

Uma revisão sobre instrumentos de avaliação do estado funcional do idoso

A review of functional status evaluation instruments in the elderly

Carlos Montes Paixão Jr.¹
Michael E. Reichenheim¹

Abstract

The Comprehensive Geriatric Assessment is characterized as a functional and interdisciplinary evaluation. The first step is the assessment of functional status, which lies at the core of the procedure and can be defined by the level at which an individual performs roles and activities of daily living. Such assessment can be carried out by instruments (questionnaires) whose conceptual framework and psychometric properties must be well-defined. The present study systematically reviewed published data on the properties of the functional status measurement instruments and their adaptation and use in Brazil, through a search of the medical literature in reference textbooks and the LILACS and MEDLINE databases. After selection of the instruments, the next step consisted of a systematic scrutiny of validation, replication, and cultural adaptation studies. Thirty instruments were included according to explicit criteria. Only two, the Medical Outcomes Study SF-36 and the Health Assessment Questionnaire, were found to have equivalence studies in Brazil. However, some of these were apparently used in this country without any formal adaptation. The cultural adaptation of functional status measurement instruments is incomplete and inconsistent in Brazil.

Geriatric Assessment; Frail Elderly; Validity

Introdução

O envelhecimento da população mundial é um fato recente, universal e inexorável. Suas causas são multifatoriais e diferentes em países desenvolvidos e em desenvolvimento, mas suas conseqüências são igualmente importantes do ponto de vista social, médico e de políticas públicas¹.

Projeções do *Bureau of Census (US Department of Commerce, Economics and Statistics Administration, 1995)* indicam que os países em desenvolvimento abrigarão mais de 470 milhões de indivíduos acima de 65 anos em torno de 2020¹. Isto representará mais que o dobro da população com a mesma idade em países desenvolvidos. A prevalência geral de doenças crônicas aumentará como conseqüência, pois são mais freqüentes neste grupo etário, transferindo a ênfase dos programas governamentais de saúde e previdência do objetivo da cura e sobrevivência, para o da melhora do estado funcional e do bem-estar. Este fato possui implicações consideráveis quanto ao treinamento e recursos materiais.

Uma avaliação geriátrica eficiente e completa, a custos razoáveis, torna-se cada vez mais premente. Seus objetivos são o diagnóstico precoce de problemas de saúde e a orientação de serviços de apoio onde e quando forem necessários, com o fim de manter as pessoas nos seus lares. A história, o exame físico e o diag-

¹ Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondência

C. M. Paixão Jr.
Instituto de Medicina Social,
Universidade do Estado
do Rio de Janeiro.
Rua Nascimento Silva 423,
apto. 101, Rio de Janeiro, RJ
22421-020, Brasil.
carlospaixaojr@openlink.com.br

nóstico diferencial tradicionais não são suficientes para um levantamento extenso das diversas funções necessárias à vida diária do indivíduo idoso. A prática clínica geriátrica deve, se preocupada com a qualidade de vida, conter também uma ampla avaliação funcional em busca de perdas possíveis destas funções.

Vários pesquisadores têm repensado o modelo de atendimento à população mais idosa, e no bojo destas discussões surgem conceitos de atendimento interdisciplinar, tais como *Avaliação Geriátrica Ampla – AGA (Comprehensive Geriatric Assessment)*, que utiliza métodos de avaliação sistemática. Esta se baseia no princípio do trabalho interdisciplinar e planejado, visando a um tratamento e acompanhamento de longo prazo². A AGA é um conjunto de técnicas, procedimentos e ambientes operacionais, onde a avaliação abrangente e estruturada se associa aos métodos clássicos não padronizados de avaliação de saúde das diversas especialidades.

As dimensões mais estudadas e utilizadas na AGA se referem ao estado funcional, saúde mental e funcionamento social. O estado funcional é a dimensão-base para a avaliação geriátrica. Esta se resume na avaliação dos fatores físicos, psicológicos e sociais que afetam a saúde dos pacientes mais idosos e frágeis. Outras dimensões, como a avaliação social e cognitiva, são também necessárias na avaliação geriátrica, mas interagem de tal forma com o estado funcional, que uma alteração nesta dimensão pode indicar problemas nestas outras. A saúde mental possui duas subdimensões importantes – cognição e humor – que também interagem com a avaliação do estado funcional.

Os métodos habituais de se realizar uma avaliação funcional estruturada consistem na observação direta (testes de desempenho) e por questionários, quer auto-aplicados ou concebidos para entrevistas face a face, sistematizados por meio de uma série de escalas que aferem os principais componentes da dimensão. Tais escalas compõem o que se tem denominado “*instrumentos de avaliação funcional*”³.

Percorrendo a literatura específica, observa-se uma proliferação de instrumentos que se propõem a avaliar o estado funcional⁴. Também há algumas evidências importantes de previsibilidade de institucionalização e mortalidade dos indivíduos com distúrbios nesta dimensão da AGA⁴. Assim sendo, por cobrir uma área de preocupação fundamental nos cuidados de saúde do paciente idoso e da qual as outras dimensões dependem, essa dimensão requer um escrutínio sistemático do seu instrumental.

No Brasil, as questões relativas ao envelhecimento e à geriatria começam apenas a ser pontuadas. Na área de avaliação funcional permanece o uso assistemático de instrumentos. Possivelmente, pouco existe em termos de adaptação de instrumentos funcionais para o contexto brasileiro, fato que exige um aprofundamento. O presente estudo objetiva identificar os instrumentos mais utilizados da dimensão de estado funcional do paciente idoso na literatura de referência. Procura sobretudo avaliá-los quanto a seus estágios de validação, confiabilidade e adaptação para o contexto brasileiro, identificando as lacunas existentes, as necessidades de aprimoramento e a utilização mais formal desses instrumentos na avaliação funcional no Brasil.

Métodos

O processo envolveu inicialmente a escolha das subdimensões relacionadas ao estado funcional visando, subseqüentemente, a identificar os respectivos instrumentos de avaliação. Na terceira parte avaliou-se o histórico do programa de investigação de cada instrumento quanto aos estudos de validade, confiabilidade e utilização, tanto na literatura internacional quanto na brasileira.

As subdimensões foram classificadas seguindo uma tipologia consagrada na literatura afim. Estas se baseiam nos conceitos de deficiência (*impairment*), incapacidade (*disability*) e desvantagem (*handicap*) propostas pela Organização Mundial da Saúde (OMS). Quatro subdimensões podem ser delineadas^{5,6}. As *Atividades Básicas de Vida Diária* (AVD) correspondem aos níveis mais graves de deficiência das aptidões físicas, relevantes, em sua maior parte, aos pacientes frágeis e/ou institucionalizados. As AVDs avaliam funções de sobrevivência, tais como se alimentar, banhar-se, higiene pessoal, vestir-se, transferir-se de um local a outro, e outras atividades básicas. As *Atividades Instrumentais de Vida Diária* (AIVD) estendem o conceito acima, abarcando também problemas um pouco mais complexos da vida cotidiana de um indivíduo menos dependente. Oferecem indicadores de funções sociais que ampliam os temas explorados pelas escalas AVDs como, por exemplo, gerir o orçamento doméstico, utilizar o telefone, sair só, fazer compras etc. As *Atividades Avançadas de Vida Diária* (AAVD) incluem também atividades voluntárias sociais, ocupacionais e de recreação. Dificuldades em participar dessas atividades podem não indicar perda funcional atual, mas risco de perdas futuras ainda mais importantes^{4,7}.

A quarta subdimensão se refere ao *Estado Geral de Saúde e Qualidade de Vida ou Medidas de Qualidade de Vida Relacionada à Saúde* (EGS-QUAL), que inclui itens ligados a aspectos sociais e interpretativos da saúde do indivíduo.

A segunda parte do processo consistiu de uma busca em livros de referência sobre os instrumentos utilizados na identificação da dimensão e respectivas subdimensões^{5,6,8}. Realizou-se também uma busca bibliográfica nas bases de dados MEDLINE entre os anos de 1966 e 2000, e LILACS entre 1990 e 2000. Foram elegíveis todos os tipos de artigo relativos aos estudos realizados em seres humanos e de qualquer faixa etária. Interessaram os produzidos nas línguas inglesa, francesa, espanhola e portuguesa. Da base MEDLINE foram utilizados os seguintes descritores: *Geriatrics [and] Function; Geriatrics [and] Instruments; Geriatrics [and] Scale; Function; Physical Disability Measurements [and] Scale; Physical Disability and Handicap [and] Scale; Activities of Daily Living [and] Scale; Instrumental Activities of Daily Living [and] Scale; Advanced Activities of Daily Living [and] Scale; Measures of Health-related Quality of Life [and] Scale; General Health Status and Quality of Life [and] Scale; Scale; Physical Disability*. Para a busca no LILACS, se utilizaram os termos equivalentes em português.

A terceira parte do estudo se iniciou pela seleção dos instrumentos, com pesquisa bibliográfica efetuada nas mesmas bases e períodos usados na segunda parte. Utilizaram-se como palavras-chave e palavras-título os nomes de cada instrumento, tanto no idioma original como em português.

A seleção dos instrumentos obedeceu aos seguintes critérios: (a) serem, em princípio, de uso amplo e inespecífico, como, por exemplo, os que avaliam distúrbios funcionais em geral. Foram excluídos aqueles que se confinavam a um único ambiente ou doença específica. Por outro lado, aceitou-se os instrumentos que já haviam sido usados para além dos limites de suas propostas originais, mesmo que, estritamente, ainda não fossem de uso generalizado; (b) necessariamente terem sido publicados até 2000 em revistas indexadas, excetuando-se os instrumentos concebidos no Brasil; (c) já terem sido validados no idioma original; (d) serem de fácil compreensão e treinamento, além de curtos, contendo menos de 100 itens e/ou com duração de aplicação igual ou menor que uma hora; (e) serem de domínio público ou com direito de cópia sem ônus, disponibilizados mediante autorização prévia do autor.

Uma vez escolhidos os instrumentos, a segunda etapa da terceira parte envolveu um es-

crutínio sistemático dos estudos de validação, confiabilidade e adaptação transcultural dos instrumentos selecionados. Procurou-se apreciar a autoria, o tipo de escala utilizado, o número de itens constituintes, duração de aplicação, se concebido para aplicação clínica e/ou pesquisa, e quanto ao método de administração (auto-preenchimento, entrevista direta ou por *proximidade*).

Alguns critérios foram estabelecidos para a apreciação dos estudos de validade e confiabilidade, tomando-se as diretrizes descritas por Streiner & Norman⁹ como base. Em relação à validade conceitual, procurou-se identificar se havia um conceito ou teoria claramente explicitada pelos proponentes do instrumento. A validade de conteúdo foi também analisada com base nas descrições dos autores. A validade de face foi aceita caso houvesse menção da utilização de grupos de especialistas na sua avaliação. A validade de construto foi avaliada pela presença de análise fatorial e/ou correlações. A adequação da validade de critério baseou-se nos tradicionais estimadores de sensibilidade e especificidade, aceitando-se medidas acima de 0,75 como boas e valores acima de 0,85 como excelentes.

Nas avaliações formais de concordância em que o estimador kappa foi utilizado, seguiu-se a orientação de Shrout¹⁰, que propõe a seguinte classificação de concordância: ausente (virtualmente) se $\kappa < 0,10$; fraca (*slight*) se $\kappa = 0,10$ a $0,40$; discreta (*fair*) se $\kappa = 0,41$ a $0,60$; moderada se $\kappa = 0,61$ a $0,80$; e substancial se $\kappa > 0,80$. A qualificação das correlações intraclasse seguiu o mesmo padrão. Para o índice de consistência interna, valores entre 0,7 e 0,8 foram considerados bons. O coeficiente de Pearson possui limitações importantes quanto à adequação de seus valores, e estudos que os utilizaram não foram considerados adequados¹⁰.

Para delinear um quadro de síntese da apreciação dos instrumentos seguiu-se a orientação de McDowell & Newell⁵, que propõem dois modos de avaliação do histórico de cada instrumento. A *Completeza* concerne uma apreciação sobre o acúmulo e intensidade dos estudos de confiabilidade ou validade do instrumento na literatura. O segundo modo envolveu a apreciação qualitativa dos resultados. As categorias de avaliação de validade e confiabilidade encontram-se nas notas de rodapé das tabelas de resultados.

Resultados

Mediante os filtros de elegibilidade apresentados na seção dos métodos, chegou-se aos ins-

trumentos exibidos nas Tabelas 1, 2 e 3. Vale apontar que não foram encontrados instrumentos estruturados que avaliassem especificamente a subdimensão AAVD.

Na subdimensão AVD foram selecionados sete instrumentos, todos de língua inglesa (Tabela 1). Foram encontrados 482 estudos relacionados, sendo que 45 de validade e/ou confiabilidade. Os instrumentos mais utilizados são o *Barthel* e o *IADL*, que correspondem praticamente a 90,0% (450) dos artigos captados. Nenhum foi adaptado para o contexto brasileiro, embora o *Barthel* tenha sido utilizado em dois estudos no Brasil.

A Tabela 2 expõe os onze instrumentos de AIVD selecionados. Dois são em língua francesa e os outros em língua inglesa. Foram encontrados 572 artigos relacionados, sendo que mais de 80,0% (467) se referem ao *HAQ* e ao *FIM*. A dimensão *incapacidade física* ou *habilidade física* do *HAQ* foi adaptada para o contexto brasileiro¹¹. Todavia, o instrumento adaptado acabou, aparentemente, não sendo utilizado em

dois estudos brasileiros, ainda que as traduções, possivelmente realizadas informalmente, sejam citadas. Nenhum dos outros instrumentos desta subdimensão foi formalmente adaptado para o Brasil, apesar de alguns terem sido utilizados em nosso contexto cultural em estudos aplicados (*HAQ*, *Géronte*, *FIM*).

Treze instrumentos foram selecionados na subdimensão de EGS-QUAL (Tabela 3). A maioria é em língua inglesa, mas dois são concebidos em português para uso no Brasil. Foram encontrados 1.556 estudos relacionados, mais de 80,0% (1.276) referindo-se ao *NHP* e *SF-36*. Este último e o *MHIQ* foram utilizados no Brasil, mas apenas o *SF-36* apresenta um estudo de adaptação transcultural¹². O *C.I.C.Ac* ainda não possui uma edição definitiva, encontrando-se em uma fase preliminar de desenvolvimento. A concepção de sua versão preliminar é fruto de uma dissertação de Mestrado, na qual são desenvolvidos critérios de validade conceitual e de conteúdo¹³. Sua concepção utiliza perguntas abertas para um grupo definido de ido-

Tabela 1

Classificação de estudos sobre instrumentos de avaliação de atividades da vida diária (AVD).

Instrumentos	Autor	Periódico	Ano	Língua original	Validade/ confiabilidade	Utilização internacional	Utilização no Brasil	Adaptação para uso no Brasil	Total de estudos	
					n*	n*	n*	n*	n**	%***
The Pulses Profile (PULSES)	Moskowitz & McCann ²⁰	J Chronic Dis	1957	Inglês	5	6	0	0	11	2,3
The Barthel Index (BARTHEL)	Mahoney et al. ²¹	South Med J	1958	Inglês	20	334	2	0	356	73,7
The Index of Independence in Activities of Daily Living (IADL)	Katz et al. ²²	JAMA	1963	Inglês	9	85	0	0	94	19,5
The Kenny Self-Care Evaluation (KENNY)	Schoening et al. ²³	Arch Phys Med Rehabil	1965	Inglês	5	5	0	0	10	2,1
The Physical Self-Maintenance Scale (PSMS)	Lawton & Brody ²⁴	Gerontologist	1969	Inglês	3	3	0	0	6	1,2
The Functional Status Rating System (FSRS)	Forer ²⁵	–	1981	Inglês	1	1	0	0	2	0,4
Medical Outcomes Study Physical Functioning Measure (MOS)	Stewart & Kamberg ²⁶	****	1992	Inglês	2	1	0	0	3	0,6
Total					45	435	2	0	482	100,0

* Número de estudos sobre cada instrumento segundo categorias de classificação.

** Total de estudos encontrados sobre o instrumento.

*** Percentual de estudos sobre o instrumento com relação ao total de estudos encontrados na busca.

**** Stewart AL, Ware Jr. JE, editors. Measuring functioning and well-being: the Medical Outcomes Study approach. Durham: Duke University Press; 1992.

Tabela 2

Classificação de estudos sobre instrumentos de avaliação de atividades instrumentais da vida diária (AIVD).

Instrumentos	Autor	Periódico	Ano	Língua original	Validade/ confiabilidade	Utilização internacional	Utilização no Brasil	Adaptação para uso no Brasil	Total de estudos	
					n*	n*	n*	n*	n**	%***
Rapid Disability Rating Scale (RDRS)	Linn ²⁷	J Am Geriatr Soc	1967	Inglês	4	8	0	0	12	2,1
The Disability Interview Schedule (DIS)	Garrad & Bennett ²⁸	Br J Prev Soc Med	1971	Inglês	4	1	0	0	5	0,9
Functional Status Index (FSI)	Jette & Deniston ²⁹	J Chronic Dis	1978	Inglês	11	6	0	0	17	3,0
The Health Assessment Questionnaire (HAQ)	Fries et al. ³⁰	Arthritis Rheum	1980	Inglês	35	144	2	1	182	31,8
The Patient Evaluation Conference System (PECS)	Harvey & Jellinek ³¹	Arch Phys Med Rehabil	1981	Inglês	5	15	0	0	20	3,5
The OECD Long-Term Disability Questionnaire (OECD)	OECD ³²	Rev Epidemiol Sante Publique	1981	Inglês	7	10	0	0	17	3,0
The Lambeth Disability Screening Questionnaire (LAMBETH)	Patrick et al. ³³	J Epidemiol Community Health	1981	Inglês	5	2	0	0	7	1,3
Géronte	Léroux et al. ³⁴	Revue Geriatr	1981	Francês	3	0	1	0	4	0,7
The Functional Activities Questionnaire (FAQ)	Pfeffer et al. ³⁵	J Gerontol	1982	Inglês	3	2	0	0	5	0,9
The Functional Independence Measure (FIM)	Hamilton & Granger ³⁶	****	1987	Inglês	22	262	1	0	285	49,8
Le Système de Mesure de l'Autonomie Fonctionnelle (SMAF)	Hébert et al. ³⁷	Age Ageing	1988	Francês	3	5	0	0	8	1,4
Total					102	455	14	1	572	100,0

* Número de estudos sobre cada instrumento segundo categorias de classificação.

** Total de estudos encontrados sobre o instrumento.

*** Percentual de estudos sobre o instrumento com relação ao total de estudos encontrados na busca.

**** Uhrer MJ, editor. Rehabilitation outcomes: analysis and measurement. Baltimore: Paul H. Brookes; 1987.

sos atendidos em uma unidade básica de saúde, estabelecendo conceitos de autocuidado. Não se identificou qualquer estudo de validação ou utilização posterior à sua concepção.

O BOAS é um instrumento multidimensional utilizado em inquéritos populacionais, concebido no Brasil com base em vários instrumentos conhecidos, tais como o *Old Americans Resources and Services Multidimensional Functional Assessment Questionnaire* (OARS) e o

Short-Care, que possuem padrões aceitáveis de validade e confiabilidade ¹⁴. É dividido em nove seções, a saber: sobre informações gerais, saúde física, utilização dos serviços médicos e dentários, atividades de vida diária, recursos sociais, recursos econômicos, saúde mental, necessidades e problemas que afetam o entrevistado, além de uma sobre a avaliação do entrevistador. Juntas, objetivam assegurar informações sobre as características principais, ne-

Tabela 3

Classificação de estudos sobre instrumentos de avaliação de estado geral de saúde e qualidade de vida (EGS-QUAL).

Instrumentos	Autor	Periódico	Ano	Língua original	Validade/ confiabilidade	Utilização internacional	Utilização no Brasil	Adaptação para uso no Brasil	Total de estudos	
					n*	n*	n*	n*	n**	%***
Physical and Mental Impairment-of-Function Evaluation (PAMIE)	Gurel et al. ³⁸	J Gerontol	1972	Inglês	1	2	0	0	3	0,2
McMaster Health Index Questionnaire (MHIQ)	Chambers et al. ³⁹	Health Serv Res	1976	Inglês	10	14	1	0	25	1,6
Quality of Life Index (QLI)	Spitzer et al. ⁴⁰	J Chronic Dis	1981	Inglês	13	85	0	0	98	6,3
Functional Assessment Inventory (FAI)	Crewe & Athelstan ⁴¹	Arch Phys Med Rehabil	1981	Inglês	3	3	0	0	6	0,4
Nottingham Health Profile (NHP)	Hunt et al. ⁴²	Soc Sci Med	1981	Inglês	34	367	0	0	401	25,8
Self-Evaluation of Life Function Scale (SELF)	Linn & Linn ⁴³	J Gerontol	1984	Inglês	2	1	0	0	3	0,2
Functional Status Questionnaire (FSQ)	Jette et al. ⁴⁴	J Gen Intern Med	1986	Inglês	8	6	0	0	14	0,9
Brazil Old Age Schedule (BOAS)	Veras ⁴⁵	Tese	1986	Português	3(Parte)	0	1	0	4	0,3
COOP Charts for Primary Care Practice (COOP)	Nelson et al. ⁴⁶	J Chronic Dis	1987	Inglês	21	16	0	0	37	2,4
Short-Form-20 Health Survey (SF-20)	Stewart et al. ⁴⁷	Med Care	1988	Inglês	13	38	0	0	51	3,3
Short-Form-36 Health Survey (SF-36)	Ware Jr. & Sherbourne ⁴⁸	Med Care	1992	Inglês	54	818	2	1	875	56,4
DUKE Health Profile (DUKE)	Parkerson Jr. et al. ⁴⁹	Med Care	1990	Inglês	10	28	0	0	38	2,4
Instrumento para a Classificação de Idosos quanto à Capacidade para o Autocuidado em atividades cotidianas (C.I.C.Ac.)	Almeida ¹³	Tese	1997	Português	0	0	0	1	1	0,1
Total					169	1.378	3	2	1.556	100,0

* Número de estudos sobre cada instrumento segundo categorias de classificação.

** Total de estudos encontrados sobre o instrumento.

*** Percentual de estudos sobre o instrumento com relação ao total de estudos encontrados na busca.

cessidades e problemas da população idosa de grandes centros urbanos. O instrumento vem acompanhado por um manual de instrução e um livro de códigos. Tal qual o primeiro instrumento brasileiro citado acima, não se encontrou qualquer estudo de validação ou confiabilidade na sua forma integral, sendo apenas en-

contrados dois estudos de validação da parte da saúde mental. Em relação à sua utilização, foram encontrados dois estudos.

A Tabela 4 oferece uma apreciação mais detalhada dos instrumentos em cada subdimensão do estado funcional. Dezesete instrumentos, ou possuem duração de até dez minutos

Tabela 4

Comparação dos instrumentos de Estado Funcional ou Incapacidade Física.

Instru- mentos	Escala	Itens	Aplicação	Administrado por (tempo)	Tipo de análise*	Confiabilidade		Tipo de análise*	Validade		
						Comple- teza**	Qualifica- ção dos re- sultados***		Comple- teza**	Qualifica- ção dos re- sultados***	
AVD											
PULSES	Ordinal	6	Clínica	Staff	Ñ Revisão: Intra, Inter	Bas	Boa	Pr, Cc	Bas	Boa	
BARTHEL	Ordinal	10	Clínica	Staff (2-5min), Autopreenchi- mento (10min)	Ci, Intra, Inter	Exc	Exc	VC, F, Pr, Cc	Exc	Boa	
IADL	Ordinal	6	Clínica	Staff	Inter, G	Bas	Bas	VC, Pr, Cc	Boa	Boa	
KENNY	Ordinal	85	Clínica	Staff	Inter	Bas	Bas	VC, Cc	Bas	Boa	
PSMS	Gutmann	6	Inquérito, Clínica	Autopreenchi- mento, Staff	Inter, G	Bas	Boa	VC, Cc	Bas	Boa	
FSRS	Ordinal	30	Clínica	Staff (15-20min)	Inter****	Bas	Boa	Pr ****	Bas	0	
MOS	Ordinal	14	Inquérito	Autopreenchi- mento	Ci	Bas	Bas	VC, F, D	Bas	Boa	
AIVD											
RDRS	Ordinal	18	Pesquisa	Staff (2min)	Inter, Intra, Ci	Boa	Boa	F, Pr, Cc	Boa	Bas	
DIS	Ordinal	17	Clínica	Entrevistador	Inter	Bas	Bas	VC, Cc	Bas	Bas	
FSI	Ordinal	45	Clínica	Entrevistador (60-90min)	Inter, Intra, Ci	Exc	Boa	VC, F, Pr, Cc	Boa	Boa	
HAQ	Ordinal	20	Clínica, Pesquisa	Autopreenchi- mento, Staff (5-8min)	Inter, Intra, Ci	Exc	Exc	VC, F,Cc,Pr,D	Exc	Exc	
PECS	Ordinal	79	Clínica	Staff	Inter	Bas	Bas	Pr, Cc, D	Bas	Bas	
OECD	Ordinal	16	Inquérito	Autopreenchi- mento	Intra	Boa	Bas	VC, Cc, D	Boa	Bas	
LAMBETH	Ordinal	25	Inquérito	Autopreenchi- mento	Intra	Bas	?	VC, Pr, Cc	Boa	Boa	
Géronte	Ordinal	28	Clínica	Staff	Intra	Bas	Bas	D	Bas	Bas	
FAQ	Ordinal	10	Inquérito	Informante leigo	Ci	Bas	Bas	VC, Pr, Cc, D	Boa	Boa	
FIM	Ordinal	18	Clínica	Especialista, Entrevistador	Inter, Intra, Ci	Boa	Exc	VC, Cnt, F, Cc, Pr, D	Boa	Boa	
SMAF	Gutmann	29	Clínica	Staff (42min)	Inter, Intra	Bas	Boa	VC, Cc, Pr	Bas	Boa	

(continua)

ou contêm até vinte itens. Sete outros duram cerca de 15 minutos ou são constituídos de 20 a 40 itens. Os demais são mais longos. O *FSI* pode até, mas não necessariamente, exceder a sessenta minutos, um dos critérios de exclusão estipulados. Os autores do *KENNY*, *PECS*, *MHIQ*, *SELF* não informam sobre a duração destes instrumentos que contêm mais de quarenta itens. Os autores do *SF-20* e *SF-36* referem que a duração de aplicação especificamente em idosos gira em torno de 15 minutos.

Quinze dos instrumentos podem ser auto-administrados. Vinte e quatro instrumentos possuem aplicação clínica. Os instrumentos *PSMS*, *HAQ*, *MHIQ*, *NHP*, *FSQ* e *SF-36* podem ser utilizados em outros ambientes operacionais

além do clínico. Um instrumento, o *BOAS*, é de uso exclusivo para inquéritos epidemiológicos. Três instrumentos, todos da subdimensão de *EGS-QUAL*, requerem autorização para seu uso (*FAI*, *FAQ* e *SF-36*). O *FAQ* necessita de um *software* fornecido pelo autor. Quanto ao *SF-36*, cabe apenas solicitar autorização prévia.

Dos instrumentos *AVD*, o *Barthel* é o que possui resultados de confiabilidade e validade mais consistentes. Vários aspectos psicométricos são reportados, tais como confiabilidade intra e interobservador, consistência interna, validade conceitual, validade de critério, preditiva e concorrente. Vinte estudos de validade e confiabilidade foram encontrados, conforme indicado na Tabela 1. Segundo os critérios apre-

Tabela 4 (continuação)

Instrumentos	Escala	Itens	Aplicação	Administrado por (tempo)	Tipo de análise*	Confiabilidade		Tipo de análise*	Validade	
						Comple- teza**	Qualifica- ção dos re- sultados***		Comple- teza**	Qualifica- ção dos re- sultados***
EGS-QUAL										
PAMIE	Ordinal	77	Clínica	Staff (10-15 min)	Ci	Bas	Bas	VC, F, D	Bas	Bas
MHIQ	Ordinal	59	Clínica, Inquérito	Autopreenchi- mento	Inter, Intra, Ci	Boa	Boa	VC, D, Cc, Pr	Boa	Bas
QLI	Ordinal	5	Clínica	Autopreenchi- mento (2min)	Inter, Intra, Ci	Boa	Boa	VC, Cnt, D, Cc	Boa	Boa
FAI	Ordinal	40	Clínica	Staff	Inter, Ci	Bas	Boa	VC, F, D, Cc, Pr	Boa	Boa
NHP	Intervalar	45	Clínica, Inquérito	Autopreenchi- mento (10-15min)	Intra, Ci	Boa	Boa	VC, F, Cnt, D, Cc, Pr	Boa	Boa
SELF	Ordinal	54	Clínica	Autopreenchi- mento	Intra	Bas	Boa	F, D, Cc, Pr	Boa	Bas
FSQ	Ordinal	34	Clínica, Triagem	Autopreenchi- mento (15min)	Ci	Bas	Boa	VC, Cns, D, Cc, Pr	Boa	Boa
BOAS	Nominal	97	Inquérito	Entrevistador	0	0	0	VC	Bas	Bas
COOP	Ordinal	9	Clínica	Autopreenchi- mento (< 5min)	Inter, Intra, Ci	Boa	Boa	VC, VF, F, Cnt, D, Cc, Pr	Exc	Boa
SF-20	Ordinal	20	Inquérito	Autopreenchi- mento (3-4min) (16min idosos)	Intra, Ci	Boa	Boa	VC, F, D, Cc, Pr	Boa	Boa
SF-36	Ordinal	36	Inquérito, Clínica	Autopreenchi- mento (5-10 min) (15min idosos)	Intra, Ci	Exc	Exc	VC, F, D, Cc, Pr	Exc	Exc
DUKE	Ordinal	17	Clínica	Autopreenchi- mento	Intra, Ci	Boa	Boa	VC, VF, F, Cnt, D, Cc, Pr	Boa	Boa
C.I.C.Ac.	Gutmann	16	Clínica	Staff	0	0	0	VC, Cnt****	Bas	Bas

* Tipo de análise: Ñ = não especificado; Intra = confiabilidade intra-observador; Inter = confiabilidade inter-observador; Ci = consistência interna; G = análise de Gutmann; F = análise fatorial; D = validade de critério, tipo discriminante; Pr = validade de critério, tipo preditiva; Cc = validade de critério, tipo concorrente; VC = validade conceitual; Cnt = validade de conteúdo; Cns = validade de construto, sem especificação; 0 = sem evidência de confiabilidade ou validade. ** Completeza: (o quão completa foi a avaliação de confiabilidade ou validade do instrumento na literatura) 0 = sem evidência de confiabilidade ou validade; Bas = informação apenas basal; Boa = vários tipos de teste ou vários estudos; Exc = todas as principais formas de teste para confiabilidade/validade descritas. *** Qualificação dos resultados: 0 = sem resultados numéricos; ? = resultados não interpretáveis; Bas = confiabilidade/validade fracas; Boa = confiabilidade/validade adequadas; Exc = confiabilidade/validade excelentes (definição, vide texto). **** Versão preliminar.

sentados nos métodos, os resultados apontam para uma boa confiabilidade e excelente validade.

Já a maioria dos outros instrumentos AVD não possui estudos de validação e confiabilidade tão completos. A qualificação dos resultados encontrados foi considerada incompleta ou pouco satisfatória. Excetua-se o instrumento *IADL* de Katz, cujos estudos de validade são relativamente robustos e nos quais há menção de validade conceitual, preditiva e concorrente. A qualidade dos resultados de validade também foi classificada como boa. Entretanto, os estudos de confiabilidade desse instrumento são incompletos e qualificados como fracos. Vale lembrar que este é o segundo instrumento mais utilizado (Tabela 1).

Quanto ao *PSMS*, apesar de ser um instrumento pouco utilizado (Tabela 1), há indícios

de boa validade e confiabilidade. Uma parte deste, conhecida como *AIVD de Lawton*, foi extraída e modificada, sendo utilizada tanto na forma avulsa quanto no conjunto de outros instrumentos como, por exemplo, o *OARS*. A busca nas bases bibliográficas revelou 23 estudos de utilização, mas nenhum de validação do *AIVD de Lawton* fora do conjunto do *PSMS* ou do *OARS* (dado não mostrado).

Dos instrumentos da subdimensão *AIVD*, aqueles com históricos mais completos de avaliação de validade e confiabilidade são o *HAQ*, *FSI* e *FIM*. Estudos de confiabilidade do *HAQ* descrevem coeficientes em torno de 0,8; resultados de estudos de validade também são classificados como muito bons.

O *HAQ* é um instrumento de concepção multidimensional, mas o autor só desenvolveu as dimensões de incapacidade ou habilidade fisi-

ca, esta adaptada para o Brasil, e a de desconforto. No estudo de adaptação são mencionados coeficientes de confiabilidade teste-reteste, interobservador, validades de face e conteúdo. Os autores se preocupam com alguns aspectos semânticos no processo de adaptação transcultural, focalizando a tradução/retradução e os significados conotativos dos itens vis-à-vis o contexto cultural brasileiro. Questões relativas às equivalências conceitual, de itens, operacional e funcional não são estudadas, embora os autores relatem existir um estudo em curso para tal ¹¹.

O *FIM* possui vários estudos de validação e confiabilidade com bom poder estatístico, dentre estes, um com 11.102 sujeitos. Os estudos que avaliaram suas propriedades psicométricas – 22 estudos (Tabela 2) – apresentaram completeza e qualificações de boas a excelentes. Não foi encontrado estudo de adaptação para o Brasil, embora um tenha utilizado o instrumento em nosso contexto.

Os instrumentos EGS-QUAL são em sua maioria, os de melhor validade e confiabilidade, sendo, a rigor, os de concepção mais recente (Tabela 4). São, em geral, mais longos que os das outras subdimensões mais antigas. Os instrumentos com avaliações mais completas são: *SF-36*, *COOP*, *DUKE*, *SF-20*, *QLI* e *NHP*.

Merece destaque a adaptação do *SF-36* realizada por Ciconelli ¹² (Tabela 3). Como parte de um projeto internacional, a adaptação foi calçada em um roteiro explícito no qual se utilizaram métodos de tradução/retradução e validação de critério. A autora procurou interpretar formalmente o conteúdo do instrumento, concentrando-se no significado conotativo dos itens e não apenas em uma tradução literal. Para fins de validação de critério, comparou o *SF-36* ao *HAQ* já adaptado ¹¹, bem como ao instrumento *NHP* que não possui adaptação para o português. Pelo menos quatro estudos de validade do *SF-36* original utilizaram mais de cem pacientes, sendo que um, no qual também se avalia a confiabilidade, utilizou 3.445 sujeitos. Os estudos de validade e confiabilidade do *SF-36* foram qualificados como excelentes (Tabela 4).

Discussão

A dimensão de estado funcional é central para a avaliação geriátrica. Os instrumentos que avaliam são variados e com objetivos diversos. Podem ser utilizados como medidas de saúde geral ou específica; servir como instrumentos de pesquisa, triagem ou avaliação clínica; medir graus de dependência mais grave e até in-

capacidade para funções físicas e sociais. A escolha de um ou mais instrumentos depende do ambiente operacional e dos objetivos. No caso desta revisão, a escolha foi feita, em princípio, por instrumentos que tendiam a uma avaliação funcional geral e curta. Isto é necessário em um ambiente de AGA que utiliza, como descrito na introdução, uma série de outros instrumentos com objetivo de avaliação multidisciplinar de várias dimensões que afetam o indivíduo idoso. Instrumentos longos, neste caso, tornariam a AGA por demais extensa e pouco eficaz. De qualquer forma, deve ser enfatizado que a decisão de se utilizar um instrumento deve se embasar em evidências formais de validade e confiabilidade.

Limitações de espaço não permitem um aprofundamento sobre cada instrumento. O leitor pode encontrar detalhes em outra publicação recente ⁵. Ainda assim, algumas questões potencialmente limitantes do estudo merecem ser aventadas. Uma diz respeito à circunscrição do estudo às bases MEDLINE e LILACS, que podem não ter sido suficientes para alcançar alguns instrumentos. Vale assinalar, no entanto, que outras não trabalhadas sistematicamente, possivelmente, não teriam muito a contribuir. A base *Cochrane* não possui ainda informações sobre instrumentos funcionais da AGA. O *Psych Abstracts* é dirigido especificamente a outras dimensões. O *Index Medicus* tem sido cada vez menos utilizado e instrumentos exclusivamente referenciados nesta base tendem a ser obsoletos e pouco utilizados. Também, uma pesquisa sumária no banco *Current Contents* revelou que, para o assunto de interesse, suas bases bibliográficas eram redundantes em relação às da MEDLINE.

Um outro aspecto concerne à delimitação aos quatro idiomas vasculhados. Todavia, ainda que alguns instrumentos possam ter ficado de fora, tudo leva a crer que os mais importantes não deixaram de ser captados, pois os respectivos programas de desenvolvimento tendem a ser divulgados nas línguas selecionadas. Também merece ser comentado que a inclusão do *BOAS* não se encaixaria no âmbito específico da AGA por ser eminentemente para uso em inquéritos epidemiológicos. Decidiu-se incorporá-lo, já que outros instrumentos selecionados possuem perfil semelhante, ainda que não exclusivamente, como por exemplo o *SF-36*.

Dentre mais de sessenta instrumentos localizados inicialmente, um pouco menos da metade acabou selecionada. A maior parte foi excluída por não possuir evidências publicadas de validade e confiabilidade. Também foram desconsiderados alguns instrumentos para os

quais, apesar de existirem evidências psicométricas, não se encontrou outro estudo que não o de apresentação do próprio instrumento, o que, a rigor, não é tão raro acontecer com medidas de saúde.

Nas três subdimensões se encontram instrumentos de boa qualidade e de administração relativamente fácil. Ao se contrastar a Tabela 4 às três primeiras observa-se que, em geral, os instrumentos com maior número de estudos de utilização são os com melhores históricos e resultados de validade e confiabilidade. Esses instrumentos são, portanto, aqueles nos quais devem se concentrar os esforços de adaptação para o Brasil.

Na subdimensão AVD ressalta-se o *Barthel*. O *IADL* de Katz é muito conhecido, com bom histórico de validade, porém com resultados de confiabilidade ainda pouco consistentes. Na subdimensão AIVD, o *HAQ*, *FIM* e *FSI* são os instrumentos mais profícuos. Na subdimensão EGS-QUAL merecem destaque os instrumentos *SF-36*, *COOP*, *DUKE*, *SF-20*, *QLI* e *NHP*.

Reichenheim & Moraes¹⁵ apontam uma gama de quesitos que torna válida uma investigação epidemiológica, dentre os quais a qualidade das informações colhidas. Importa a adequação do processo de coleta, bem como do instrumento utilizado. Ressaltam que, se já é laborioso estabelecer a validade de um instrumento concebido em seu próprio contexto cultural e lingüístico, isto é ainda mais complexo quando este pretende ser utilizado em outro¹⁵. Usar um instrumento não adaptado formalmente pode invalidar todo um processo de investigação.

Os resultados do presente estudo apontam para a raridade e a pouca preocupação em se adaptar formalmente instrumentos na dimensão de estado funcional no Brasil. Parece haver uma utilização assistemática no contexto brasileiro. O *FIM*, *Barthel*, *Géronte*, *HAQ* e *MCMAS-TER* foram mencionados em estudos brasileiros, embora não haja referências sobre suas adaptações. Isto leva a suspeitar que tenham sido traduzidos sem formalidade ou, se o foram, os achados não estão publicados. O *SF-36*, embora se tenha identificado o estudo de adaptação de Ciconelli¹², foi utilizado em dois estudos sem menção dessa adaptação. Os dois instrumentos concebidos no Brasil (*C.I.C.Ac* e *BOAS*) não foram explorados formalmente do ponto de vista de validade e confiabilidade, o que reforça a idéia da pouca importância dada à avaliação de propriedades psicométricas, mesmo quando novos instrumentos estão sendo desenvolvidos.

Reichenheim & Moraes¹⁶, ao revisarem a questão da adaptação transcultural de instru-

mentos segundo vários autores, apontam que ainda não existe consenso quanto à melhor estratégia de adaptação. Diferentes abordagens teóricas são responsáveis pela diversidade de propostas metodológicas encontrada na literatura. Questões como o número de traduções necessárias, o perfil dos tradutores, o papel das retraduições, o que fazer quando essa etapa não tem bons resultados e como escolher entre termos semelhantes, ainda merecem reflexão e debate. Segundo os autores, somente baseando-se em avaliações sistemáticas dos resultados das diferentes estratégias existentes será possível propor um roteiro mínimo que seja útil nas investigações sobre diferentes temas e em contextos culturais distintos.

Ainda assim, o roteiro de investigação de equivalência transcultural proposto por Herdman et al.¹⁷ merece ser destacado pela sua clareza. Os autores propõem uma sistemática que envolve um aprofundamento cronologicamente ordenado de vários subtipos de equivalência. Segundo este modelo, inicialmente, a adaptação deve contemplar uma apreciação da pertinência dos conceitos e dimensões apreendidos pelo instrumento original na cultura alvo da nova versão (equivalência conceitual). O modelo também prescreve que se avalie a adequação de cada item do instrumento original em termos de sua capacidade, para representar tais conceitos na população onde o instrumento pretende ser utilizado (equivalência de itens). Somente então, deve-se partir para a avaliação da equivalência semântica entre esta e o original (equivalência semântica). A adequação e pertinência de aspectos operacionais na nova cultura – forma de administração, número de opções de resposta etc. – também devem ser observadas (equivalência operacional). Na seqüência, avalia-se a equivalência entre as propriedades psicométricas do instrumento original e de sua nova versão (equivalência de mensuração). Segundo os autores, somente após percorrer os diferentes aspectos de equivalência apontados acima, é possível declarar que está estabelecida a adaptação transcultural da nova versão.

Retomando a apreciação dos instrumentos, Ciconelli¹², no estudo de adaptação do *SF-36*, segue uma perspectiva mais próxima do que Herdman et al.¹⁸ qualificam de *absolutista*. Esta parte do princípio de que a cultura possui um impacto mínimo no construto medido, sendo, conseqüentemente, invariante de um contexto para outro. Ciconelli utiliza critérios clínicos de artrite reumatóide para trabalhar a validade construto, além de comparações com outros dois instrumentos já adaptados para o

uso no Brasil – HAQ¹¹ e *Arthritis Impact Measurement Scale*¹² – e a um terceiro desprovido de adaptação formal para o português (NHP). Dois aspectos merecem reflexão. O NHP não foi adaptado formalmente para uso no Brasil. O HAQ¹¹, por sua vez, se restringiu à validação de conteúdo e avaliação de confiabilidade teste-reteste e inter-observador. Reconhecendo que esses instrumentos têm bastante difusão internacional, valeria recomendar um maior aprofundamento das respectivas agendas de investigações transculturais.

Não se pode perder de vista que instrumentos de avaliação como escalas ou testes de desempenho têm inserções e finalidades bem definidas, e não devem substituir avaliações não padronizadas como é o caso da anamnese e exame físico. Estas permitem uma maior flexibilidade e oportunidade de avaliar de maneira mais profunda e rápida respostas não esperadas⁴. Por outro lado, as medidas padronizadas possuem a vantagem de ser menos suscetíveis a vieses, mais acuradas e reprodutíveis, mesmo quando utilizadas por equipes com diferentes níveis de treinamento. O uso de um teste padronizado permite ao profissional de saúde estabelecer uma linha de base bem definida sobre a qual futuras decisões podem ser calcadas. A identificação de modificações na função pelo uso desses testes em reavaliações periódicas permite discernir melhor se algum aprofundamento propedêutico deve ou não ser realizado, ou mesmo para propor uma ação imediata na eventualidade de um declínio funcional importante. Um grande benefício desse tipo de avaliação estruturada é a atenção às variações patológicas que, na sua ausência, poderiam passar despercebidas.

No capítulo final de seu livro sobre instrumentos funcionais, McDowell & Newell⁵ sugerem uma série de normas para o desenvolvimento de medidas padronizadas de saúde. Referem que não há motivos para que a saúde seja aferida de forma menos acurada que outros campos do conhecimento. Os autores propõem que se usem métodos mais novos nas concepções dos instrumentos, o desenvolvimento de um corpo de especialistas para oferecer orientações sobre assuntos pertinentes e, se possível, a promoção de uma publicação periódica específica para esta matéria.

Por fim, é importante que se indague sobre a real necessidade de se conceber novos instrumentos e se uma possível pletoira não limita uma das vantagens das medidas sistematizadas de saúde que é a garantia de comparabilidade de resultados. Hedrick et al.¹⁹ sugerem que deva haver progresso continuado na procura de medidas uniformes, objetivando uma maior comparabilidade e reprodutibilidade de dados em preferência a revisões e reinvenções sem fim. Essas preocupações não devem passar ao largo do Brasil, principalmente no que concerne ao investimento em adaptações transculturais permanentes e de qualidade.

Um “mapeamento” mais sistemático e elaborado de instrumentos de avaliação de idoso servirá de auxílio na tomada de decisões mais robustas sobre quais estudar de forma mais aprofundada. A partir daí lançar-se-ão bases para recomendações subseqüentes sobre quais instrumentos utilizar em nosso meio. O desenvolvimento do instrumental de aferição relativo ao idoso e à AGA em particular, certamente se beneficiará desta perspectiva.

Resumo

A avaliação geriátrica se baseia em diagnósticos funcionais interdisciplinares, sendo correntemente denominada Avaliação Geriátrica Ampla. Inicia-se pela avaliação do estado funcional, dimensão central neste contexto. Este pode ser definido como o nível com o qual a pessoa desempenha funções e atividades da vida diária. São utilizados instrumentos que, por sua vez, necessitam estar bem estudados. Importam questões psicométricas como validade e confiabilidade. Este estudo procurou revisar o histórico, utilização e adaptação dos instrumentos de avaliação funcional no âmbito internacional e no Brasil. Utilizou-se uma revisão sistemática da literatura médica em livros, textos e nas bases de dados LILACS e MEDLINE. Escolhi-

dos os instrumentos, realizou-se, em seguida, um escrutínio sistemático dos estudos de validação, confiabilidade e adaptação transcultural dos instrumentos selecionados. Dos trinta instrumentos escolhidos por critérios explícitos, apenas dois, o Medical Outcomes Study SF-36 e o Health Assessment Questionnaire, possuem adaptação para o português. Entretanto, alguns dos instrumentos revisados vêm sendo utilizados em nosso meio sem adaptação formal prévia. Conclui-se que a adaptação de instrumentos de avaliação do estado funcional é incompleta e pouco sistematizada no Brasil.

Avaliação Geriátrica; Idoso Débil; Validade

Colaboradores

C. M. Paixão Jr. participou das seguintes etapas: escolha do tema, organização do trabalho, montagem das tabelas de revisão, procura e revisão das publicações, escolha dos instrumentos para revisão, discussão, organização do artigo, revisão e edição final do artigo. M. E. Reichenheim participou das etapas: escolha do tema, organização do trabalho, discussão, organização e correção do artigo, revisão e edição final do artigo.

Referências

1. Kinsella K. Demographic aspects. In: Ebrahim S, Kalache A, editors. *Epidemiology in old age*. London: BMJ Publishing Group; 1996. p. 32-40.
2. Rubenstein LZ, Wieland D, Bernabei R. *Geriatric assessment technology: state of the art*. Milano: Editrice Kurtis; 1995.
3. Rubenstein LV, Calkins DR, Greenfield S, Jette AM, Meenan RF, Nevins MA, et al. Health status assessment for elderly patients. Report of the Society of General Internal Medicine Task Force on Health Assessment. *J Am Geriatr Soc* 1989; 37:562-9.
4. Hedrick SC. Assessment of functional status: activities of daily living. In: Rubenstein LZ, Wieland D, Bernabei R, editors. *Geriatric assessment technology: state of the art*. Milano: Editrice Kurtis; 1995. p. 311.
5. McDowell IW, Newell C. *Measuring health: a guide to rating scales and questionnaires*. 2nd Ed. New York: Oxford University Press; 1996.
6. Hazzard WR, Blass JP, Ettinger WHJ, Halter JB, Ouslander JG. *Principles of geriatric medicine and gerontology*. 4th Ed. New York: McGraw-Hill; 1999.
7. Reuben DB, Wieland D, Rubenstein LZ. Functional status assessment of older persons: concepts and implications. *Facts and Research in Gerontology* 1993; 7:231-40.
8. Wade DT. *Measurement in neurological rehabilitation*. Oxford: Oxford University Press; 1996.
9. Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use*. 2nd Ed. New York: Oxford Medical Publications; 1995.
10. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998; 7:301-17.
11. Ferraz MB, Oliveira LM, Araújo PM, Atra E, Tugwell P. Cross-cultural reliability of the physical ability dimension of the Health Assessment Questionnaire. *J Rheumatol* 1990; 17:813-7.
12. Ciconelli RM. Tradução para o português e validação do questionário genérico de avaliação de qualidade de vida "Medical Outcomes Study 36-Item Short-Form Health Survey (Sf-36)" [Tese de Doutorado]. São Paulo: Universidade Federal de São Paulo; 1997.
13. Almeida MHMD. C.I.C.Ac.: instrumento para classificação de idosos quanto à capacidade para o autocuidado [Dissertação de Mestrado]. São Paulo: Universidade de São Paulo; 1997.
14. Veras RP, Souza CAM, Cardoso RF, Milioli R, Silva SDD. Pesquisando populações idosas. A importância do instrumento e o treinamento de equipe: uma contribuição metodológica. *Rev Saúde Pública* 1988; 22:513-8.
15. Reichenheim ME, Moraes CL. Alguns pilares para a apreciação da validade de estudos epidemiológicos. *Rev Bras Epidemiol* 1998; 1:131-48.
16. Reichenheim ME, Moraes CL. Buscando a qualidade das informações em pesquisas epidemiológicas. In: Minayo MCS, Deslandes SF, organizadoras. *Caminhos do pensamento: epistemologia e método*. Rio de Janeiro: Editora Fiocruz; 2002. p. 227-54.
17. 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; 7:323-35.
18. Herdman M, Fox-Rushby J, Badia X. Equivalence and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res* 1997; 6:237-47.
19. Hedrick SC, Katz S, Stroud MW. Patient assessment in long-term care: is there a common language? *Aged Care Serv Rev* 1980/1981; 2:1-19.
20. Moskowitz E, McCann CB. Classification of disability in the chronically ill and aging. *J Chronic Dis* 1957; 5:342-6.
21. Mahoney FI, Wood OH, Barthel DW. Rehabilitation of chronically ill patients: the influence of complications on the final goal. *South Med J* 1958; 51:605-9.
22. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185:914-9.
23. Schoening HA, Anderegg L, Bergstrom D, Fonda M, Steinke N, Ulrich P. Numerical scoring of self-care status of patients. *Arch Phys Med Rehabil* 1965; 46:689-97.
24. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; 9:179-86.

25. Forer S. Revised functional status rating instrument. Glendale: Rehabilitation Institute, Glendale Adventist Medical Center; 1981.
26. Stewart AL, Kamberg CJ. Physical functioning measures. In: Stewart AL, Ware Jr. JE, editors. Measuring functioning and well-being: the Medical Outcomes Study approach. Durham: Duke University Press; 1992. p. 86-101.
27. Linn MW. A rapid disability rating scale. *J Am Geriatr Soc* 1967; 15:211-4.
28. Garrad J, Bennett AE. A validated interview schedule for use in population surveys of chronic disease and disability. *Br J Prev Soc Med* 1971; 25:97-104.
29. Jette AM, Deniston OL. Inter-observer reliability of a functional status assessment instrument. *J Chronic Dis* 1978; 31:573-80.
30. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23:137-45.
31. Harvey RF, Jellinek HM. Functional performance assessment: a program approach. *Arch Phys Med Rehabil* 1981; 62:456-60.
32. McWhinnie JR. Disability assessment in population surveys: results of the O.E.C.D. Common Development Effort. *Rev Epidemiol Sante Publique* 1981; 29:413-9.
33. Patrick DL, Darby SC, Green S, Horton G, Locker D, Wiggins RD. Screening for disability in the inner city. *J Epidemiol Community Health* 1981; 35:65-70.
34. Leroux R, Viau G, Fournier M, Bergeot R, Attali G. Visualisation d'une échelle simple d'autonomie: géronte. *Revue Geriatr* 1981; 6:433-6.
35. Pfeffer RI, Kurosaki TT, Harrah Jr. CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982; 37:323-9.
36. Hamilton BB, Granger C. A uniform national data system for medical rehabilitation. In: Uhrer MJ, editor. *Rehabilitation outcomes: analysis and measurement*. Baltimore: Paul H. Brookes; 1987. p. 137-47.
37. Hebert R, Carrier R, Bilodeau A. The Functional Autonomy Measurement System (SMAF): description and validation of an instrument for the measurement of handicaps. *Age Ageing* 1988; 17:293-302.
38. Gurel L, Linn MW, Linn BS. Physical and mental impairment-of-function evaluation in the aged: the PAMIE scale. *J Gerontol* 1972; 27:83-90.
39. Chambers LW, Sackett DL, Goldsmith CH, MacPherson AS, McAuley RG. Development and application of an index of social function. *Health Serv Res* 1976; 11:430-41.
40. Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Shepherd R, et al. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. *J Chronic Dis* 1981; 34:585-97.
41. Crewe NM, Athelstan GT. Functional assessment in vocational rehabilitation: a systematic approach to diagnosis and goal setting. *Arch Phys Med Rehabil* 1981; 62:299-305.
42. Hunt SM, McKenna SP, McEwen J. The Nottingham Health Profile: subjective health status and medical consultations. *Soc Sci Med* 1981; 15:221-9.
43. Linn MW, Linn BS. Self-evaluation of life function (self) scale: a short, comprehensive self-report of health for elderly adults. *J Gerontol* 1984; 39:603-12.
44. Jette AM, Davies AR, Cleary PD, Calkins DR, Rubenstein LV, Fink A, et al. The Functional Status Questionnaire: reliability and validity when used in primary care. *J Gen Intern Med* 1986; 1:143-9.
45. Veras RP. A survey of the health of elderly people in Rio de Janeiro, Brazil [PhD Thesis]. London: London University; 1992.
46. Nelson E, Wasson J, Kirk J, Keller A, Clark D, Dietrich A, et al. Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings. *J Chronic Dis* 1987; 40 Suppl 1:55S-69S.
47. Stewart AL, Hays RD, Ware Jr. JE. The MOS short-form general health survey. Reliability and validity in a patient population. *Med Care* 1988; 26:724-35.
48. Ware Jr. JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
49. Parkerson Jr. GR, Broadhead WE, Tse CK. The Duke Health Profile. A 17-item measure of health and dysfunction. *Med Care* 1990; 28:1056-72.

Recebido em 10/Fev/2004

Aprovado em 25/Ago/2004