tradução, avaliação da retrotradução, pré-teste do questionário e proposição final do instrumento. Esta compreende a tarefa mais complexa da tradução, e inclui o estudo do significado das palavras, avaliando o vocabulário e a gramática, ou seja, a transferência do significado através dos diferentes idiomas. Realiza-se a tradução, retrotradução, avaliação da equivalência semântica entre o original e a retrotradução, discussão com a população-alvo e com especialistas e pré-teste da nova versão<sup>6</sup>.

O processo de tradução linguística da versão original para a língua desejada é avaliado por uma equipe de trabalho formada por dois tradutores nascidos no país para o qual está sendo feita a tradução bilíngue na língua original. O questionário é traduzido de forma independente pelos dois tradutores, cientes dos objetivos da pesquisa, sendo enfatizada a necessidade da tradução ser conceitual, e não estritamente literária. A tradução deve essencialmente abranger os mesmos conceitos e itens que o instrumento original, com expressões que deveriam ser aceitas e culturalmente relevantes. As duas traduções são então convertidas na chamada versão de consenso 1 pela equipe do trabalho.

A versão de consenso 1 é então confrontada com a edição original por um profissional da área de saúde, fluente em ambos os idiomas, nascido no país de origem do questionário. A equivalência é avaliada de acordo com a seguinte classificação: inalterado, pouco alterado, muito alterado e completamente alterado, podendo ser utilizado o modelo de formulário empregado para examinar a equivalência semântica<sup>7</sup>.

### RETROTRADUÇÃO

A partir da tradução de consenso e sem conhecer a versão original do questionário, um tradutor cujo

idioma materno seja o da versão original realizará a retrotradução do instrumento. Este processo de validação tem por objetivo checar se a versão traduzida reflete o mesmo conteúdo das versões originais. Entretanto, a concordância entre a retrotradução e a versão original do instrumento não garante uma tradução satisfatória, pois esta pode ter sido realizada de maneira incorreta; esta apenas assegura uma tradução consistente<sup>8</sup>. A retrotradução é apenas uma maneira de validação, apontando inconsistências grosseiras ou erros conceituais da tradução<sup>3</sup>.

Alguns pesquisadores sugerem que a retrotradução também seja realizada por dois tradutores independentes cuja língua materna seja a do questionário original (Inglês, por exemplo)<sup>3</sup>.

### **AValiação DA TRADUÇÃO POR ESPECIALISTAS**

A formação de uma comissão de especialistas é de suma importância para se garantir uma equivalência transcultural. A comissão deve ser composta por profissionais de saúde (de preferência especialistas no tema a ser abordado), profissionais em linguística e os tradutores. Os desenvolvedores do questionário original devem permanecer em contato íntimo com a comissão de especialistas durante esta fase.

O papel da comissão de especialistas é agregar todas as versões do questionário e desenvolver aquela que será considerada a versão pré-final do questionário.

As decisões que devem ser tomadas pela comissão para atingir a equivalência entre o questionário original e a versão são as seguintes<sup>6</sup>:

- *Equivalência semântica*: compreende a tarefa complexa da tradução. Ela contempla o significado das palavras, avaliando-se o vocabulário e a gramática, ou seja, a transferência do significado através dos diferentes idiomas. As estratégias para sua realização são: tradução, retrotradução, avaliação da equivalência semântica entre o original e a retrotradução, discussão com a população-alvo e com especialistas e pré-teste da nova versão.
- *Equivalência idiomática*: coloquialismos são difíceis de ser traduzidos. A comissão precisa formular uma expressão equivalente.
- *Equivalência da experiência*: os itens devem capturar e avaliar as experiências da vida diária, entretanto em um país ou cultura diferentes, uma determinada atividade pode não ser realizada (mesmo que possa ser traduzida). O item no questionário deve ser substituído por um item semelhante que de fato seja realizado na cultura da versão a ser traduzida. Por exemplo: "você tem dificuldades para comer com um garfo?", quando na cultura a ser traduzida este utensílio não é utilizado.

- *Equivalência conceitual*: a importância dos itens abordados varia entre as populações, e mesmo termos usuais podem abordar aspectos muito diferentes em diversas culturas. A aceitabilidade das questões também pode variar, ou seja, certos aspectos podem ser agressivos em uma cultura ou considerados tabus em outras. Em certas situações, questões ou partes delas podem necessitar substituições.

### **PRÉ-TESTE**

O estágio final do processo da adaptação transcultural é o pré-teste, que compreende o teste de campo utilizando o novo questionário com a população-alvo. O número de indivíduos sugerido varia dependendo da literatura. Baiardini e cols. sugerem que no mínimo 10 pacientes participem desta fase, enquanto que Beaton e cols. relatam que o ideal seria testar 30-40 pacientes<sup>3,9</sup>.

Cada indivíduo preenche o questionário, e posteriormente deve ser entrevistado para avaliar cada item do questionário e suas respostas. Esta etapa garante que a versão adaptada mantém as características de avaliação do questionário original. A distribuição das respostas deve ser examinada para identificar itens não respondidos e respostas individuais.

Deve ser avaliado o percentual de pacientes que apresentarem os escores mais baixos e os escores mais altos em cada domínio do questionário, e também as perguntas que não vierem a ser respondidas<sup>10</sup>.

### **ACEITABILIDADE, VALIDADE E CONFIABILIDADE**

Na adaptação transcultural de um instrumento, após a tradução, retrotradução e aplicação do pré-teste, é realizado o estudo psicométrico, em que se deve estudar a aceitabilidade, confiabilidade e validade do instrumento em cada país, para apresentar qualidade internacional<sup>10</sup>.

#### **Aceitabilidade**

Na avaliação da aceitabilidade é calculado o percentual dos itens dos instrumentos avaliados que apresentaram os escores mais baixos (efeito piso) e os escores mais altos (efeito teto) e também os itens não respondidos. Efeito piso ou teto excedendo 15% é considerado alto. Efeitos piso e teto elevados apontam problemas na compreensão do instrumento e no cálculo do escore<sup>10</sup>.

#### **Validade**

A validade diz respeito ao grau de exatidão do resultado de uma medida. A validade de construto

avalia em que grau uma mensuração corresponde ao conceito teórico relativo ao fenômeno que está sendo estudado.

As técnicas de análise fatorial exploratória e confirmatória proporcionam evidências de validade para questionários comumente usados para diversos fins, sejam eles de respostas dicotômicas ou escalas tipo Likert.

A validade de construto interna pode ser medida através da análise de fatores exploratória (AFE) e confirmatória (AFC), que avaliam a estrutura dimensional do instrumento e a adequação dos itens<sup>10</sup>.

A análise fatorial exploratória analisa o padrão de correlações existentes entre as variáveis e utiliza esses padrões de correlações para agrupar suas variáveis em fatores, os quais são variáveis não observadas que se pretende medir a partir das variáveis observadas.

Na análise fatorial confirmatória há uma teoria prévia sobre quais variáveis medem quais fatores e é testado se a estrutura fatorial teórica se adequa aos dados observados. Para isso utiliza-se a modelagem de equações estruturais, na qual a relação entre variáveis observadas e fatores é modelada em uma série de regressões lineares, utilizando para isso uma matriz de covariância.

A análise de fatores é uma técnica que permite determinar se todas as variáveis estudadas (itens do instrumento) podem ser explicadas por um número menor de fatores. Quando um problema considera muitas variáveis, a análise utilizando cada variável individualmente pode levar a erros em virtude das inter-relações entre as variáveis não terem sido consideradas. A análise de fatores é uma técnica multivariada que considera estas correlações e tem como objetivos transformar as variáveis originais em novas variáveis (fatores, domínios ou dimensões) não correlacionadas e obter uma redução no número de variáveis, isto é, algumas poucas novas variáveis, sendo responsáveis por uma alta percentagem da variabilidade dos dados observados<sup>7,10</sup>.

A validade de construto externa pode ser investigada com base em critérios clínicos e demográficos derivados de experiência clínica e prática médica, e também através da correlação com outros instrumentos existentes, idealmente o padrão-ouro.

### **Confiabilidade**

A confiabilidade diz respeito à reprodutibilidade e à estabilidade da medida, sendo um índice da extensão na qual as medidas de indivíduos obtidas em diferentes circunstâncias levam a resultados semelhantes. O instrumento deve medir aquilo que se propõe de uma maneira reproduzível. É também uma forma de avaliar

a quantidade de erro inerente a qualquer medida. Há várias maneiras de avaliarmos a confiabilidade, entre elas, a consistência interna e a reprodutibilidade.

A consistência interna é baseada em uma única administração da medida e representa a média de correlações entre todos seus itens (por exemplo, entre as perguntas de um questionário). Ela reflete a extensão em que todos os itens que compõem uma medida refletem o mesmo construto subjacente. Ela pode ser aferida através do coeficiente alfa de Cronbach que é um importante indicador estatístico da homogeneidade de um instrumento psicométrico<sup>10</sup>. Quanto maior a correlação entre os itens, maior o valor do alfa de Cronbach<sup>7</sup>.

### **Reprodutibilidade**

A reprodutibilidade refere-se ao grau que um instrumento produz resultados estáveis com intervalo de tempo curto entre as duas medidas assumindo que não ocorreu nenhuma mudança clínica. Ela pode ser aferida através da avaliação da concordância entre dois ou mais observadores (confiabilidade interobservador) e da concordância entre as observações feitas por um mesmo observador, ou instrumento em diferentes ocasiões (confiabilidade intraobservador, teste-reteste respectivamente)<sup>11</sup>. Uma das abordagens mais comuns é a avaliação da reprodutibilidade intra e interobservador. Existem muitos métodos de análise, incluindo *inter alia*; as avaliações de concordância via *kappa*<sup>12-15</sup>; os estimadores *pos p* e *neg p* propostos por Cicchetti e Feinstein<sup>16</sup>; a enorme gama de correlações intraclasse<sup>17,18</sup>; os modelos log-lineares<sup>19-20</sup>; o coeficiente de concordância de Lin<sup>21</sup>; ou, ainda, o método de Bland & Altman<sup>22</sup>.

Com relação ao coeficiente kappa, admite-se: concordância quase perfeita para valores acima de 0,8; concordância substancial entre 0,61 e 0,80; concordância moderada entre 0,41 e 0,6; concordância regular entre 0,21 e 0,4; e concordância leve abaixo de 0,21<sup>23</sup>.

O coeficiente de correlação intraclasse (CCIC) pode ser utilizado para investigar a estabilidade temporal do instrumento através do método de teste-reteste, e valores > 0,75 e entre 0,4 e 0,75 indicam reprodutibilidade excelente e boa, respectivamente.

### **DISCUSSÃO**

O processo de ATC e validação é importante, pois a tradução literal poderia levar ao desenvolvimento de um instrumento não equivalente ao questionário original. A falta de equivalência limita a comparação de respostas entre populações separadas pela língua e pela cultura. Neste artigo, apresentamos um roteiro para o processo

de adaptação e validação de questionários para uso na área da saúde.

O processo descrito no artigo requer a tradução e, quando necessário, a substituição de itens ou escalas para tornar o instrumento relevante e válido na nova cultura. Os métodos sugeridos permitem a adaptação, porém vale a pena ressaltar que deve-se atentar não apenas para a equivalência de itens, mas para todos os tipos de equivalência. O ponto-chave do artigo é sinalizar que a tradução simples não fornece uma medida válida para a nova cultura, e deve passar cuidadosamente por todo o processo descrito.

A adaptação de um questionário para uso em uma nova população culturalmente distinta requer tempo e é dispendiosa. Entretanto, acreditamos que é a melhor e mais adequada maneira de garantir uma equivalência. Desta maneira, permite a coleta de dados em diversos países ou num mesmo país de indivíduos de culturas diferentes, evitando, assim, um viés de seleção.

## REFERÊNCIAS

1. WHOQOL Group 1995. The World Health Organization Quality of Life Assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med*. 1995;41(10):1403-9.
2. Casas-Anguitta J, Ramón Repullo Labrador J, Pereira Candel J. Medidas de calidad de vida relacionada con la salud. Conceptos básicos, construcción y adaptación cultural. *Med Clin*. 2001;116(20):789-96.
3. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186-91.
4. 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.
5. Reichenheim ME, Moraes CL. Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. *Rev Saude Publica*. 2007;41(4):665-73.
6. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-32.
7. Aires MT. Adaptação transcultural do “Vécu et Santé Perçue de l’adolescent”, um instrumento utilizado para a avaliação da qualidade de vida relacionada à saúde dos adolescentes. Rio de Janeiro, 2009. 147p. Tese de doutorado. Universidade do Estado do Rio de Janeiro, Instituto de Medicina Social.
8. Leplege A, Verdier A. The adaptation of health status measures: a discussion of certain methodological aspects of the translation procedure. In: *The International Assessment of Health-Related Quality of Life: Theory, Translation, Measurement and Analysis*. Oxford, UK: Rapid Communication of Oxford; 1995:93-101.
9. Baiardini I, Bousquet PJ, Brzoza Z, Canonica GW, Compalati E, Fiocchi A, et al. Recommendations for assessing patient-reported outcomes and health-related quality of life in clinical trials on allergy: a GA(2) LEN taskforce position paper. *Allergy*. 2010 Mar;65(3):290-5.
10. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32(1):40-66.
11. Fayers PM, Machin D. Sample size: how many patients are necessary? *Br J Cancer*. 1995 Jul;72(1):1-9.
12. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46.
13. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70(4):213-20.
14. Fleiss JL. *Statistical methods for rates and proportions*. 2nd. ed. New York: John Wiley & Sons, 1981. 321, xvii p.
15. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic: confidence interval construction, significance-testing and sample size estimation. *Stat Med*. 1992;11(11):1511-9.
16. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43(6):551-8.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-8.
18. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res*. 1998;7(3):301-17.
19. Goodman LA. Simple models for the analysis of association in cross-classification having ordered categories. *J Am Stat Assoc*. 1979;74(367):537-52.
20. Silva EF, Pereira MG. Avaliação das estruturas de concordância e discordância nos estudos de confiabilidade. *Rev Saude Publica*. 1998;32(4):383-93.
21. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255-68.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
23. Szklo M, Nieto FJ. Quality assure and control. In: Szklo M, Nieto FJ (ed.). *Epidemiology: Beyond the basics*. Maryland: Aspen Publishers, 2000. p. 297-348.