

Validação da versão em português europeu de questionário de avaliação funcional multidimensional de idosos

Rogério Manuel Clemente Rodrigues¹

Como citar Rodrigues RMC. Validação da versão em português europeu de questionário de avaliação funcional multidimensional de idosos. Rev Panam Salud Publica. 2008;23(2):109–15.

RESUMO

Objetivo. Validar a versão em português europeu do Older Americans Resources and Services, um instrumento para avaliar de forma multidimensional a qualidade de vida dos idosos.

Método. O estudo, autorizado pelos autores da versão original em inglês, começou com a tradução e a adaptação cultural do questionário para a língua portuguesa. A versão resultante foi validada por peritos e aplicada a 302 idosos (147 residentes em lares e 155 frequentadores de centros de dia) na região centro de Portugal (distritos de Aveiro, Castelo Branco, Coimbra, Leiria, Guarda e Viseu). Depois disso, foi feito o estudo psicométrico global do instrumento (coerência interna, validade da construção da versão em português europeu, validade de critério e reprodutibilidade). A amostra foi aleatória e estratificada por idade, gênero e tipo de apoio institucional.

Resultados. A coerência interna (alfa de Cronbach) para as cinco escalas de avaliação funcional do instrumento situou-se entre 0,64 e 0,91. Para a validade de critério, os valores relativamente baixos da correlação de Pearson ($< 0,700$) obtidos na correlação com o Medical Outcomes Study 36-Item Short-Form General Health Survey demonstram a dificuldade em dispor, na avaliação em saúde, de medidas de critério com as quais novos instrumentos possam ser comparados. No estudo da reprodutibilidade, os valores da correlação de Pearson situaram-se entre 0,859 e 1,000.

Conclusões. O instrumento de avaliação da qualidade de vida para idosos resultante deste trabalho — o Questionário de Avaliação Funcional Multidimensional de Idosos — é útil para levantar informações significativas que auxiliem a centrar as intervenções comunitárias e a promover estratégias de intervenção mais adequadas a esse grupo populacional.

Palavras-chave

Avaliação geriátrica, qualidade de vida, saúde pública, Older Americans Resources and Services Program, validação, Portugal.

O Older Americans Resources and Services Program (OARS) — um programa de recursos e serviços para idosos — é o resultado dos trabalhos desenvolvi-

dos, desde 1972, na universidade Duke, nos Estados Unidos da América (1). A metodologia OARS foi desenvolvida para avaliar a capacidade funcional em cinco áreas fundamentais para a qualidade de vida do idoso: recursos sociais, recursos econômicos, saúde mental, saúde física e atividades de vida diária. Mede, ainda, a utili-

zação e a necessidade percebida de vários tipos de serviços, o que permite a avaliação de programas de intervenção e a tomada de decisões fundamentadas sobre o impacto de recursos e serviços (1).

Os idosos estão sujeitos a uma multiplicidade de diagnósticos de saúde nos quais os aspectos físicos, mentais e

¹ Escola Superior de Enfermagem, Coimbra, Portugal. Correspondência: Rua Central 6-A, Espírito Santo das Touregas 3040/144, Coimbra, Portugal. E-mail: rogerio@esenfc.pt, rmcrc@sapo.pt

sociais estão fortemente interligados. Portanto, as medidas de avaliação multidimensionais são as mais apropriadas para essa população. De igual modo, as medidas de capacidade funcional que avaliam o grau de independência apesar da doença, das limitações físicas ou mentais e do apoio social são indicadores úteis para quem cuida de idosos (2).

Além disso, a revolução demográfica em curso, com o aumento crescente da população idosa, exige uma nova abordagem desse grupo etário (3), passando de um modelo centrado na doença para um modelo centrado nas necessidades das pessoas com incapacidades ou em risco de incapacidade (4). Em função dessas novas solicitações, as intervenções comunitárias tenderão a ser ajustadas às necessidades e à procura de serviços (5, 6).

Partindo desses pressupostos, o modelo conceitual do OARS inclui três elementos:

- avaliação da capacidade funcional, de forma a agrupar indivíduos com idêntica capacidade funcional;
- avaliação do uso e da necessidade de serviços, tendo em conta a diferenciação desses aspectos, definindo o seu uso atual e a necessidade percebida;
- avaliação de alternativas de serviços a implementar em função da capacidade funcional.

O instrumento desenvolvido para operacionalizar esse modelo conceitual, o questionário OARS, é uma das ferramentas mais importantes atualmente disponíveis para a avaliação de idosos (2). O presente estudo teve como objetivo validar a versão em português europeu do OARS.

MATERIAIS E MÉTODOS

O estudo, autorizado pelos autores do instrumento original, começou com a tradução e a adaptação cultural do OARS para português europeu.

O OARS é dividido em duas partes: a parte A (*Multidimensional Functional Assessment Questionnaire*, ou OARS-

MFAQ) permite a avaliação funcional multidimensional (áreas de recursos sociais e econômicos, saúde mental, saúde física e atividades de vida diária). A parte B (*Services Assessment Questionnaire*) avalia a utilização e a necessidade percebida de serviços (7). As duas partes podem ser utilizadas em conjunto ou separadamente. Para relacionar a informação de ambas as partes, estas devem ser administradas ao mesmo tempo, já que alguns itens da parte B se relacionam diretamente com a parte A (utilização de consultas médicas, por exemplo). No entanto, é possível a utilização apenas das questões específicas de cada área, como demonstra a validação da versão brasileira, em que só foram aplicadas as questões referentes à área de saúde mental (8). No presente estudo, ambas as partes foram validadas.

Avaliação funcional multidimensional

A parte A do OARS avalia cada uma das cinco áreas funcionais e permite uma classificação em escalas que posicionam os idosos avaliados segundo as suas incapacidades em cada área. É constituída por 101 perguntas (já que uma entre 102 questões se refere à parte B), das quais 70 são dirigidas ao idoso (questões 1 a 70), 10 são sobre o idoso, mas são respondidas por um informante (questões 72 a 82), e 14 são sobre o idoso, mas são respondidas pelo entrevistador (questões 83 a 96). Cada uma das cinco escalas funcionais apresenta pontuação de 1 a 6 (questões 97 a 101). A incapacidade global é determinada a partir do somatório das pontuações de incapacidade em cada uma das áreas (questão 102).

A aplicação da parte A é iniciada com um questionário preliminar, o *Short Portable Mental Status Questionnaire* (SPMSQ). De forma rápida, em 10 questões, o SPMSQ avalia o déficit da função cognitiva do avaliado, permitindo definir se está apto para continuar a responder o questionário ou se é necessário obter as respostas de um informante qualificado. Mais de quatro erros no SPMSQ indicam que o en-

trevistado não pode prestar informações. Nesse caso, o informante deve responder a todas as questões, com exceção das consideradas subjetivas, que questionam sentimentos dos sujeitos.

Quanto à avaliação dos recursos sociais, esta é centrada na quantidade e adequação da interação social e na disponibilidade de ajuda em caso de doença ou incapacidade. É composta por nove questões (questões 6 a 14) sobre estado civil, composição do agregado familiar, relações sociais, existência de confidente e sentimento de solidão e auto-avaliação de recursos sociais.

A avaliação dos recursos econômicos dirige-se aos rendimentos e à auto-avaliação da sua adequação às necessidades. É composta por 16 questões (questões 15 a 30) sobre a atual situação laboral, a ocupação principal e a ocupação principal do cônjuge, a origem e o valor dos rendimentos, o número de pessoas dependentes desses rendimentos, habitação, a relação rendimentos/gastos, a necessidade de ajuda econômica ou o aumento de rendimentos, a necessidade de ajuda econômica para compra de alimentos, a disponibilidade de seguro de saúde, a comparação da situação econômica com a de pessoas da mesma idade, auto-avaliação de recursos econômicos, a possibilidade de comprar pequenos extras e a perspectiva de rendimentos futuros.

A avaliação da saúde mental inclui os resultados do SPMSQ aplicado no início da parte A (9), uma avaliação da satisfação com a vida, sintomas psiquiátricos (10) (*Short Psychiatric Evaluation Schedule*, ou SPES, composto por 15 itens), auto-avaliação da saúde mental atual e auto-avaliação, por comparação, do estado atual com o existente há 5 anos.

A avaliação da saúde física inclui 19 questões (questões 37 a 55) sobre consultas médicas nos últimos 6 meses, dias de incapacidade física nos últimos 6 meses, dias de hospitalização nos últimos 6 meses, dias de internação noutra instituição de saúde nos últimos 6 meses, necessidade percebida de consultas médicas, uso de medicamentos com receita no último mês, doenças atuais e sua interferência nas ativida-

des diárias, incapacidades físicas, déficit visual e auditivo, uso e necessidade de dispositivos de apoio e próteses, problemas pelo consumo de álcool, prática de atividades físicas ou de esportes cansativos, auto-avaliação da saúde física, comparação do estado atual com o existente há 5 anos e interferência dos problemas de saúde nas atividades diárias (11).

A área de atividades de vida diária (AVD) inclui 15 questões (questões 56 a 70) sobre a capacidade para AVD instrumentais (telefonar, utilizar meios de transporte, fazer compras, preparar refeições, fazer as tarefas de casa, tomar os medicamentos, lidar com o dinheiro) e AVD físicas (comer, vestir-se e despir-se, cuidar da aparência, andar, levantar-se e deitar na cama, tomar banho ou ducha, manter continência de esfínteres) e sobre a existência de alguém disponível para ajudar.

Ao completar o questionário, o entrevistador emite um juízo quantitativo em função dos parâmetros definidos para cada uma das cinco áreas. Essa avaliação final, com a atribuição de uma pontuação de incapacidade funcional (PIF), permite enquadrar o sujeito no conjunto de características predefinidas para cada área, classificando-o com base em toda a informação recolhida ao longo da entrevista. Essa classificação é facilitada pelo conjunto de questões dirigidas ao entrevistador (questões 83 a 96). Assim, para cada escala, a capacidade funcional do idoso pode ser classificada na categoria 1 (muito bom), 2 (bom), 3 (ligeiramente insatisfatório), 4 (moderadamente insatisfatório), 5 (fortemente insatisfatório) e 6 (totalmente insatisfatório).

Os perfis funcionais caracterizam grupos funcionalmente equivalentes. Considera-se que todas as pessoas com um mesmo perfil têm idênticas incapacidades e, portanto, um funcionamento similar (7). A combinação da capacidade satisfatória com a capacidade insatisfatória, aplicada às cinco áreas, resulta em 32 possíveis combinações ou perfis funcionais, que vão desde capacidade satisfatória em todas as áreas a capacidade insatisfatória em todas as áreas.

Utilização e necessidade percebida de serviços

A parte B, o grupo de questões 71.1 a 71.19, avalia 19 serviços distribuídos em seis categorias: serviços gerais de apoio; serviços sociais e recreativos; serviços de saúde; serviços de apoio econômico; serviços de apoio, avaliação e coordenação; e serviços não classificados (7). Avalia-se a utilização atual (nos últimos 6 meses), a frequência com que são utilizados e a necessidade percebida. Aos 19 serviços são ainda adicionados quatro serviços já inseridos na parte A do questionário (consultas médicas, dispositivos de apoio e próteses, ajuda econômica em geral e ajuda para a compra de comida).

Validação da versão em português europeu e estudo psicométrico

A versão original do questionário OARS foi traduzida para português europeu. Posteriormente retrovertida, considerou-se, nos pontos de desacordo, a opinião de um conjunto de três juízes (portugueses com conhecimento da língua inglesa, experiência na utilização de escalas de medição em saúde e experiência profissional em cuidados de saúde primários e cuidados a idosos). Após alterações pontuais sugeridas, a escala, denominada Questionário de Avaliação Funcional Multidimensional de Idosos (QAFMI), foi considerada como apta a ser utilizada. Na elaboração desta versão foram mantidas todas as características da versão original. No questionário, cada conjunto de perguntas é precedido por uma nota introdutória que situa o respondente quanto ao tema que será abordado e à forma como deve fornecer a resposta. Não há limite de tempo para elaborar ou selecionar a resposta pretendida.

O QAFMI foi aplicado a 302 idosos (estudo psicométrico) com apoio institucional (amostra aleatória e estratificada por idade, gênero e tipo de apoio institucional) na região centro de Portugal (distritos de Aveiro, Castelo Branco, Coimbra, Leiria, Guarda e Viseu). Dos 302 idosos, 155 frequenta-

vam centros de dia e 147 moravam em lares para idosos. A coleta de informação foi autorizada pelas instituições e teve o consentimento expresso dos inquiridos.

Para o estudo do questionário, foram avaliadas a coerência interna, a validade da construção da versão em português europeu, a validade de critério e a reprodutibilidade.

Para comprovar a existência de coerência interna e a validade de construção da adaptação da versão portuguesa foi realizada a análise fatorial (alfa de Cronbach) dos itens subjetivos, agrupando-os tal como na versão original. Essa técnica de análise permite demonstrar a relação das diferentes dimensões com um conjunto de variáveis que lhe estão subjacentes (12). Um alfa de Cronbach superior a 0,80 indica uma boa consistência interna.

Uma dificuldade para avaliar a validade de critério reside no fato de não existirem, para a maioria das medidas utilizadas em ciências sociais, variáveis de critério significativas (12). Ainda, algumas das medidas com o estatuto de padrão apresentam, quando examinadas com detalhe, menor confiabilidade e validade que as desejadas (13).

O presente estudo adotou como medida de critério um instrumento multidimensional, o *Medical Outcomes Study 36-Item Short-Form General Health Survey* (MOS SF-36), utilizado para avaliar a qualidade de vida. Utiliza os conceitos de função física, desempenho físico, desempenho emocional, dor física, saúde mental, função social, vitalidade e saúde em geral. No seu conjunto, esses conceitos constituem dois componentes: físico e mental (12). É considerado como um dos questionários de avaliação, para a população em geral, com melhores perspectivas de aplicação no campo da saúde. Estudos de confiabilidade e de validade realizados com idosos (14) demonstram a sua capacidade de avaliação também com este grupo etário. Para a comparação entre o OARS e o MOS SF-36 utilizou-se a correlação de Pearson, sendo considerada boa uma correlação com valor superior a 0,700.

A reprodutibilidade do instrumento, ou verificação da estabilidade

TABELA 1. Análise de Cronbach dos itens subjetivos do *Older Americans Resources and Services Program* em três línguas

Itens subjetivos da escala OARS-MFAQ ^a (número da questão)	Inglês norte-americano (7) ^b	Espanhol europeu (15)	Português europeu (18)
Recursos sociais	(...)	0,69	0,64
Fator de interação (8 a 10)	0,56	0,61	0,63
Fator de confiança (11, 14)	0,69	0,59	0,80
Fator de sentimentos (12, 13)	0,71	0,47	0,32
Recursos econômicos	(...)	0,86	0,89
Percepção de recursos econômicos (21, 22, 27 a 30)	0,86	0,86	0,89
Saúde mental	(...)	0,88	0,83
Satisfação com a vida (32; 33; 34.2; 34.7; 34.15; 35)	0,70	0,80	0,81
Transtornos do sono (34.1; 34.6)	0,58	0,71	0,81
Fator de energia (34.5; 34.9; 34.10; 34.11; 34.13; 34.14)	0,71	0,76	0,65
Fator de paranóia (34.3; 34.4; 34.8)	0,52	0,58	0,51
Saúde física	(...)	0,77	0,85
Auto-percepção de saúde física (53, 55)	0,74	0,77	0,85
Atividades de vida diária	(...)	0,96	0,91
Fator instrumental (56 a 62)	0,87	0,94	0,89
Fator físico (63 a 68)	0,84	0,94	0,82

^a *Multidimensional Functional Assessment Questionnaire*.

^b Resultados globais não disponíveis.

dos resultados no tempo (15), foi avaliada através do método do teste repetido (teste-reteste). Para o estudo da reprodutibilidade foram selecionados 30 indivíduos em centros de dia, a quem foi aplicado o questionário na sua totalidade, em dois momentos, com intervalo de 1 semana. Esse intervalo de tempo foi selecionado (13) porque no espaço de 1 semana não ocorrem fenômenos capazes de alterar a percepção de uma população sobre o conjunto de áreas a avaliar, além de ser tempo suficiente para que as questões anteriormente formuladas e as respostas fornecidas não sejam lembradas. Os próprios indivíduos inquiridos não sabiam que, 1 semana após a primeira entrevista, seriam novamente confrontados com as mesmas questões.

O estudo incidiu sobre as questões consideradas chave pelos autores do instrumento original. Essas questões são as referentes aos valores atribuídos pelo entrevistador nas escalas de classificação final (questão 97 a 101) e sobre a auto-avaliação dos sujeitos nas áreas de recursos sociais, recursos econômicos, saúde mental e saúde física (não existe nenhuma pergunta de auto-avaliação para a área de AVD). Essas questões foram selecionadas

pelo fato de as escalas de classificação final funcionarem como resumo de toda a avaliação efetuada pelo entrevistador e pela possibilidade de avaliar a persistência das respostas a questões subjetivas como são a auto-avaliação nas áreas referidas.

RESULTADOS

Na tabela 1 são apresentados os resultados da análise de Cronbach para os fatores integrados em cada escala, comparando os resultados obtidos àqueles da versão original norte-americana (OARS-MFAQ) e da versão espanhola (OARS-MFAQ-VE). Para a versão original não dispomos do valor global de cada escala.

Na escala de recursos sociais, o valor global obtido (0,64) é inferior ao da versão espanhola. Os fatores de interação e confiança apresentam valores comparativos mais elevados, 0,63 e 0,80 respectivamente, sendo o fator sentimentos o mais baixo das três versões. Tomando como referência um alfa de 0,80 como indicativo de bons resultados de coerência interna, e considerando somente as escalas globais, observa-se um valor inferior para a escala global de recursos sociais.

Na escala de recursos econômicos, o alfa global obtido (0,89) é mais elevado do que nas outras versões. Na escala de saúde mental o valor global (0,83) é inferior ao da versão espanhola. Os fatores de satisfação com a vida e os transtornos do sono apresentam os alfas mais elevados das três versões (0,81). Os fatores de energia (0,65) e de paranóia (0,51) são comparativamente mais baixos, com o fator de paranóia muito próximo da versão original.

Na escala de saúde física, o estudo do fator de auto-percepção de saúde física apresenta o valor mais elevado das três versões (0,85). Na escala de atividades de vida diária o valor global obtido (0,91) é inferior ao da versão espanhola. O fator instrumental (0,89) é ligeiramente superior ao da versão original, mas inferior ao da versão espanhola. O fator físico (0,82) é inferior ao das outras versões, mas muito próximo ao da versão original.

A aplicação da versão em português europeu mostra os valores globais mais elevados na escala de recursos econômicos e saúde física, não apresentando as outras versões valores muito distintos. Numa análise geral das cinco escalas, e apesar das diferenças acima analisadas, a tendência é de acompanhar os valores das outras versões.

TABELA 2. Matriz da correlação de Pearson entre o *Medical Outcomes Study 36-Item Short-Form General Health Survey* e o Questionário de Avaliação Funcional Multidimensional de Idosos

MOS SF-36	Escala do Questionário de Avaliação Funcional Multidimensional de Idosos ^a								
	S	E	M	F	AVD	q.14	q.28	q.35	q.53
Função física	0,089	0,196	0,223	0,483	0,569	0,301	0,059	0,323	0,426
Desempenho físico	0,161	0,061	0,223	0,373	0,145	0,145	0,105	0,148	0,439
Dor física	0,161	0,292	0,460	0,397	0,114	0,304	0,193	0,458	0,620
Saúde em geral	0,224	0,194	0,495	0,402	0,130	0,361	0,039	0,469	0,655
Vitalidade	0,125	0,352	0,591	0,328	0,386	0,223	0,230	0,524	0,677
Função social	0,099	0,083	0,522	0,377	0,160	0,195	0,225	0,462	0,522
Desempenho emocional	0,298	0,199	0,325	0,119	0,051	0,082	0,242	0,085	0,302
Saúde mental	0,346	0,452	0,611	0,276	0,600	0,219	0,350	0,030	0,522

^a Versão em português europeu do *Older Americans Resources and Services Program*. AVD = atividades de vida diária; E = recursos econômicos; F = saúde física; M = saúde mental; q.14 = auto-avaliação de recursos sociais; q. 28 = auto-avaliação de recursos econômicos; q.35 = auto-avaliação de saúde mental; q.53 = auto-avaliação de saúde física; S = recursos sociais.

TABELA 3. Correlação de Pearson para a aplicação em dois tempos diferentes

Escala de classificação de recursos sociais (questão 97)	0,962
Escala de classificação de recursos econômicos (questão 98)	0,964
Escala de classificação de saúde mental (questão 99)	0,979
Escala de classificação de saúde física (questão 100)	1,000
Escala de classificação de desempenho de atividades de vida diária (questão 101)	1,000
Auto-avaliação de recursos sociais (questão 14)	1,000
Auto-avaliação de recursos econômicos (questão 28)	1,000
Auto-avaliação de saúde mental (questão 35)	0,859
Auto-avaliação de saúde física (questão 53)	0,866

Em geral, os resultados indicam que as três versões estão medindo os mesmos fenômenos. Os dados de que atualmente dispomos ainda não nos permitem uma explicação satisfatória para os mais baixos valores observados em algumas dimensões (mesmo quando comuns às três versões, como nos fatores interação e paranóia), ou para as diferenças entre as três versões (de que são exemplos os fatores confiança e sentimentos), apesar de os peritos estarem de acordo quanto ao conteúdo, a formulação, a pertinência e o agrupamento das questões que constituem estas dimensões.

Validade de critério

Os dados acerca da validade de critério foram obtidos a partir da mesma população e no primeiro momento em que foi utilizado o QAFMI para o

estudo da reprodutibilidade (teste-reteste), gerando a matriz da correlação de Pearson apresentada na tabela 2. Para o MOS SF-36 foram utilizados todos os oito conceitos, e para o QAFMI foram selecionadas as classificações finais (PIF) em cada uma das áreas avaliadas (recursos sociais, recursos econômicos, saúde mental, saúde física e AVD) e a auto-avaliação efetuada pelos respondentes nas áreas de recursos sociais (questão 14), recursos econômicos (questão 28), saúde mental (questão 35) e saúde física (questão 53).

Os resultados mostram que os valores mais elevados se encontram na correlação da dor física (0,620), da saúde em geral (0,655) e da vitalidade (0,677) com a auto-avaliação da saúde física. A função física (-0,569) e a saúde mental (-0,600) também se correlacionam com a PIF na área de AVD. A vitalidade (-0,591) e a saúde mental (-0,611)

correlacionam-se com a PIF na área de saúde mental. Esses resultados (com valores inferiores a 0,700 tomados como referência) reforçam a dificuldade já apontada de encontrar variáveis de critério significativas.

Reprodutibilidade

Verifica-se que, para as escalas de classificação de saúde física (1,000) e de desempenho de atividades da vida diária (1,000), não existiram diferenças de avaliação nos dois momentos, como demonstram os valores obtidos (tabela 3). Quanto às escalas de classificação de recursos sociais (0,962), recursos econômicos (0,964) e saúde mental (0,979), os valores obtidos são elevados, permitindo afirmar uma idêntica avaliação pelo entrevistador, do mesmo indivíduo, nos dois diferentes tempos.

Dos resultados obtidos verifica-se, para todas as perguntas selecionadas, que as respostas não foram influenciadas pelo intervalo de tempo decorrido entre a primeira e segunda medição.

DISCUSSÃO

A partir deste estudo, ficou disponível um instrumento de avaliação da qualidade de vida para idosos falantes de português europeu que pode fornecer informações significativas para centrar as intervenções comunitárias e promover mais adequadas estratégias

de intervenção nesse grupo populacional (16). A aplicação do QAFMI em outros países de língua portuguesa, como o Brasil, exigirá um estudo das particularidades de vocabulário, já que o estudo aqui apresentado se refere ao português escrito e falado em Portugal. A versão do OARS em português europeu pode ser solicitada, sem encargos, através do e-mail do autor deste artigo.

São várias as considerações acerca do processo de validação do OARS. Uma vantagem desse instrumento é que as questões são fechadas, o que torna comparáveis as respostas à mesma questão, com menos variabilidade. As questões são de mais fácil resposta e de mais fácil análise, codificação e informatização. O foco central na elaboração dessas questões é garantir a qualidade da informação verbal através da compreensão da natureza do comportamento pergunta/resposta, que reflete aquilo que o respondente pensa que o investigador quer saber e que pretende fazer com a informação fornecida.

Considera-se o período de 1 ano como o limite aceitável para a memória sobre acontecimentos com relevância para os respondentes, e 30 dias o limite credível para assuntos pouco relevantes (17). Neste questionário, o limite máximo de apelo à memória é de 6 meses, e o período mínimo reporta-se a 1 mês.

Para o estudo da confiabilidade de um instrumento, é necessária atenção à coerência interna, reprodutibilidade e concordância entre observadores. Ainda não foi possível o estudo da concordância entre observadores para a versão em português europeu, já que nos estudos efetuados (18–20) apenas se dispunha do autor para a realização das entrevistas. Entretanto, a informação disponível nesse campo (7, 13) demonstra que os resultados das avaliações nas diferentes áreas não são influenciados pelo entrevistador.

Quanto à validade do instrumento de medida (que inclui a validade de conteúdo, construção e critério), esta representa a capacidade para medir o que nos propomos. Um instrumento

pode ser válido para o efeito que levou ao seu desenvolvimento, mas não adequado para algo semelhante ou não equivalente. A validade deve ser analisada em função do contexto da utilização prática a dar ao instrumento (15). Tendo partido de um instrumento já validado em contextos sociais e culturais, quer distintos, quer próximos dos nossos, e com estudos profundos da validade de conteúdo e validade de construção, a nossa atenção centrou-se na avaliação da validade de critério, estudando a correlação de algumas das variáveis mais significativas com outra medida multidimensional do campo da saúde, não específica para determinado grupo etário ou patologia. Esses procedimentos levaram em conta as recomendações genéricas para estudos de validação de instrumentos de medição em saúde.

Por outro lado, não existem procedimentos padronizados para demonstrar a validade de conteúdo, sendo que no campo da saúde esta tarefa é ainda mais difícil. A necessidade de um consenso quanto aos domínios a incluir no estudo de cada dimensão torna aconselhável a existência de um número representativo de juizes capazes de propor e selecionar os mais significativos.

A validade de construção pretende demonstrar que a construção do instrumento foi baseada em paradigmas teóricos com verdadeira correspondência às observações. É dividida em validade convergente (requer correlação com variáveis relacionadas) e validade discriminante (requer que não haja correlação com variáveis não relacionadas) (12, 13). Pretende-se saber se um instrumento mede realmente aquilo que se pretende medir. O modelo de desenvolvimento do OARS foi feito por uma equipe multidisciplinar e baseado na adaptação de algumas escalas já existentes, criando as condições para um consenso entre técnicos de diversas áreas que se traduz por boas condições de validade (1, 7, 21).

A adaptação de instrumentos de medição em saúde já desenvolvidos e validados para diferentes contextos sociais e culturais é uma alternativa à

criação de novas formas de avaliação (12). A tradução do OARS da sua versão original para outras línguas tem-se mostrado uma tarefa não isenta de alguma complexidade, mas que se torna simplificada (apesar da sua extensão) pela linguagem acessível do instrumento. Essa característica é apontada como facilitadora da sua divulgação noutros países (15), comprovada pelas versões existentes em diversas línguas, incluindo as latinas (8, 13, 21).

Os resultados do estudo psicométrico efetuado com a aplicação do OARS na sua versão portuguesa pretendem demonstrar que essa versão é equivalente à versão original do instrumento. A análise psicométrica de um instrumento de medida, definindo a sua qualidade, compreende o estudo da confiabilidade e validade (22). Existem dois conceitos independentes no estudo da confiabilidade. Por um lado, a existência de coerência interna, implicando homogeneidade do conteúdo, e a não existência de erro aleatório. O segundo conceito diz respeito à reprodutibilidade, ou capacidade de obter uma mesma resposta em tempos diferentes (reprodutibilidade teste-reteste) (12). Quanto à consistência interna, e comparativamente às outras versões apresentadas (norte-americana e espanhola), os resultados são bons, e em linha com os obtidos por essas versões. Quanto à reprodutibilidade, os resultados obtidos neste estudo são excelentes, com algumas classificações e auto-avaliações a coincidir totalmente (1,000), sendo o valor mais baixo 0,859, como indica a tabela 3.

Em resumo, pode-se dizer que, pela metodologia, pela qualidade e pelo volume de informação coletada, o QAFMI pode desempenhar um importante papel no estudo dessa população, identificando problemas e orientando intervenções (23). A análise dos dados indica que podemos considerar a versão em português europeu uma versão equivalente à original, com resultados do estudo psicométrico idênticos à versão original e a outras versões em línguas latinas.

REFERÊNCIAS

- Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS Multidimensional Functional Assessment Questionnaire. *J Gerontol.* 1981;36(4):428-34.
- George LK, Fillenbaum GG. OARS methodology. A decade of experience in geriatric assessment. *J Am Geriatr Soc.* 1985;33(9):607-15.
- Kane RL. The public health paradigm. Em: Hickey T, Speers MA, Prohaska TR, eds. *Public health and aging.* Baltimore: Johns Hopkins University; 1997. Pp. 3-16.
- Kutza EA. Implications of an aging society for the organization and evaluation of public health services. Em: Hickey T, Speers MA, Prohaska TR, eds. *Public health and aging.* Baltimore: Johns Hopkins University; 1997. Pp. 293-307.
- Wacker RR, Roberto KA, Piper LE. Community resources for older adults: programs and services in an era of change. 2nd ed. Thousand Oaks: Pine Forge; 2002.
- Cox CB. Community care for an aging society: issues, policies, and services. New York: Springer; 2005.
- Fillenbaum GG. Multidimensional functional assessment of older adults. *The Duke Older Americans Resources and Services Procedures.* Hilldale: Lawrence Erlbaum; 1988.
- Blay SL, Ramos LR, Mari Jde J. Validity of a Brazilian Version of the Older Americans Resources and Services (OARS) Mental Health Screening Questionnaire. *J Am Geriatr Soc.* 1988;36(8):687-92.
- Liang J, Levin JS, Krause NM. Dimensions of the OARS mental health measures. *J Gerontol.* 1989;44(5):127-38.
- Gatz M, Pedersen NL, Harris J. Measurement characteristics of the mental health scale from the OARS. *Older Americans Resources and Services.* *J Gerontol.* 1987;42(3):332-5.
- Whitlaw NA, Liang J. The structure of the OARS physical health measures. *Med Care.* 1991;29(4):332-47.
- Ferreira PL, Marques FB. Avaliação psicométrica e adaptação cultural e lingüística de instrumentos de medição em saúde: princípios metodológicos gerais. Coimbra: Centro de Estudos e Investigação em Saúde; 1998.
- Patiño PE. Valoración multidimensional de la tercera edad en el municipio de Vigo. Santiago de Compostela: Universidad de Santiago de Compostela; 1994.
- Stadnyk K, Calder J, Rockwood K. Testing the measurement properties of the Short Form-36 Health Survey in a frail elderly population. *J Clin Epidemiol.* 1998;51(10):827-35.
- Wilkin D, Hallam L, Dogget MA. Measures of need and outcome for primary health care. New York: Oxford University; 1994.
- Ferreira PL, Rodrigues RMC, Nogueira DMS. Avaliação multidimensional em idosos. Coimbra: Mar da Palavra; 2006.
- Foddy W. Como perguntar: teoria e prática da construção de perguntas em entrevistas e questionários. Oeiras: Celta; 1996.
- Rodrigues RMC. Avaliação funcional multidimensional de idosos: estudo das incapacidades funcionais e utilização de serviços. Coimbra: Universidade de Coimbra; 1999.
- Nogueira DMS. Avaliação multidimensional de idosos. Estudo de uma população idosa não institucionalizada: caracterização das incapacidades funcionais e utilização de serviços. Lisboa: ISCTE; 2003.
- Silva SMDT. Avaliação multidimensional dos idosos inscritos no Centro de Saúde de Vale de Cambra. Coimbra: Universidade de Coimbra; 2005.
- Fibla GG, Patiño PE, Domínguez AC. Versión española del OARS Multidimensional Functional Assessment Questionnaire: adaptación transcultural y medida de la validez. *Atencion Primaria.* 1996;17(8):486-95.
- Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 2ª ed. Nova Iorque: Oxford University; 1995.
- Van Hook MP, Berkman B, Dunkle R. Assessment tools for general health care settings: Prime-MD, OARS, and SF-36. Primary care evaluation of mental health disorders. *Older Americans Resources and Services Questionnaire; Short Form-36.* *Health Soc Work.* 1996; 21(3):230-4.

Manuscrito recebido em 19 de novembro de 2006. Aceito em versão revisada em 20 de julho de 2007.

ABSTRACT

Validation of the European Portuguese version of the Older Americans Resources and Services instrument

Objective. To validate the European Portuguese version of the Older Americans Resources and Services (OARS) instrument, used for the multidimensional evaluation of the quality of life of the elderly.

Method. The study was authorized by the authors of the original English version. First, the questionnaire was translated into European Portuguese and culturally adapted. The resulting version was then validated by experts and administered to 302 elderly persons (147 living in nursing homes and 155 at day care centers) in the central region of Portugal (districts of Aveiro, Castelo Branco, Coimbra, Leiria, Guarda, and Viseu). Subsequently, a global psychometric study was conducted on the instrument, evaluating internal coherence, validity of construction, criterion validity, and reproducibility. The sample was randomly selected and stratified by age, sex, and type of institutional support.

Results. The internal coherence (Cronbach's alpha) was between 0.64 and 0.91 for each of the five functional assessment scales included in the instrument. For criterion validity, the relatively low Pearson correlation values obtained (< 0.700) through the Medical Outcomes Study 36-Item Short-Form General Health Survey reflect the challenges associated with establishing a criterion by which new health evaluation instruments might be compared. Pearson's correlation for reproducibility was between 0.859 and 1.0 for all the scales.

Conclusions. The European-Portuguese version of the OARS (*Questionário de Avaliação Funcional Multidimensional de Idosos*) is useful for collecting valid information that can help focus community interventions and promote strategies that specifically target this population group.

Key words

Geriatric assessment, quality of life, public health, Older Americans Resources and Services Program, validation, Portugal.