

POPULAÇÃO IDOSA NO RIO DE JANEIRO (BRASIL): ESTUDO-PILOTO DA CONFIABILIDADE E VALIDAÇÃO DO SEGMENTO DE SAÚDE MENTAL DO QUESTIONÁRIO BOAS*

Renato P. Veras**
Evandro Coutinho***
George Ney Jr.****

VERAS, R. P. et al. População idosa no Rio de Janeiro (Brasil): estudo-piloto da confiabilidade e validação do segmento de saúde mental do questionário BOAS. *Rev. Saúde públ.*, S. Paulo, 24: 156-63, 1990.

RESUMO: São discutidos os procedimentos propostos no estudo-piloto de confiabilidade e validação do segmento de saúde mental do questionário BOAS. Sugere-se a necessidade de se ajustar os instrumentos numa primeira etapa pela análise pormenorizada de cada entrevista, com o propósito de prepará-los para a etapa posterior de um estudo de confiabilidade e validação, onde prevalecerá uma abordagem de caráter quantitativo. São propostos alguns delineamentos para estudos de confiabilidade e validade para pesquisa epidemiológica.

DESCRITORES: Idoso. Saúde mental. Métodos epidemiológicos. Questionários. Validez.

INTRODUÇÃO

Os problemas da velhice normalmente estão associados aos países desenvolvidos. É neles que a proporção da população idosa**** já atingiu, há várias décadas, níveis bastante expressivos. No entanto, não vem sendo dada a devida atenção ao fato de que desde o início da década de 70 mais da metade das pessoas que chegam aos sessenta anos vive em países do Terceiro Mundo. As projeções demográficas indicam ainda que, de 1980 até o final do século, cerca de três quartos do aumento da população idosa ocorrerão em tais países, fazendo com que este seja o grupo etário que mais crescerá na maioria dos países menos desenvolvidos^{4, 11, 15}. A longo prazo as perspectivas são ainda mais impressionantes. De acordo com os critérios atuais que definem os países quanto ao seu grau de desenvolvimento, estima-se que num prazo de 40 anos 8 países, hoje tidos como em desenvolvimento, estarão incluídos entre os 11 com as maiores populações de idosos do mundo. Haverá, portanto, uma substituição; as grandes populações idosas dos países europeus cederão lugar a países caracteristicamente jovens como Nigéria, Brasil, Paquistão, entre outros^{24,26}.

Este acontecimento não tem precedentes. Por

exemplo, o Brasil ocupava no ano de 1950 o 16º lugar no ranking mundial, com 2,2 milhões de idosos. Em 1980, essa população aumentou para 7,4 milhões, colocando o Brasil no 11º lugar. No ano de 2025, o país deverá estar ocupando o sexto lugar no ranking, com uma população acima de sessenta anos estimada em 33,8 milhões. Nesse período, compreendido entre o ano de 1950 a 2025, a população brasileira terá aumentado cinco vezes, enquanto o número de idosos será quinze vezes maior¹¹.

Atualmente no Brasil pouco se conhece sobre este grupo etário. Sabe-se, no entanto, que os recursos disponíveis são escassos e a demanda é crescente. Nos países onde a proporção de idosos é alta há mais tempo, e existe tradição de pesquisa, os estudos indicam algumas áreas onde são necessárias medidas urgentes para minorar alguns dos problemas da população idosa²⁷. Entre outras, destacam-se as doenças mentais, que aliás já foram objeto de alguns estudos no Brasil^{2,20}.

O presente artigo propõe-se a padronizar os conceitos de confiabilidade, utilizando-se para tal as definições presentes em Bartko e Carpenter⁵, Grove e col.⁸ e Almeida F²¹, este último em nosso meio, e de apresentar alguns desenhos de confia-

* Financiado pela FINEP (Proc. 42.86.0911.00) e FAPERJ (Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - Proc. 170.226/88)

** Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro - Rua São Francisco Xavier, 524 - 7º andar - 20550 - Rio de Janeiro, RJ - Brasil.

*** Departamento de Epidemiologia da Escola Nacional de Saúde Pública - Rua Leopoldo Bulhões, 1480 - 21041 - Rio de Janeiro, RJ - Brasil.

**** Serviço de Psiquiatria do Hospital Universitário Pedro Ernesto da Universidade do Estado do Rio de Janeiro - Rua 28 de Setembro, 77 - 20551 - Rio de Janeiro, RJ - Brasil.

***** Como população idosa ou terceira idade, considera-se indivíduos de sessenta anos ou mais.

bilidade, suas aplicações e dificuldades.

METODOLOGIA

Objetivou-se descrever os procedimentos utilizados no estudo piloto da versão em português do Short-CARE¹⁰ (segmento de saúde mental do questionário multidimensional Boas*), como parte de um conjunto de "controle de qualidade", com vistas à etapa seguinte do estudo quando será procedida a avaliação de sua confiabilidade e validade. Face a essa opção de desenho foi realizado um pequeno número de entrevistas onde não se conduziu uma análise estatística dos dados, mas concentrou-se no estudo caso a caso (qualitativo), visando ajustar os instrumentos e os procedimentos a serem utilizados por ocasião do estudo de confiabilidade e validação. Em síntese, esta fase é a de "sintonia fina", prévia a um outro momento quando será realizado os estudos tradicionais de confiabilidade e de validação do instrumento.

Fonte de dados

O estudo piloto foi realizado no distrito de Copacabana em 18 idosos, selecionados a partir de um levantamento das pessoas com 60 anos ou mais, residentes nos setores censitários incluídos aleatoriamente no universo da pesquisa²⁵. A perda de 4 dessas pessoas deveu-se a dificuldades decorrentes do intervalo entre as entrevistas**.

Coleta de dados

Para o piloto do estudo de confiabilidade e validação foram feitas duas visitas à residência do entrevistado. Na primeira vez, o entrevistador aplicou integralmente o questionário BOAS. Por ocasião da segunda visita, o mesmo entrevistador foi acompanhado por um dos psiquiatras da equipe do projeto. Nesta ocasião o entrevistador aplicou novamente o segmento de saúde mental do BOAS, enquanto o psiquiatra, usando o mesmo instrumento, anotava de modo independente do entrevistador, as respostas fornecidas. O objetivo desta etapa foi de testar a estratégia para a obtenção de estimativas de confiabilidade do tipo re-teste e de confiabilidade inter-avaliador.

Em seguida, o psiquiatra avaliou o entrevistado quanto a presença de síndrome cerebral orgânica e depressão, de acordo com os critérios definidos pelo DSM-III (Diagnostic and Statistical Manual

of Mental Disorders)³. O objetivo deste procedimento foi o de estabelecer o *golden standard* a fim de testar a capacidade discriminatória do segmento de saúde mental do BOAS.

A escolha do DSM-III como critério para validação do BOAS decorre da sua longa utilização, não só nos Estados Unidos onde foi desenvolvido pela Associação Psiquiátrica Americana³, mas também em inúmeros países. Desde a sua publicação, cerca de 2.000 artigos fizeram menção a ele e muitos de seus aspectos básicos foram incluídos no capítulo de distúrbios mentais da CID-10.

Para alcançar uma boa homogeneidade na avaliação dos psiquiatras, adotou-se o uso do check-list do DSM-III, cuja tradução no Brasil foi feita por Miranda e col.¹⁷. O check-list é um roteiro de sinais e sintomas relativos aos diagnósticos de distúrbios mentais, acompanhado de critérios bem definidos para inclusão ou exclusão de indivíduos em diversas categorias diagnósticas. Alguns desses sinais e sintomas podem ser avaliados através de perguntas padronizadas extraídas do questionário DIS (Diagnostic Interview Schedule)²¹, um instrumento estruturado de diagnóstico psiquiátrico criado para ser utilizado por entrevistadores leigos.

Na ausência de perguntas referidas ao DIS, os sinais e sintomas foram previamente definidos pelos psiquiatras de forma clara, com base na literatura especializada. Com isto, acredita-se ter melhorado a homogeneidade entre os dois psiquiatras.

De modo esquemático, os métodos utilizados para avaliar validação e confiabilidade podem ser observados na Figura.

Análise das entrevistas

Conforme assinalado anteriormente, não é objetivo deste trabalho apresentar uma análise estatística dos dados em decorrência do pequeno universo estudado (n = 14), mas sim ajustar o instrumento e a metodologia a ser utilizado em um universo maior de entrevistas para então estimar as medidas de confiabilidade e validade do segmento de saúde mental do BOAS.

O pequeno número de avaliações nos deu a oportunidade de analisar, caso a caso, as discrepâncias

* O questionário multidimensional visa cobrir as áreas consideradas as mais importantes da vida do idoso, a saber: saúde física, saúde mental, condição social, condição econômica e atividades do dia-a-dia. Brazil Old Age Schedule - BOAS¹⁹ é um questionário multidimensional de avaliação funcional, desenvolvido para a população idosa de um centro urbano (Rio de Janeiro), baseado em outros instrumentos que obtiveram padrões aceitáveis de validade e confiabilidade, como o questionário CARE⁹, para o segmento de saúde mental, e o OARS⁶ e PAHO¹⁸ para os demais segmentos.

** Ver discussão relativo ao estudo de confiabilidade do teste re-teste (Anexo).

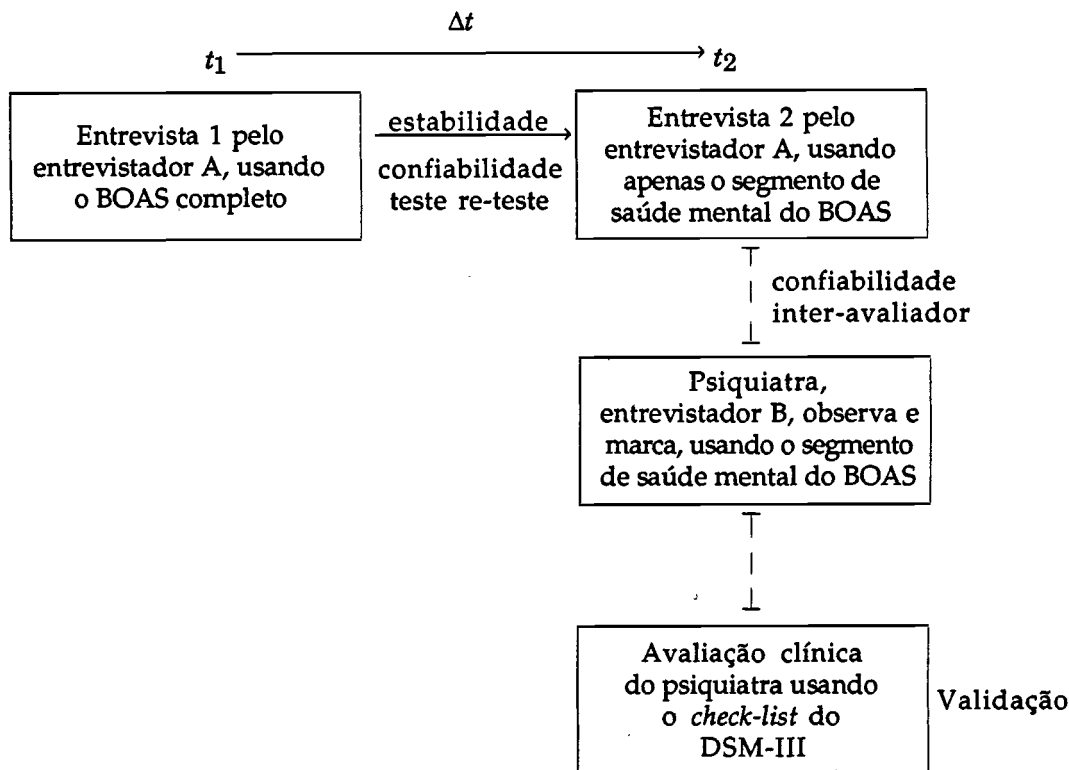


Figura — Delineamento para estudos de confiabilidade e validação: método utilizado no estudo piloto do segmento de saúde mental do questionário BOAS

ocorridas entre os resultados alcançados pelo BOAS nas diversas aplicações, assim como a discordância entre o BOAS e o diagnóstico psiquiátrico mediante o uso do *check-list* do DSM-III.

Pelo fato deste estudo ser apenas um piloto visando testar o delineamento proposto para o estudo de confiabilidade e validação do BOAS, não se estabeleceu pontos de corte do BOAS (segmento saúde mental) para o diagnóstico de suspeição de síndrome cerebral orgânica e de depressão mediante a utilização da análise ROC (Relative Operating Characteristic)^{7, 22, 23}.

Na ausência de um número de entrevistas suficientes para determinar os pontos de corte mais satisfatórios para o BOAS*, optou-se por aqueles sugeridos no estudo inglês coordenado por Mur-

phy¹². Foi utilizado (ainda que temporário) os pontos de corte propostos na investigação Guy's and Age Concern Psychogeriatric Research Project¹², onde entrevistados com escore superior a 2 na seção de síndrome cerebral orgânica e escore superior a 7 na seção de depressão foram considerados casos.

RESULTADOS E DISCUSSÃO

No estudo caso a caso para confiabilidade demonstrou haver apenas uma discordância, ocorrida no delineamento teste re-teste. Na revisão destas entrevistas atribuiu-se a possibilidade de ter ocorrido uma modificação real no estado de saúde mental do entrevistado. O longo tempo entre as duas entrevistas pode ser apontado como um fator que contribuiu para a discrepância no escore deste caso. Com relação a confiabilidade inter-avaliador,

* A forma comum de apresentar resultados num estudo de validade é mostrar um único valor para sensibilidade e especificidade. Segundo esse método, uma curva ROC é obtida plotando sensibilidade contra resultado falso-positivo para todos os pontos possíveis de corte. Um programa, denominado ROCFIT¹⁶ (1984) é usado para estimar parâmetros para a curva ROC ajustada. Este procedimento foi usado por Mari e Williams¹³ quando do estudo do SRQ-20 na versão brasileira¹⁴.

não houve discordância.

O critério de pontos de cortes fez com que, dos 14 entrevistados, houvesse quatro discordâncias entre o diagnóstico sugerido pelo BOAS e aquele definido pelos psiquiatras usando os critérios da DSM-III, com o auxílio do *check-list*.

Com relação ao diagnóstico de depressão, entre os 5 entrevistados cujo BOAS apresentou escore superior a 7 ("caso"), 3 não foram confirmados pelos critérios adotados pelos psiquiatras. O estudo detalhado destes "falsos-positivos" mostrou que dois preenchiam os critérios de desordens distímicas da DSM-III. O terceiro, ainda que não preenchesse os requisitos para o diagnóstico de distímia, foi considerado limítrofe. Portanto, é possível que o BOAS estivesse avaliando depressão num contexto mais amplo**, aliás mais adequado para os estudos comunitários do que os psiquiatras, restritos apenas à depressão maior. Ao se incluir o diagnóstico de distímia do *check-list* da DSM-III na entrevista psiquiátrica, há uma maior adequação dos critérios a serem adotados para o estudo de validação do segmento de saúde mental do BOAS.

Em relação ao diagnóstico de síndrome cerebral orgânica, houve apenas uma discordância entre o BOAS e o diagnóstico psiquiátrico. Enquanto o primeiro identificou o entrevistado como caso, o psiquiatra não confirmou o diagnóstico. O estudo desta entrevista apontou como possível causa da discordância o fato de a seção de síndrome cerebral orgânica do BOAS concentrar-se mais nos distúrbios de memória, enquanto o *check-list* do DSM-III valoriza também outros aspectos para o diagnóstico desta síndrome, tais como alteração do pensamento abstrato, do juízo, da personalidade e das funções corticais superiores. Portanto, apesar de os distúrbios de memória serem sensíveis à identificação de casos de síndrome cerebral orgânica, eles

não são específicos na identificação desta síndrome. Isto pode ser considerado como uma característica desejável, pelo fato de o instrumento ter sido elaborado para estudos populacionais, com a preocupação de identificar o maior número possível de casos com esta síndrome.

CONCLUSÃO

O conhecimento da confiabilidade do segmento de saúde mental do BOAS, através do cálculo dos coeficientes de concordância, conforme sugerido por Bartko e Carpenter⁵, assim como o estabelecimento dos pontos de corte para definição de casos e não-casos, só poderá ser considerado numa etapa posterior do estudo, quando um número maior de entrevistas tiver sido realizado. O mesmo podemos dizer da capacidade discriminatória do instrumento através da análise ROC.

Em que pese a ausência de um tratamento quantitativo dos dados, este estudo-piloto, através de uma abordagem caso a caso, permitiu identificar a existência de algumas deficiências e possibilitou a adoção de medidas para superá-las.

Confirmou-se a inadequação para os estudos de confiabilidade teste re-teste quando a segunda entrevista é realizada em um período muito distante da primeira. Verificou-se também que para o estudo da validação do BOAS (relativo a depressão), utilizando-se os critérios de *check-list* do DSM-III, mostrou-se a necessidade da entrevista psiquiátrica incluir os itens relativos aos critérios para desordens distímicas, ao lado do segmento relativo a depressão.

Estes achados sugerem a importância de uma abordagem qualitativa, numa etapa inicial do estudo, afim de se evitar a utilização de critérios e instrumentos inadequados quando da realização do estudo de confiabilidade e validação.

VERAS, R. P. et al. [The elderly population of Rio de Janeiro (Brazil): pilot-study as to the trustworthiness and validity of the mental health segment of the BOAS questionnaire]. Rev. Saúde públ., S. Paulo, 24:156-63, 1990.

ABSTRACT: A first step in the pilot-study undertaken as part of a research project on psychiatric morbidity — depression and organic brain syndrome — in the elderly population of Rio de Janeiro is described. The main objective is to "fine-tune" the instrument used (check list of DSM-III) and testing validity of the BOAS questionnaire (Brazilian version of Short-Care). The methodological aspects, and conceptual definitions are focused as well some procedures for the next stage of the project are suggested.

KEYWORDS: Aged. Mental health. Epidemiologic methods. Questionnaires. Validity.

* Em consequência de termos apenas 14 casos, tais resultados são extremamente instáveis, isto é, sujeitos a modificações importantes se movermos o ponto de corte estabelecido de um ponto para mais ou para menos. Por exemplo, a mudança do ponto de corte de 7/8 para 8/9, nos casos de depressão, faz com que a sensibilidade caísse de 100% para zero. Por outro lado, a inclusão de apenas um ou dois novos entrevistados pode provocar mudanças de grande magnitude em nossos valores de sensibilidade e especificidade.

** Esta categoria "depressão ampla" visa designar todas as formas leves de depressão, incluindo portanto, tristeza, processos reacionais etc.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALMEIDA FILHO, N. de *Epidemiologia sem números: uma introdução crítica a ciência epidemiológica*. Rio de Janeiro, Ed. Campus, 1989.
2. ALMEIDA FILHO, N. de; SANTANA, V.S.; PINHO, A.R. de Estudo epidemiológico dos transtornos mentais em uma população de idosos — área urbana de Salvador, BA. *J. bras. Psiq.*, 33: 114-20, 1984.
3. AMERICAN PSYCHIATRIC ASSOCIATION. *Committee on Nomenclature and Statistics: diagnostic and statistical manual of mental disorders*. 3rd. ed. Washington, D.C., 1980.
4. ANZOLA-PEREZ, E. Ageing in Latin American and the Caribbean. In: Pan-American Health Organization. *Toward the wellbeing of the elderly*. Washington, D.C., 1985. p. 9-23 (PAHO - Scientific Publication, 492).
5. BARTKO, J. J. & CARPENTER, W. T. On the methods and theory of reliability. *J. nervous Dis.*, 163:649-62, 1976.
6. DUKE UNIVERSITY OARS. *Multidimensional functional assessment: the OARS methodology*. 2nd ed. Durham, N C, Duke University, Centre for the Study of Ageing and Human Development, 1978.
7. ERDREICH, L. & LEE, E. T. Use of relative operating characteristic analysis in epidemiology. *Amer. J. Epidemiol.*, 114:649-62, 1981.
8. GROVE, W. M.; ANDREASEN, N. C.; DONALD-SCOTT, P. M.; KELLER, M.B.; SHAPIRO, R. W. Reability studies of psychiatric diagnosis. *Arch. gen. Psychiatry*, 38: 408-13, 1981.
9. GURLAND, B.; KURIANSKY, J.; SHARPE, L.; SIMON, R.; STILLER, P.; BIRKETT, P. The Comprehensive Assessment and Referral Evaluation (CARE): rationale, development and reability. *Int. J. Aging hum. Dev.*, 8:9-42, 1977/78.
10. GURLAND, B.; GOLDEN, R. R.; TERESI, J. A.; CHALLOP, J. The short-care: an efficient instrument for the assessment of depression, dementia and disability. *J. Geront.*, 39:166-9, 1984.
11. KALACHIE, A.; VERAS, R. P.; RAMOS, L. R. O envelhecimento da população mundial; um desafio novo. *Rev. Saúde públ.*, S. Paulo, 21:200-10, 1987.
12. LINDESAY, J. & MURPHY, E. Prevalence rates of cognitive impairment, depression and anxiety in a urban elderly community. *Int. J. geriat. Psychiatry*, 1989 [In press].
13. MARI, J. & WILLIAMS, P. A comparison of the validity of two psychiatric screening questionnaires(ghq12 and srq-20) in Brazil, usind Relative Operating Characteristic (ROC) analisys. *Psychol. Med.*, 15:651-9, 1985.
14. MARI, J. & WILLIAMS, P. A validity study of a psychiatry screening questionnaire (sqr-20) in primary care in the city of São Paulo. *Brit. J. Psychiatry*, 148:23-6, 1986.
15. MARTIN, L. G. The aging of Asia. *J. Geront. soc. Sci.*, 43(4):99-113, 1988.
16. MERTZ, C. E.; WANG, P. L.; KRONMAM, H.B. *Rocifit, Department of Radiology and the Franklin McLean Memorial Research Institute*. Chicago, University of Chicago, 1984.
17. MIRANDA, C. T.; IACOPONI, E.; GALENDER, G.; RAMOS, L. R. Diagnostic Inter Schedule (DIS) version iii. São Paulo, Departamento de Psiquiatria da Escola Paulista de Medicina, 1985. [Adaptação para o Brasil da versão portuguesa original feita pelo Departamento de Saúde Mental da Faculdade de Medicina do Porto].
18. PAN AMERICAN HEALTH ORGANIZATION. Epidemiology of ageing in Latin America and Caribbean. Washington, D.C. 1985. [Mimeografado].
19. QUESTIONNAIRE BOAS: Brazil Old Age Schedule(Multidimensional questionnaire for the elderly population). Rio de Janeiro, Instituto de Medicina Social da UERJ, 1987.
20. RAMOS, L. R. Health care for the elderly - an unmet need. Assessing health and living conditions of an elderly living in a large urban center, S. Paulo. London, 1986. [PhD Thesis - London School of Hygiene and Tropical Medicine. University of London].
21. ROBBINS, L. N.; HELZER, J. E.; CROUGHAN, J.; RA-TELIFF, R. S. National institute of mental health diagnostic interview schedule: its history, charactersitics and validity. *Arch. gen. Psychiatry*, 38:381-9, 1981.
22. SWETS, J.A. Roc analysis applied to the evaluation of medical imaging techniques. *Ass. Univ. Radiol.*, 14 (2):109-21, 1979.
23. SWETS, J.; PICKETT, R. M.; WHITEHEAD, S. F.; GETTY, D. J.; SCHNER, J. A., SWETS, J. B.; FREEMAN, B. A. Assessment of diagnostic technologies. *Amer. Ass. Adv. Sci.*, 205:753-9, 1979.
24. VERAS, R.P. Considerações acerca de um jovem país que envelhece. *Cad. Saúde públ.*, 4:382-97, 1988.
25. VERAS, R. P.; DUTRA, S.; SOUZA, C. A. M.; MILIOLI, R.; VENTURA, F. Proposta metodológica para inquérito domiciliar com populações idosas em um centro urbano do Estado do Rio de Janeiro (Brasil). *Rev. Saúde públ.*, S. Paulo, 23:429-38, 1989.
26. WARNES, T. The elderly in less-develop word regions. *Ageing and Soc.*, 6:373-80, 1986.
27. WORLD HEALTH ORGANIZATION. Scientific Group on Epidemciology in Aging, Geneva, 1983. *Uses of epidemiology in aging; Report*. Geneva, 1984. (Technical Report Series, 706)

Recebido para publicação em 11/4/1989
 Reapresentado em 20/10/1989
 Aprovado para publicação em 15/2/1990

ANEXO

ELABORAÇÃO DO QUESTIONÁRIO

Levando-se em conta o crescimento da população de mais 60 anos em nosso país, o projeto ora em curso do programa de Epidemiologia da Terceira Idade do Instituto de Medicina Social da UERJ pretende conhecer as características funcionais da população idosa de alguns distritos da cidade do Rio de Janeiro¹⁴, utilizando-se para tal de um questionário multidimensional. O objetivo desse estudo é investigar a veracidade da hipótese de que a saúde física e fatores sociais e econômicos estão associados à depressão e à síndrome cerebral orgânica na velhice¹⁴. Por isso, na elaboração deste questionário deu-se especial atenção a estes dois transtornos mentais, não só pela importante prevalência entre os idosos⁹, mas também por estarem estas doenças associadas ao desencadeamento e/ou aparecimento de outros distúrbios neste grupo etário¹⁰. O questionário BOAS, desenvolvido para o presente estudo¹⁵, foi elaborado levando em consideração a sugestão de Fillenbaum³. Esta pesquisadora propõe que antes de se partir para o desenho de um novo instrumento deve-se consultar aqueles já existentes que mostraram padrões aceitáveis de validade e confiabilidade. Para a elaboração do segmento de saúde mental, foram consultados o CARE (Comprehensive Assessment and Referral Evaluation⁵), o PSE (Present State Examination¹⁶), o MSQ (Measures for the determination of mental Status in the aged Questionnaire⁸), o SPMSQ (A Short Portable Mental State Questionnaire¹⁷) e o GMS (The Geriatric Mental State²). A opção recaiu sobre o CARE, por quatro razões principais: a) por ser o mais completo e recente instrumento; b) de ser estruturado e de ter sido desenvolvido para a população idosa na comunidade; c) de ter sido testado em estudos comparativos entre diferentes países⁶; d) de possuir uma versão reduzida e de também ter sido validada em diversos estudos⁷. A administração do CARE dura aproximadamente 90 minutos e requer profissional de saúde treinado. A versão reduzida da CARE, denominada de Short-CARE, cobre as áreas de depressão e síndrome cerebral orgânica. Tem ainda como característica poder ser aplicado por uma pessoa leiga, com um tempo de administração que não ultrapassa 20 min. Outro motivo para a escolha deste instrumento refere-se ao fato de que, no momento, realiza-se uma investigação domiciliar, em Londres, (Guy's and Age Concern Psychogeriatric Research Project), utilizando uma versão "inglesa" do Short-CARE*. Nesta investigação, são adotados os mesmos procedimentos me-

todológicos do nosso estudo no Rio de Janeiro. Pretendemos, num momento seguinte, comparar os resultados obtidos.

Nas línguas neolatinas existe uma versão para o italiano^{12,13}. A versão comentada no presente trabalho é a primeira tradução em português do Short-CARE, na qualidade de segmento de saúde mental do questionário BOAS.

Para a utilização do Short-CARE em nosso meio, fez-se uma tradução em três tempos. Inicialmente o questionário foi traduzido para o português. Em seguida, esta versão foi retraduzida para a língua inglesa por uma pessoa leiga que não o conhecia. Finalmente a versão original em inglês e a retraduzida foram revisadas por um especialista em epidemiologia e psiquiatria, com experiência direta no uso de questionário como instrumentos de pesquisa em população.

CONFIABILIDADE E VALIDADE

Para que um instrumento seja considerado adequado para um estudo epidemiológico, é necessário não apenas que exista concordância entre os resultados obtidos (confiabilidade), mas também que haja concordância entre esses resultados e o padrão definido como real (validade).

CONFIABILIDADE

Confiabilidade pode ser definido como a capacidade de um instrumento não variar seus resultados ao ser utilizado por diferentes pessoas ou em momentos diferentes do tempo¹.

Cronbach (citado por Almeida Filho)¹ utiliza duas idéias: estabilidade e equivalência. A primeira refere-se à obtenção dos mesmos resultados em diversas aplicações dos instrumentos. Ela expressa a capacidade que o instrumento tem de reproduzir os mesmos resultados quando a situação não se modifica. Já a equivalência tem a ver com o desempenho de diferentes instrumentos ao medir a mesma característica ou de um instrumento aplicado por diversas pessoas ao mesmo tempo ou, ainda, avaliado por diferentes juízes. Reflete, portanto, a capacidade de diferentes instrumentos ou diferentes avaliadores chegarem a uma mesma conclusão.

Grove e col.⁴ apontam 4 delineamentos de pesquisa de confiabilidade:

* Pequenas adaptações foram realizadas nesta versão do Short-CARE, pelo grupo de pesquisa da York Clinic, Guy's Hospital Medical School, (Universidade de Londres), e Age Concern, sob a coordenação da Prof. E. Murphy.

vinheta de caso — Aqui um caso é apresentado por escrito para múltiplos avaliadores. Tem como principal vantagem ser o método mais barato de delineamento, além da simplicidade e facilidade de sua execução. No entanto, ao apresentar os dados do caso por escrito ele elimina as fontes de discordâncias associadas com a técnica de entrevista, como a variação na informação fornecida pelo paciente em ocasiões distintas, interpretação não verbal. As únicas fontes de variações possíveis de influência sobre a concordância entre os avaliadores são a variação de interpretação dos critérios diagnósticos e na interpretação das vinhetas.

estudos de vídeo-tape — Apresenta-se um caso através de vídeo-tape ou filme para múltiplos avaliadores. É útil no estudo do viés de observação por permitir que muitos avaliadores observem um mesmo paciente simultaneamente. Seu uso em estudos multicêntricos também é interessante pois é mais econômico enviar o tape com a entrevista para os diferentes centros do que deslocar os avaliadores de diversos centros a fim de conduzi-rem as entrevistas. O maior problema deste delineamento é que o entrevistador, no vídeo-tape, pode fornecer pistas sobre a avaliação que ele faz do entrevistado, influenciando na opinião dos entrevistadores.

concordância interavaliadores — Neste caso, um avaliador aplica o questionário enquanto o outro observa. Depois, ambos avaliam independentemente. Trata-se do deliameento mais eficiente que os anteriores quanto a relação custo-benefício. Ao contrário dos estudos de vídeo-tape, não requer equipamentos e é capaz de mensurar as mesmas fontes de variação. Tem como desvantagens não permitir a presença de muitos observadores por entrevistado, além da possibilidade de que estes observadores sejam sugestionados pelo entrevistador, isto é, por aquele que conduz a entrevista.

teste reteste — Trata-se do método mais rigoroso, isto é, aquele que costuma apresentar os maiores índices de discordâncias. Nele os entrevistados são reavaliados algum tempo após a primeira entrevista, podendo haver 3 (três) fontes de variação: do entrevistado (mudança da história relatada de uma ocasião para a outra), do entrevistador (mudança na entrevista, na percepção) e do quadro nosológico (mudança real da sintomatologia). As variações devidas ao entrevistador podem ser minimizadas através de treinamento adequado da equipe, além do uso de questionários padronizados¹⁵. Outras fontes de variação podem

ser reduzidas se os intervalos entre as 2 (duas) entrevistas não for longo (recomenda-se até 20 dias). Intervalos longos também aumentam as possibilidades de recusas para a segunda entrevista, mas esta dificuldade pode ser reduzida se, no primeiro contacto, for estabelecido uma boa relação com o entrevistado.

Almeida Filho¹ considera a confiabilidade teste re-teste como sinônimo de estabilidade. Chama de confiabilidade de avaliação a concordância inter-avaliadores, e confiabilidade de aplicação a equivalência entre os resultados alcançados pela mesma aplicação do instrumento por diferentes entrevistadores.

VALIDADE

Em epidemiologia este termo expressa a capacidade de um instrumento medir corretamente uma dada característica. Para tal, exige-se um padrão contra o qual a avaliação obtida pelo instrumento é confrontada. A validade expressa-se numericamente através de duas medidas:

sensibilidade — É a capacidade de identificar como positivos todos os que possuem a característica em estudo. Portanto, quanto maior a sensibilidade, menor será a proporção de falsos-negativos (pessoas com a característica, classificadas pelo instrumento como não a possuindo).

especificidade — É a capacidade de identificar como negativos todos os que não possuem a característica em estudo. Portanto, quanto maior a especificidade, menor será a proporção de falsos-positivos (pessoas sem a característica, classificadas pelo instrumento como portadores desta).

Para que um instrumento seja considerado válido, é preciso haver confiabilidade, isto é, concordância entre os avaliadores quanto ao resultado obtido (homogeneidade de conceitos, critérios de avaliação). No entanto, uma confiabilidade elevada por si só, não garante a existência de validade, pois 2 (dois) avaliadores podem concordar entre si, ainda que estejam errados em suas avaliações. A confiabilidade (concordância entre os avaliadores) não é uma finalidade em si própria mas sim uma condição indispensável para se atingir um fim: a validade.

Em síntese, pode-se ter baixa validade com elevada confiabilidade mas a recíproca não é verdadeira.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALMEIDA FILHO, N. *Epidemiologia sem números: uma introdução crítica a ciência epidemiológica*. Rio de Janeiro, Ed. Campus, 1989.
2. COPERLAND, J. R. M.; KELLEHER, M. J.; GOURLAY, A. J.; GURLAND, B. J.; FLEISS, J. L.; SHARPE, L. A semistructured clinical interview for the assessment of diagnosis and mental health status in the elderly; the geriatric mental state schedule. *Psychol. Med.*, 6: 439-49, 1976.
3. FILLENBAUM, G. *The wellbeing of the elderly: approaches to multidimensional assessment*. Geneva, World Health Organization, 1984. (Technical Report, 84).
4. GROVE, W. M.; ANDREASEN, N. C.; DONALD-SCOTT, P. M.; KELLER, M. B.; SHAPIRO, R. W. Reability studies of psychiatric diagnosis. *Arch. gen. Psychiatry*, 38: 408-13, 1981.
5. GURLAND, B.; KURIANSKY, J.; SHARPE, L.; SIMON, R.; STILLER, P.; BIRKETT, P. The Comprehensive Assessment and Referral Evaluation (CARE): rationale, development and reliability. *Int. J. Aging hum. Dev.*, 8: 9-42, 1977/78.
6. GURLAND, B.; COPERLAND, J.; KELLEHER, M.; KURIANSKY, J.; SHAPE, L.; DEAN, L. *The mind and mood of aging: mental health problems of the community elderly in New York and London*. New York, Haworth Press; 1982.
7. GURLAND, B.; GOLDEN, R. R.; TERESI, J. A.; CHALLOP, J. The short-care: an efficient instrument for the assessment of depression, dementia and disability. *J. Geront.*, 39: 166-9, 1984.
8. KAHN, R. L.; GOLDFARB, A. I.; POLLACK, M.; PECK, A. Brief objective measures for determination of mental status in the aged. *Amer. J. Psychiatry*, 107: 326-8, 1960.
9. KAY, D. W. & BURROWS, G. D. *The epidemiology of mental disorders in the aged: handbook of studies on psychiatry and old age*. Amsterdam, Elsevier Science Publishers, 1984.
10. MURPHY, E. Social origins of depression in old age. *J. Psychiatry*, 141: 135-42, 1982.
11. PFEIFFER, E. A short portable mental status questionnaire for the assessment of organic brain deficit in the elderly patients. *J. Amer. geriat. Soc.*, 23: 433-41, 1975.
12. SPAGNOLI, A.; FORESTI, G.; MAC DONALD, A.; WILLIAMS, P. Italian version of the organic brain syndrome and the depression scales from care: evaluation of their performance in geriatric institutions. London, Institute of Psychiatry. University of London. Maudsley Hospital, 1986. [Mimeografado].
13. SPAGNOLI, A.; FORESTI, G.; MAC DONALD, A.; WILLIAMS, P. Dementia and depression in italian geriatric institution. *Int. J. Geriat. Psychiatry*, 1(1): 15-23, 1988.
14. VERAS, R. P. A multidimensional assessment of the elderly population living in Rio de Janeiro. London, Guy's Hospital Medical School. London University, 1986. [Protocolo de Projeto de Doutorado].
15. VERAS, R. P.; SOUZA, C. A. M.; CARDOSO, R. F.; MILLIOLI, R.; DUTRA, S. Pesquisando populações idosas: a importância do instrumento e o treinamento de equipe; uma contribuição metodológica. *Rev. Saúde públ.*, S. Paulo, 22: 513-8, 1988.
16. WING, J. K.; COOPER, J. E.; SARTORIUS, N. *The measurement and classification of psychiatric symptoms*. London, Cambridge University Press, 1974.