

Capacidade funcional de idosos: análise das questões de mobilidade, atividades básicas e instrumentais da vida diária via Teoria de Resposta ao Item

Functional capacity in the elderly: analyzing questions on mobility and basic and instrumental activities of daily living using Item Response Theory

La capacidad funcional de los ancianos: análisis de las preguntas sobre actividades de la vida diaria, utilizando la Teoría de Respuesta al Ítem

Cibele Comini César ^{1,2}
 Juliana Vaz de Melo Mambrini ²
 Fabiane Ribeiro Ferreira ³
 Maria Fernanda Lima-Costa ²

Resumo

This article aims to evaluate the items used to measure functional capacity in the 2010 Health Survey in Greater Metropolitan Belo Horizonte, Minas Gerais State, Brazil, using Item Response Theory. We analyzed the scale's dimensionality, the items' position and discriminatory power, and the precision of the functional capacity estimate. The study was based on a sample of 2,174 individuals aged 60 to 99 years, 61% of whom were women, with a median of 4 years of schooling. The instrument with 21 items (4 response options) showed a Cronbach's alpha coefficient of 0.98. Factor analysis identified one factor that explained 92% of the variability between the items. The results indicated: (i) redundancy in items; (ii) precise estimation of functional capacity only for elderly below the median on the scale; and (iii) inability of the elderly to differentiate between response options. The analysis suggests the need to discuss items included in the instrument in order to cover a wider range of the scale and seek more appropriate response options.

Health of the Elderly; Evaluation of Research Programs and Tools; Multivariate Analysis

Resumo

O objetivo deste estudo foi avaliar as questões sobre capacidade funcional utilizadas no Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, realizado em 2010, utilizando a Teoria de Resposta ao Item. Foram analisados: dimensionalidade da escala, posição dos itens na escala, poder de discriminação dos itens e precisão da estimativa da capacidade funcional. O estudo teve como base uma amostra de 2.174 idosos, com idades entre 60 e 99 anos, 61% mulheres e escolaridade mediana de 4 anos de estudos. O instrumento de 21 itens com 4 opções de respostas forneceu coeficiente alfa de Cronbach de 0,98; a análise fatorial mostrou existir um fator explicando 92% da variabilidade entre as questões. Os resultados indicaram: (i) redundância nas questões; (ii) precisão na estimação da capacidade funcional apenas para os idosos que estão abaixo da mediana na escala; e (iii) dificuldade dos idosos em discernir entre as opções de resposta. A análise sugere a necessidade de se discutir as questões a serem incluídas no instrumento, com vistas a uma maior cobertura da escala e melhor adequação das opções de respostas.

Saúde do Idoso; Avaliação de Programas e Instrumentos de Pesquisa; Análise Multivariada

¹ Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
² Núcleo de Estudos em Saúde Pública e Envelhecimento, Fundação Oswaldo Cruz/Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
³ Programa de Pós-graduação em Ciências da Reabilitação, Universidade Federal de Minas Gerais, Belo Horizonte.

Correspondência
 C. C. César
 Departamento de Estatística,
 Universidade Federal de Minas Gerais.
 Av. Antônio Carlos 6627,
 Belo Horizonte, MG 30123-970,
 Brasil.
 cibelescomini@yahoo.com.br

Introdução

Mudanças na estrutura da população mundial, observadas a partir do início do século passado, modificaram o perfil de morbidade da população, com as doenças crônicas e suas consequências passando a preponderar sobre os processos agudos. Esse novo perfil de adoecimento resultou em uma ampliação do conceito de saúde, que passou a considerar o bem-estar do indivíduo, englobando as dimensões física, mental e social e não apenas a ausência de doença¹. No que tange à população idosa, um aspecto importante do bem-estar e, portanto do estar saudável, é a capacidade funcional^{2,3}, que surge como um novo paradigma de saúde, sendo a dimensão base para a avaliação geriátrica^{4,5,6}.

Há diversas maneiras de se avaliar a capacidade funcional. Nos estudos epidemiológicos é frequente considerar a habilidade para realizar atividades básicas da vida diária (ABVD) e/ou atividades instrumentais da vida diária (AIVD) e/ou atividades relacionadas à mobilidade. A função é definida como a capacidade de um indivíduo se adaptar aos problemas cotidianos, ou seja, aquelas atividades que lhe são requeridas por seu entorno imediato, incluindo a sua participação como indivíduo na sociedade, ainda que apresente alguma limitação física, mental ou social⁷. A capacidade para realizar ABVD é uma medida importante para avaliar a demanda por assistência, cuidados e apoio^{8,9}. As ABVD incluem atividades tais como alimentar-se, vestir-se, tomar banho, usar o sanitário, transferir-se da cama para uma cadeira e caminhar em um cômodo do mesmo andar. As AIVD compreendem outras atividades indicativas da capacidade para levar uma vida independente na comunidade, incluindo realizar tarefas domésticas, fazer compras, administrar as próprias medicações e manusear dinheiro, entre outras¹⁰. A mobilidade compreende a capacidade para locomoção, agachar-se/ajoelhar-se, subir escadas, entre outras, que refletem disfunções criadas por condições crônicas, bem como o funcionamento cognitivo e afetivo. Os estudos utilizam as atividades de vida diária de diversas maneiras, podendo tomar como base os relatos de dificuldades para realizar diferentes ABVD e/ou AIVD^{3,11,12} ou, o que é mais frequente, considerar, de alguma forma, o número de atividades realizadas com dificuldade^{13,14,15,16}.

Em Belo Horizonte, capital do Estado de Minas Gerais, Brasil, foi examinada a associação entre a capacidade funcional de idosos (60 anos ou mais) e a percepção da vizinhança. A limitação funcional foi definida pela soma de 15 ABVD e AIVD realizadas com dificuldade. En-

tre os 1.611 participantes, 52,9% não apresentaram dificuldade para realizar nenhuma das atividades consideradas¹⁷. Clarke & Smith¹⁸ estudaram o efeito do contexto macrosocial e econômico no bem-estar e na saúde individual de 3.784 participantes do *Health and Retirement Study* (HRS), conduzido nos Estados Unidos, e de 3.860 participantes do *English Longitudinal Study on Ageing* (ELSA), realizado na Inglaterra, ambos inquéritos conduzidos em 2006 com indivíduos de 50 anos ou mais. Entre 13 atividades consideradas, 7 AIVD e 6 ABVD, 51,9% dos americanos e 66,7% dos ingleses não apresentaram nenhuma dificuldade para realizá-las. Dados da primeira onda do *Canadian Study of Health and Ageing* (CSHA), realizada no Canadá em 1991 com 9.008 idosos não institucionalizados, encontraram que 84,6% dos idosos relataram não ter dificuldade para realizar as ABVD e 67,6% para as AIVD. A capacidade funcional foi avaliada por meio de 14 itens referentes às atividades básicas e instrumentais, sendo 7 itens para cada grupo de atividades¹⁶.

Nos estudos relatados, apesar do número elevado de atividades da vida diária utilizadas na construção da variável resposta, dado o alto percentual de entrevistados que relatou não ter dificuldade em nenhuma atividade, não foi possível diferenciar o nível de capacidade funcional de parcela importante da amostra.

Uma abordagem alternativa para estimar a funcionalidade é a Teoria da Resposta ao Item (TRI). A TRI propõe modelos que transformam respostas dicotômicas ou ordinais de questões relacionadas a um mesmo construto em um escore contínuo. Em contraste com a abordagem tradicional, que se baseia na Teoria Clássica da Medida (TCM) e que tem como referência o instrumento, a TRI adota a perspectiva de que a resposta positiva de um indivíduo a uma questão está associada a níveis mais altos desse escore contínuo que representa uma variável ou traço não observável do indivíduo: a variável latente¹⁹. Os modelos da TRI, por serem expressos com base na resposta a cada questão e não do escore total no instrumento, possibilitam o desenvolvimento de instrumentos mais curtos ou até mesmo de instrumentos adaptados ao perfil do respondente, com ótimas propriedades psicométricas²⁰. Ao analisar a resposta dada à questão, é possível determinar qual nível da variável latente está mais associada a ela. Por meio dessa análise, a TRI permite que se identifique questões que representam diferentes posições da escala da variável de interesse, auxiliando o pesquisador na construção do instrumento. No processo de construção de uma escala, espera-se que existam questões que cubram todo o espectro da variá-

vel latente. Por outro lado, questões associadas à mesma posição na escala da variável latente não acrescentarão muita informação à medida. Uma das principais aplicações da TRI é a geração de escore. Por considerar as características do item na criação do escore, a TRI fornece escores mais confiáveis do que, por exemplo, a contagem do número de respostas corretas. Uma característica importante do escore gerado pela TRI é que estimativas comparáveis da variável latente podem ser obtidas mesmo de testes construídos com itens não equivalentes²⁰. Na utilização dos modelos da TRI pode-se supor a existência de mais de uma variável latente, o que determina a dimensionalidade do modelo. Um estudo anterior utilizando ABVD e AIVD concluiu pela existência de uma variável latente²¹.

Este artigo tem por objetivo avaliar 21 questões relativas à capacidade funcional, em uma amostra representativa de idosos residentes na Região Metropolitana de Belo Horizonte, utilizando a TRI. A avaliação das questões visa a elucidar a dimensionalidade do instrumento utilizado, a posição destas questões ao longo da escala de funcionalidade, a capacidade de cada questão em diferenciar idosos de diferentes níveis de funcionalidade e avaliar a precisão da escala na estimação da funcionalidade.

Metodologia

Amostra

Os dados analisados foram provenientes do 2º *Inquérito de Saúde da Região Metropolitana de Belo Horizonte*. Trata-se de um estudo de base populacional que foi conduzido entre maio e junho de 2010, em uma amostra representativa dos adultos residentes nos municípios que compõem a Região Metropolitana de Belo Horizonte²². A coleta de dados para a pesquisa foi realizada mediante questionário suplementar à *Pesquisa de Emprego e Desemprego* (PED), periodicamente conduzida na região. A PED é realizada por amostragem em domicílios na área urbana de 26 municípios. A coleta de dados é feita por meio de delineamento amostral probabilístico por conglomerado, estratificada em dois estágios, sendo o setor censitário a unidade primária de amostragem e o domicílio a unidade amostral. A amostra continha 7.500 domicílios e 24 mil moradores. Para este trabalho, foram selecionados todos os participantes com 60 anos ou mais de idade, totalizando 2.271 indivíduos.

Considerações éticas

O *Inquérito de Saúde da Região Metropolitana de Belo Horizonte* foi aprovado pelo Comitê de Ética em Pesquisas do Centro de Pesquisas René Rachou, unidade da Fundação Oswaldo Cruz localizada em Belo Horizonte (protocolo nº 10/2009).

Itens

Foram selecionados todos os 21 itens do inquérito de saúde que avaliaram a capacidade funcional do participante, compreendendo indicadores de mobilidade, AIVD e ABVD. Os primeiros incluíram: caminhar 2 ou 3 quarteirões, subir dez degraus de escada, inclinar-se/agachar-se/ajoelhar-se, levantar/carregar objetos pesando em torno de 5kg. As AIVD consideradas foram: tomar remédios, administrar o próprio dinheiro, ir às compras, cortar as unhas dos pés, sair de ônibus, realizar tarefas domésticas leves, realizar tarefas domésticas pesadas e preparar a própria refeição. As ABVD incluídas na análise foram: alimentar-se, vestir-se, tomar banho, pentear o cabelo/lavar o rosto, usar o sanitário, caminhar em um cômodo no mesmo andar, levantar de uma cadeira sem braços, levantar da cama, transferir-se da cama para a cadeira. Todos os itens foram mensurados em escala ordinal de 4 categorias, correspondendo a nenhuma dificuldade, alguma dificuldade, muita dificuldade e somente com a ajuda de outra pessoa.

O modelo

A TRI é uma família de funções matemáticas que modelam a probabilidade da resposta à questão, considerando as características da questão e características não observáveis do indivíduo chamadas de variáveis, traços ou fatores latentes. Variáveis latentes são construções teóricas que explicam comportamentos consistentes em contextos variáveis¹⁹. Podem representar, por exemplo, proficiência acadêmica, atitude ou crença²⁰.

A função que relaciona a probabilidade de escolha de uma determinada alternativa com a variável latente é conhecida como curva característica do item (CCI). Descreve como a probabilidade de resposta à questão está relacionada com a variável latente. Existem modelos da TRI para respostas dicotômicas e politômicas (nominais ou ordinais). Para os modelos da TRI é feita a suposição de independência condicional entre as questões, ou seja, embora exista correlação entre as respostas do mesmo indivíduo, esta é inteiramente explicada pelas variáveis latentes. O número de variáveis latentes subjacentes ao conjunto de questões fornece a dimensionalidade

de do instrumento. Embora existam alternativas para se modelar a CCI, a opção mais frequente é a função logística. Um dos objetivos da TRI é encontrar um pequeno conjunto de questões que mapeie toda a extensão da variável latente.

O modelo geral da TRI para uma resposta binária considerando uma função logística é representado por:

$$P_i(Y=1|\theta) = c + (1-c) \frac{\exp[a_i(\theta - b_i)]}{1 + \exp[a_i(\theta - b_i)]} \quad (1)$$

em que $P_i(Y=1|\theta)$ representa a probabilidade de um indivíduo cujo valor da variável latente é θ dar resposta "1" à *i*-ésima questão. No contexto deste artigo, se a resposta a cada uma das atividades fosse 1 se a pessoa declarasse não ter dificuldade e 0 se tivesse dificuldade ou não conseguisse executar, $P(Y=1|\theta)$ seria a probabilidade de um idoso com nível de capacidade funcional igual a θ declarar que não tem dificuldade para a atividade sendo avaliada.

Essa probabilidade é modelada como função da variável latente θ e dos parâmetros da questão: a , b e c . O parâmetro b está associado com a posição da atividade na escala da variável latente sendo medida. Indica para qual valor da variável latente a atividade é realizada sem dificuldade por 50% dos indivíduos. Esse parâmetro está associado com o grau de complexidade da atividade: quanto mais alto o valor de b , maior a capacidade funcional necessária para que 50% dos indivíduos consigam realizar a atividade. O parâmetro a mede a inclinação da CCI no ponto b . Indica quão bem a atividade é capaz de classificar o idoso como tendo ou não dificuldade em realizar a tarefa. Em outras palavras, se a é alto, há uma diferenciação muito clara entre o grupo que é capaz de realizar a atividade do grupo que não consegue realizá-la. Nos modelos para itens binários, o parâmetro a indica a habilidade do item em discriminar os indivíduos com relação à variável latente no ponto b . Mostra como o item está relacionado com a variável latente. O parâmetro c é a assíntota inferior da curva, importante em situações em que a probabilidade de dar uma resposta afirmativa é superior a 0 nos níveis mais baixos da variável latente. No ambiente educacional esse parâmetro se propõe a captar a probabilidade de acerto ao caso, que pode ocorrer quando o examinando não sabe a questão. No ambiente de saúde, o parâmetro tem mais de uma interpretação: pode captar a resposta de indivíduos que querem se passar por saudáveis ou erro de resposta, sendo esta a interpretação mais aceita^{21,23,24}.

O modelo apresentado, conhecido como modelo de 3 parâmetros, pode ser simplificado. Se o parâmetro c for eliminado, temos o modelo de 2

parâmetros. Se todas as questões apresentarem a mesma discriminação temos o modelo de 1 parâmetro, equivalente ao modelo de Rasch²⁵.

A métrica dos parâmetros da TRI é arbitrária, sendo definida com base na identificação do centro da escala e da unidade de medida. Em geral, são estabelecidos como centro e unidade de medida a média e o desvio padrão estimados de θ , que recebem os valores 0 e 1, respectivamente. O parâmetro b tem a mesma métrica de θ , variando teoricamente no intervalo $(-\infty, \infty)$ mas, na prática, assumindo valores no intervalo $(-3, 3)$. O intervalo de variação teórico do parâmetro de discriminação (a) é também $(-\infty, \infty)$. Entretanto, não se espera observar valores positivos muito elevados ou valores negativos. Itens com parâmetros de discriminação muito baixos, menores que 0,4, em geral, são eliminados do instrumento durante o processo de desenvolvimento²⁰. O parâmetro c embora possa variar no intervalo $(0, 1)$ é, em geral, inferior a 0,3, sendo quase sempre nulo para questões na área de saúde.

Os modelos para respostas politômicas são generalizações do modelo binário. O Modelo de Respostas Graduadas (MRG)²⁶ é uma generalização do modelo de 2 parâmetros. É adequado às situações nas quais as questões foram mensuradas em escala ordinal, como nas escalas do tipo Likert. No MRG modela-se a probabilidade do respondente no nível θ da variável latente escolher a resposta k ou superior:

$$P_{ik}(Y \geq k|\theta) = \frac{\exp[a_i(\theta - b_{ik})]}{1 + \exp[a_i(\theta - b_{ik})]}, \quad k = 0, 1, \dots, t \quad (2)$$

Nesse modelo, b_{ik} representa o limite da categoria k da questão i e corresponde ao valor de θ para o qual 50% dos entrevistados escolhem a categoria k ou superior. Espera-se que a probabilidade de resposta a uma categoria superior esteja associada a níveis mais altos da variável latente, ou seja, os b_{ik} são ordenados. No MRG a discriminação da questão é função do parâmetro de inclinação a e da localização relativa dos parâmetros b_k ²⁰.

Na TRI, o erro padrão do escore é obtido com base na função de informação do instrumento, que varia segundo o valor do escore e das características do conjunto de questões que fazem parte do instrumento. Quanto maior a informação, menor o erro padrão do escore. A informação do instrumento é obtida pela soma da informação das questões que o compõem. Para questões binárias, a informação atinge o máximo em $\theta = b$, concentrando-se em torno deste ponto se o valor do parâmetro a é alto. Se a é baixo, a questão fornece pouca informação em um espectro mais amplo do escore, o que é retratado por uma função de informação achatada. Em questões

politômicas, todas as categorias contribuem para a informação. Se as categorias estiverem bem dispersas, cada uma delas fornecerá informação em um ponto diferente da escala dos escores, o que se refletirá em diversos picos na função de informação. Nesse caso, a questão fornecerá informação sobre θ em um espectro amplo da escala. Se as categorias estiverem próximas, haverá um ponto de informação máxima, que ocorrerá próximo ao centro dos parâmetros b e a questão funcionará como se fosse dicotômica²⁰.

A abordagem da TRI, ao centrar a avaliação na questão e não no instrumento, tem diversas características desejáveis. Do ponto de vista teórico, os modelos da TRI fornecem escores medidos na escala intervalar, e as estimativas dos parâmetros das questões e da variável latente são invariantes às características da amostra e do instrumento. Do ponto de vista prático, a confiabilidade do instrumento não está relacionada ao seu tamanho; as estimativas da variável latente estão ancoradas nas questões, permitindo a comparação entre populações que foram submetidas a instrumentos distintos, desde que possuam algumas questões comuns ou ainda entre indivíduos da mesma população que tenham sido submetidos a instrumentos totalmente distintos; facilita avaliar se as questões têm o mesmo significado para subgrupos populacionais distintos e a inclusão de questões com formatos diferentes^{27,28}. Essas propriedades do escore da TRI dependem da verificação das seguintes condições: especificação correta da dimensionalidade do instrumento, independência local e especificação correta do modelo para a probabilidade da resposta. A dimensionalidade do instrumento se refere ao número de variáveis latentes medidas pelas questões. A independência local estabelece que a correlação existente entre as respostas é totalmente explicada pelos fatores latentes avaliados pelo instrumento.

Análise de dados

Os itens foram avaliados com relação à unidimensionalidade por meio do coeficiente alfa de Cronbach, que mede a consistência interna, e pela análise fatorial exploratória da matriz de correlação policórica entre os 21 itens de capacidade funcional.

A independência condicional foi avaliada por meio da análise da matriz de correlação entre os resíduos obtidos após o ajuste da análise fatorial.

Para a construção da escala foi utilizado o modelo para respostas graduadas. Foram consideradas duas opções de modelos: modelo de 1 parâmetro, ou seja, com inclinação constante para todos os itens; e modelo de 2 parâmetros,

que considera inclinação variável. O teste da razão das verossimilhanças foi utilizado para escolher o modelo que melhor se ajustava aos dados.

São apresentados os parâmetros e as curvas características dos itens, bem como a curva de informação para o conjunto dos 21 itens utilizados na construção da escala.

Todas as análises foram realizadas no software R (The R Foundation for Statistical Computing, Viena, Áustria; <http://www.r-project.org>). Para o ajuste dos modelos foi utilizado o pacote *ltm*²⁹.

Resultados

Características da amostra

O inquérito de saúde teve como base uma amostra de 2.271 participantes com 60 anos ou mais. Desses, 97 (aproximadamente 5%) não responderam aos itens referentes à avaliação de saúde e não foram incluídos na análise.

Entre os 2.174 participantes da pesquisa, a idade variou de 60 a 99 anos, (mediana = 68, IQ = [63;75]), sendo 60,9% mulheres. A mediana da escolaridade foi igual a 4 anos (IQ = [3;11]), com 26,7% dos entrevistados declarando escolaridade inferior a 4 anos e 36,9% entre 4 e 7 anos de estudos.

Descrição dos itens

O percentual de não resposta nas questões relativas à capacidade funcional foi inferior a 0,6%. Houve concentração das respostas na categoria “nenhuma dificuldade”, com o percentual nesta categoria variando entre 62% (“inclinarse/ agacharse/ ajoelhar-se”) e 92% (“alimentarse” e “escovar os dentes/pentear os cabelos/lavar o rosto”). Essa porcentagem diminuiu com o aumento da idade (resultados não apresentados). Para a pior condição, ou seja, “só com a ajuda de outra pessoa”, a porcentagem variou entre 3 e 12%, com o maior percentual sendo observado para alimentar-se, como pode ser visto na Tabela 1.

Pressupostos do modelo da TRI

As correlações entre as 21 atividades foram positivas e estatisticamente significativas, variando de 0,73 a 0,98. Entre as 210 correlações, 54 (25,7%) apresentaram valor de pelo menos 0,90; 68,6% ficaram entre 0,80 e 0,90.

O coeficiente alfa de Cronbach foi 0,98 considerando os 21 itens da escala e não mostrou alteração com a exclusão dos itens individualmente.

Na decomposição da matriz de correlação policórica segundo seus autovalores e autovetores,

Tabela 1

Distribuição (%) do grau de dificuldade para a realização de atividades de mobilidade, atividades instrumentais e básicas da vida diária. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte*, Minas Gerais, Brasil, 2010.

Atividade	Nenhuma dificuldade	Alguma dificuldade	Muita dificuldade	Só faz com ajuda
Comer	91,6	3,8	1,7	2,9
Escovar os dentes, pentear os cabelos ou lavar o rosto	91,6	3,6	1,6	3,2
Usar o sanitário	91,1	3,9	1,5	3,5
Tomar banho	89,8	4,8	1,8	3,6
Tomar remédios	89,6	4,3	1,6	4,5
Vestir-se	87,2	6,8	2,6	3,4
Caminhar de um cômodo a outro no mesmo andar	86,9	6,1	3,3	3,7
Transferir-se da cama para uma cadeira	85,3	7,9	3,4	3,4
Administrar o próprio dinheiro	84,8	6,3	2,5	6,4
Levantar da cama	84,2	8,8	3,6	3,4
Preparar o próprio alimento	82,8	7,0	3,7	6,5
Levantar-se de uma cadeira sem braço	82,7	9,5	3,9	3,9
Sair e fazer compras	76,8	8,7	4,0	10,5
Fazer trabalhos domésticos leves	73,9	12,3	6,6	7,2
Caminhar 2 ou 3 quarteirões	73,5	14,2	6,5	5,8
Sair de ônibus	73,4	12,1	4,4	10,1
Cortar as unhas dos pés	70,0	14,0	4,5	11,5
Subir 10 degraus de escada	65,4	18,5	9,4	6,7
Levantar ou carregar objetos leves	64,4	16,0	10,0	9,6
Fazer trabalhos domésticos pesados	62,6	15,2	10,2	12,0
Inclinar-se, agachar-se ou ajoelhar-se	61,9	20,6	10,9	6,6

observaram-se os valores 18,46 e 0,71 para os dois primeiros autovalores, correspondendo a uma razão de 26. Esses valores indicaram para a conveniência do modelo de análise fatorial com 1 fator. O ajuste do modelo de 1 fator forneceu autovalor, explicando 92% da variância e cargas fatoriais variando entre 0,86 e 0,97.

A análise da matriz de correlação residual obtida após o ajuste do modelo de análise fatorial não evidenciou inadequação do modelo com um fator. Foram observados valores baixos para a correlação residual entre itens, com 79,9% das correlações com valores absolutos inferiores a 0,03 (43,3% menores que 0,02). Somente um coeficiente apresentou valor absoluto superior a 0,10: correlação residual de 0,11 entre “levantar ou carregar objetos leves” e “fazer trabalhos pesados”.

Estimativa dos parâmetros e ajuste do modelo

O que melhor se ajustou aos itens foi o modelo de 2 parâmetros – que considera discriminação variável. Esse se mostrou superior, na avaliação

pelo teste da razão das verossimilhanças, ao modelo de 1 parâmetro no qual a discriminação é constante para todas as atividades (valor de $p < 0,001$).

A Tabela 2 apresenta as estimativas dos parâmetros das questões considerando o modelo de 2 parâmetros. O parâmetro de inclinação variou entre 2,55 e 5,68, sendo 7 (33,3%) deles iguais ou superiores a 4. Os limites das categorias foram todos negativos, variando de -2,56 a -0,36. O menor valor foi observado para a transição “somente com a ajuda de outra pessoa” e “muita dificuldade” para a atividade “comer”, e o maior valor para a transição entre as categorias “alguma dificuldade” e “nenhuma dificuldade” da atividade “inclinar-se, agachar-se ou ajoelhar-se”. Essas foram as atividades que apresentaram, respectivamente, as maiores e menores porcentagens de respondentes que declararam não ter dificuldade para realizá-las, ou seja, as atividades mais fácil e mais difícil, respectivamente, pela TCM.

Os valores estimados de θ variaram entre -3,07 e 0,63, com média de -0,16 (DP = 0,90), mediana de 0,14 (IQ = [-0,77;0,63]). Os resultados apontam para a assimetria da distribuição dos

Tabela 2

Estimativas da inclinação e limites das categorias para o modelo de dois parâmetros. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, 2010.*

Atividade	b_{11}	b_{12}	b_{13}	a_i
Comer	-2,56	-2,17	-1,72	3,79
Escovar os dentes, pentear os cabelos ou lavar o rosto	-2,52	-2,17	-1,75	3,68
Usar o sanitário	-2,41	-2,09	-1,65	4,51
Tomar banho	-2,37	-2,03	-1,55	4,64
Tomar remédios	-2,31	-2,06	-1,63	3,40
Vestir-se	-2,47	-1,96	-1,40	4,16
Caminhar de um cômodo a outro no mesmo andar	-2,33	-1,81	-1,36	5,12
Transferir-se da cama para uma cadeira	-2,40	-1,80	-1,27	5,68
Administrar o próprio dinheiro	-2,02	-1,75	-1,31	3,33
Levantar da cama	-2,41	-1,80	-1,21	5,42
Preparar o próprio alimento	-2,01	-1,65	-1,20	3,54
Levantar-se de uma cadeira sem braço	-2,31	-1,74	-1,14	4,92
Sair e fazer compras	-1,64	-1,38	-0,93	3,40
Fazer trabalhos domésticos leves	-1,91	-1,40	-0,79	3,61
Caminhar 2 ou 3 quarteirões	-2,04	-1,46	-0,77	3,73
Sair de ônibus	-1,65	-1,37	-0,78	3,47
Cortar as unhas dos pés	-1,60	-1,33	-0,67	3,03
Subir 10 degraus de escada	-1,96	-1,26	-0,49	3,42
Levantar ou carregar objetos leves	-1,74	-1,16	-0,48	2,93
Fazer trabalhos domésticos pesados	-1,62	-1,07	-0,42	2,55
Inclinar-se, agachar-se ou ajoelhar-se	-1,99	-1,20	-0,36	3,31

escores, com maior concentração na porção direita da curva. Uma grande parcela dos entrevistados (48,9%) obteve o escore máximo, referente à realização de todas as atividades de vida diária sem dificuldade. A distribuição dos escores é apresentada na Figura 1.

As CCI são apresentadas nas Figuras 2, 3 e 4. A análise das curvas evidencia que, para quase todos os itens, os respondentes com nível de funcionalidade a partir de zero respondem “não tem dificuldade” com probabilidade próxima de 1. Os únicos itens para os quais a probabilidade dessa categoria está posicionada um pouco mais à direita, na vizinhança do nível 0,5, são “cortar as unhas dos pés”, “subir dez degraus de escada”, “levantar ou carregar alguma coisa pesando até 5kg” e “inclinar-se, ajoelhar-se, agachar-se”. Quanto às categorias de resposta, em 8 dos 21 itens (38%) as 4 alternativas de respostas foram de fato utilizadas de maneira distinta pelos respondentes, sendo 3 itens de mobilidade e 5 de atividade instrumental (“Caminhar dois ou três quarteirões”, “Subir dez degraus de escada”, “Inclinar-se, ajoelhar-se, agachar-se”, “Caminhar de um cômodo a outro em um mesmo andar”, “Levantar de uma cadeira sem braço”, “Levantar

da cama”, “Transferir-se da cama para uma cadeira”). Em 5 atividades avaliadas (24%) foram utilizadas efetivamente 3 alternativas de resposta, sendo a categoria “alguma dificuldade” englobada totalmente pela alternativa “não consegue” ou partilhada pelas duas alternativas adjacentes (“Levantar ou carregar alguma coisa pesando até 5kg”, “Fazer trabalhos domésticos leves”, “Fazer trabalhos domésticos pesados”, “Sair de ônibus”, “Tomar banho”). Nas 8 atividades restantes (38%), somente as duas categorias extremas foram efetivamente utilizadas. Nesse último grupo, estão três atividades básicas e cinco atividades instrumentais (“Escovar os dentes, pentear os cabelos ou lavar o rosto”, “Ir ao banheiro”, “Comer”, “Tomar remédios”, “Sair e fazer compras”, “Cortar as unhas dos pés”, “Preparar o próprio alimento”, “Administrar o dinheiro”).

A Figura 5 apresenta a Curva de Informação do Teste, que é obtida com base na soma da informação de cada um dos itens. Observa-se que o teste fornece muita informação para os valores negativos da variável latente, mas praticamente nenhuma para os valores positivos. Lembrando que a precisão da estimativa do escore é dada pelo inverso da informação, concluímos

Figura 1

Escores de capacidade funcional estimados pela Teoria da Resposta ao Item. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, 2010.*

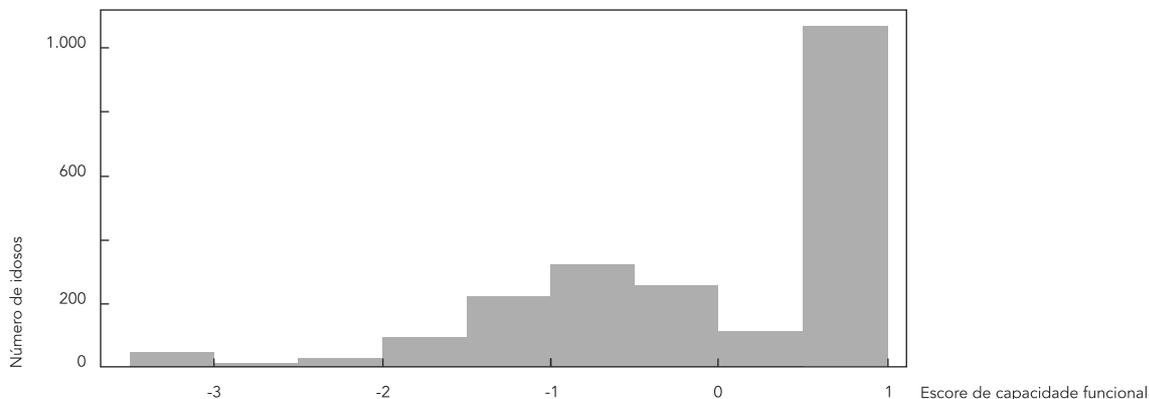
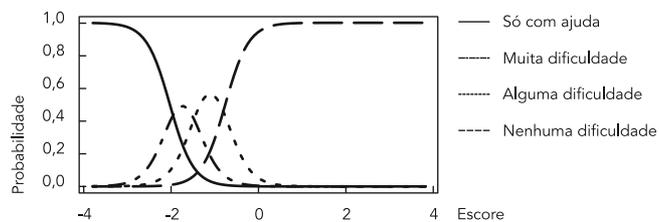


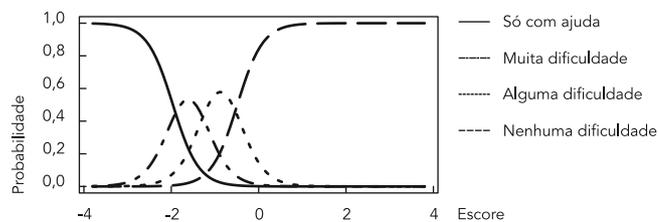
Figura 2

Curva característica dos itens referentes às atividades de mobilidade. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, 2010.*

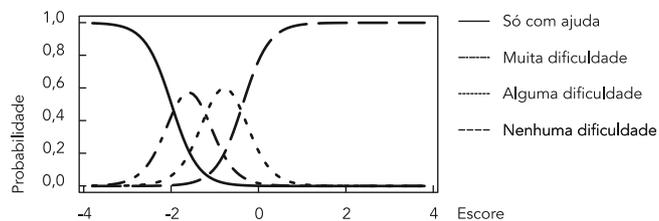
2a) Caminhar 2 ou 3 quarteirões



2b) Subir 10 degraus de escada



2c) Inclinar-se, agachar-se ou ajoelhar-se



2d) Levantar ou carregar objetos leves

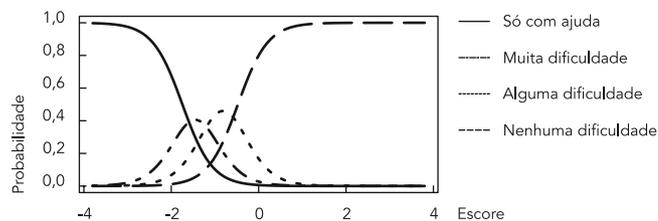
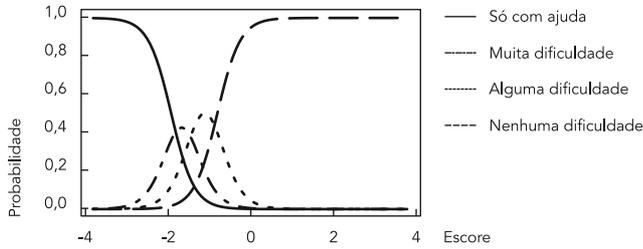


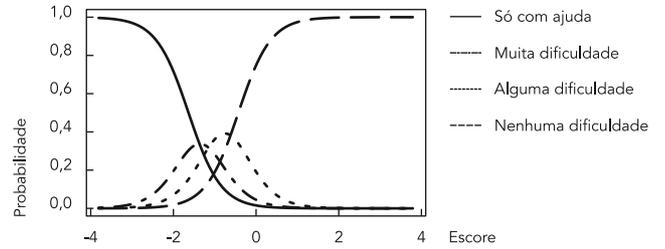
Figura 3

Curva característica dos itens referentes às atividades instrumentais da vida diária. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, 2010.*

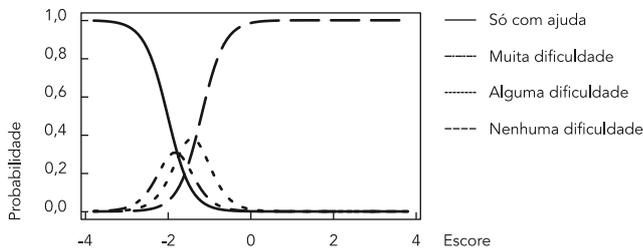
3a) Fazer trabalhos domésticos leves



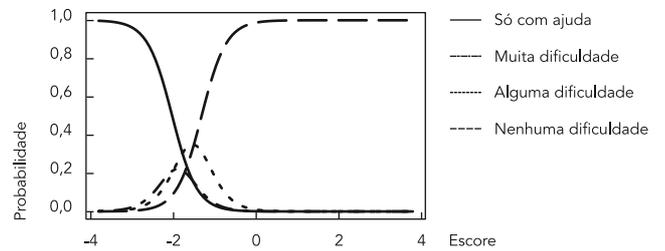
3b) Fazer trabalhos domésticos pesados



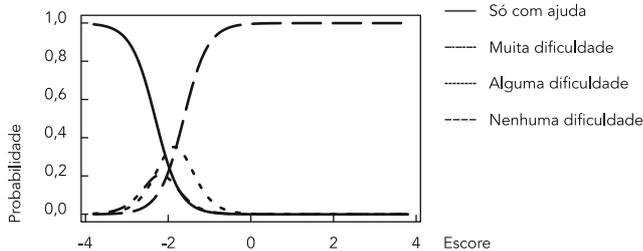
3c) Preparar seu próprio alimento



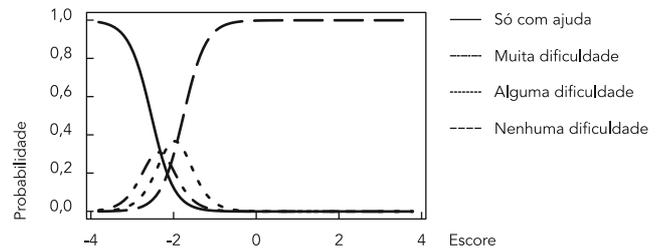
3d) Administrar o seu dinheiro



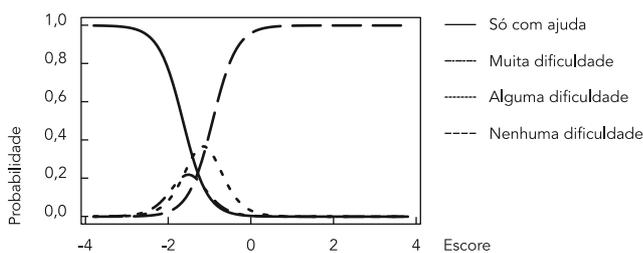
3e) Tomar seus remédios



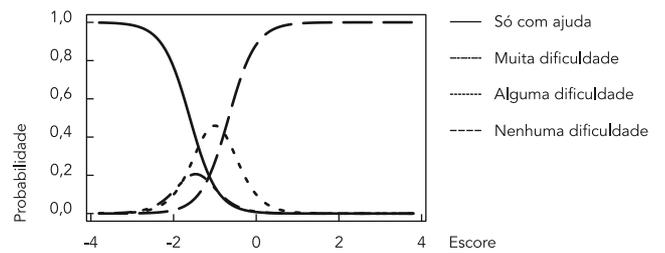
3f) Escovar os dentes, pentear os cabelos ou lavar o rosto



3g) Sair para fazer compras



3h) Cortar as unhas dos pés



3i) Sair de ônibus

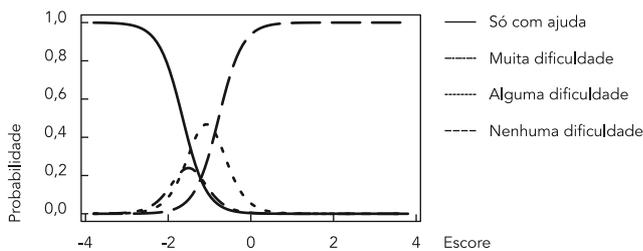
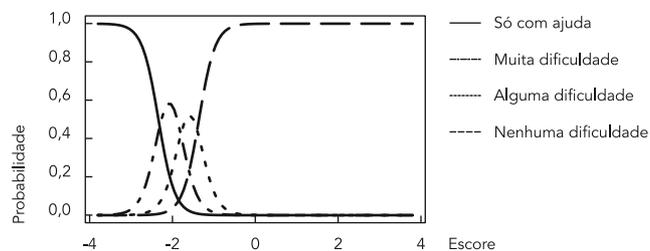


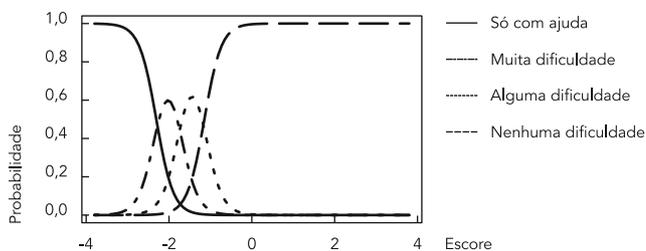
Figura 4

Curva característica dos itens referentes às atividades básicas da vida diária. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil, 2010.*

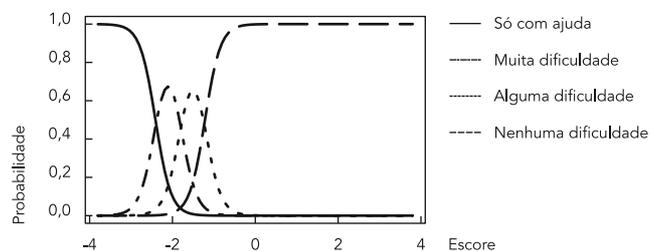
4a) Caminhar de um cômodo a outro num mesmo andar



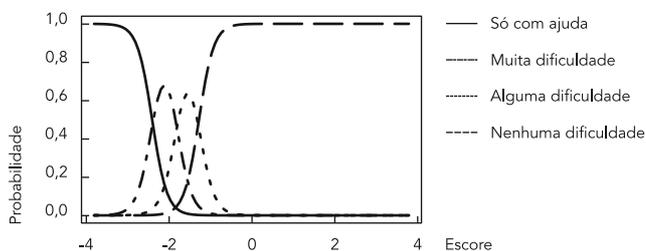
4b) Levantar-se de uma cadeira sem braço



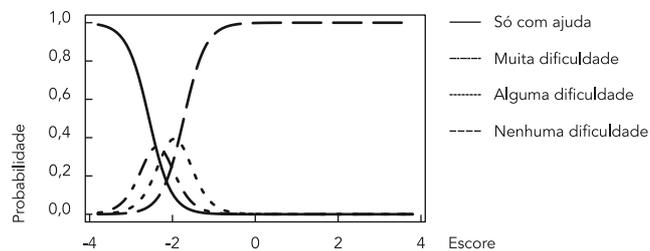
4c) Levantar-se da cama



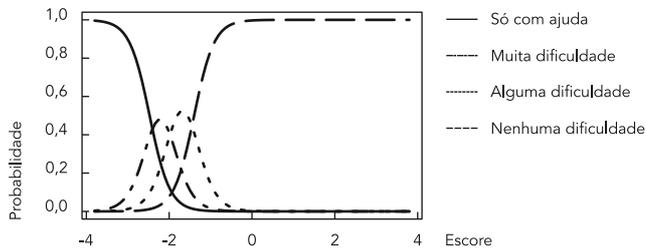
4d) Transferir-se da cama para uma cadeira



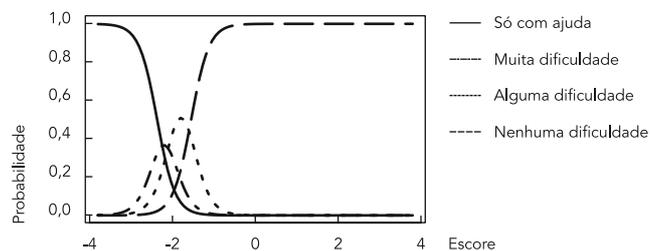
4e) Comer



4f) Vestir-se



4g) Tomar banho



4h) Usar o sanitário

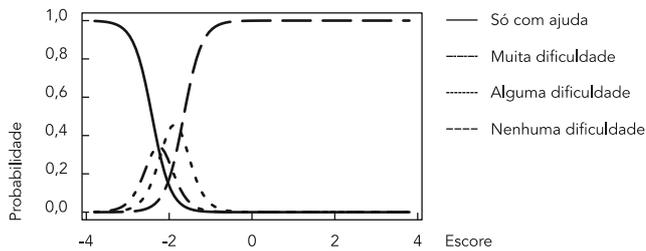
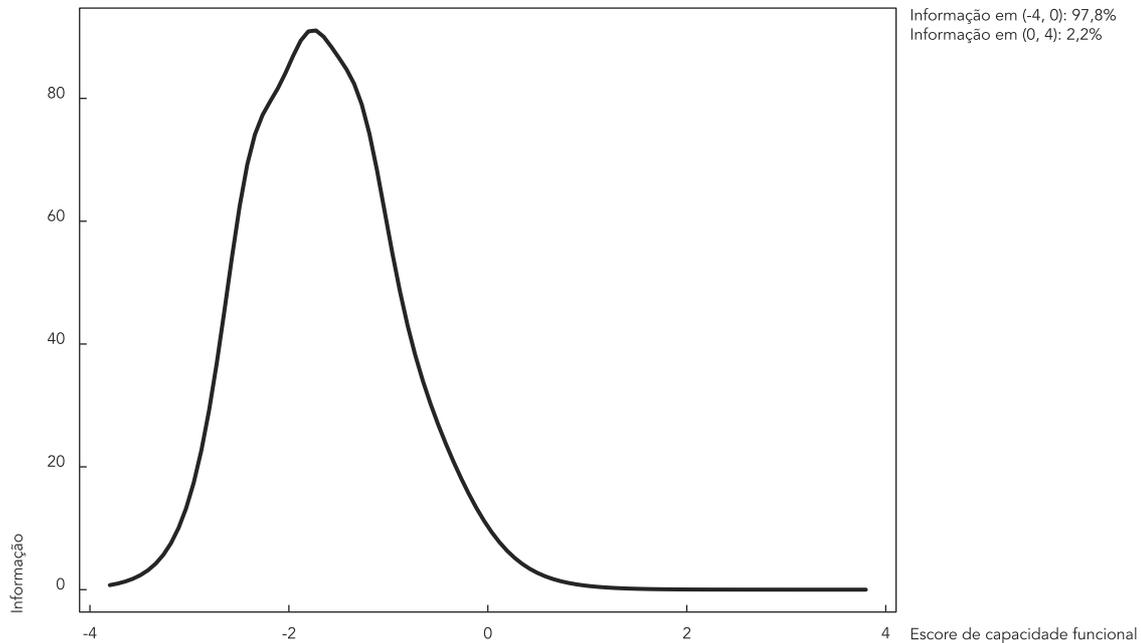


Figura 5

Curva de informação do teste. *Inquérito de Saúde da Região Metropolitana de Belo Horizonte*, Minas Gerais, Brasil, 2010.



que o escore de capacidade funcional é estimado com alta precisão para indivíduos situados abaixo da mediana. Para os indivíduos situados acima da mediana a precisão é muito pequena. Pouco mais de 2% da informação do teste está concentrada nessa faixa de capacidade funcional (Figura 5).

Discussão

Este artigo teve por objetivo avaliar os itens de capacidade funcional utilizados no *Inquérito de Saúde da Região Metropolitana de Belo Horizonte* de 2010, utilizando a TRI. Dos resultados da análise emergiram quatro conclusões importantes: (i) a análise fatorial da matriz de correlação policórica e o coeficiente alfa de Cronbach evidenciaram que o conjunto dos 21 itens está relacionado a uma única característica latente, ou seja, o instrumento é unidimensional; (ii) os itens não fornecem informação para toda a escala de funcionalidade, estando concentrados na sua metade inferior; (iii) há redundância nas questões incluídas no instrumento; (iv) os idosos não conseguem discernir entre as diversas categorias

de respostas que lhes são apresentadas para a maioria das questões. As três últimas conclusões foram obtidas baseando-se na TRI.

Existe controvérsia com relação à dimensionalidade das escalas de capacidade funcional. Um estudo realizado recentemente entre idosos brasileiros identificou três fatores ao avaliar a dimensionalidade de um instrumento contendo vinte atividades da vida diária, que incluíam ABVD, AIVD e atividades de mobilidade³⁰. Unidimensionalidade foi observada em outros trabalhos^{21,31}.

Na amostra estudada os limites das categorias foram todos negativos, variando de -2,56 a -0,36. Além disso, todas as questões apresentaram valores altos para o parâmetro de inclinação, indicando o alto poder de discriminação da questão na vizinhança dos limites das categorias. Esses resultados evidenciam que, embora tenham sido observadas 21 atividades com quatro opções de respostas, somente foi possível captar perdas consideráveis de capacidade funcional. Entrevistados com pequena limitação não tiveram sua situação avaliada pelo conjunto de questões utilizado, enquanto aqueles que apresentavam perda mais acentuada tiveram sua situação auferida

repetidas vezes. Observa-se, portanto, um cenário de redundância em um intervalo da escala e escassez no restante. Situação essa diagnosticada também pelo coeficiente alfa de Cronbach, que apresentou valor de 0,98. É sabido que valores do coeficiente muito superiores a 0,80 indicam a necessidade de redução da escala^{24,32}. Esse resultado é coerente com o que se observa nos estudos que avaliam a capacidade funcional por meio do número de atividades realizadas com dificuldade e encontra-se mais de 50% de respondentes relatando não apresentar dificuldade em nenhuma delas^{16,17,18}. A mesma situação de redundância em escala de capacidade funcional foi detectada por outros autores^{30,33}.

É importante salientar que muitas vezes utilizamos questões categóricas para mensurar variáveis que são de fato contínuas. É o que ocorre com as atividades da vida diária. É sabido que ao se categorizar uma variável perdemos informação, e o instrumento se torna menos eficiente para detectar efeitos ou associações com outras variáveis. Quanto menor o número de categorias maior será essa perda. Entretanto, o número de categorias é limitado pela capacidade que o respondente tem de discriminar entre elas²⁴. A análise das curvas características dos itens apresentadas nas Figuras 2, 3 e 4 sugere que, para a maioria das atividades analisadas, o número de opções de resposta superou a capacidade de discriminação dos avaliados. Para a maior parte das atividades, o número de categorias efetivamente utilizadas pelos respondentes foi inferior às quatro opções apresentadas. Mesmo nas situações em que foram usadas as quatro opções, a pequena distância entre elas na escala da capacidade funcional, para diversos itens, nos leva a concluir pela equivalência ao item dicotômico. Essa constatação nos indica a necessidade de avaliação das alternativas de resposta e o significado que elas possam ter para os respondentes.

A construção de instrumentos para avaliação de atributos complexos não é tarefa simples. Instrumentos que apresentam maior precisão e abrangência tendem a ser longos e complexos, o que pode comprometer a qualidade da informação. Na busca por formulários menores, é necessário um compromisso entre o tamanho do instrumento, a precisão da medida e a cobertura da escala do atributo.

No contexto de capacidade funcional, essa inquietação se coloca. Um instrumento adequado para uma população de doentes crônicos provavelmente será insuficiente para detectar dificuldades na população geral ou mesmo entre idosos saudáveis. Um instrumento desenvolvido para um estudo epidemiológico da população

em geral terá características distintas daquele desenvolvido para acompanhamento clínico de pacientes em um serviço de reabilitação musculoesquelética que, por sua vez, poderá ser distinto daquele sugerido para acompanhamento de pacientes de um serviço neurológico.

Apesar das propostas de mudanças de paradigmas relacionadas às questões do processo saúde/doença e da funcionalidade humana, instrumentos para avaliação da capacidade funcional têm sido construídos tendo como foco doentes crônicos. Assim, torna-se relativamente fácil mensurar o nível funcional de pessoas com deficiência grave, mas, por outro lado, aumenta-se a dificuldade de obter informação no outro extremo da escala, resultando na observação de um efeito piso considerável quando se aplica o instrumento³³. Estudos de base populacional, em que indivíduos sadios e com diversos níveis de comprometimento funcional são avaliados, requerem escalas de medidas de maior extensão, exigindo que se considere um espectro de atividades mais amplo e de dificuldade crescente no que se refere à mobilidade, força, resistência e cognição³⁴.

Por outro lado, é inquestionável a necessidade de se desenvolver instrumentos que forneçam resultados comparáveis, quando aplicados em diferentes momentos em uma mesma população ou a diferentes populações. Se desejamos avaliar a perda de capacidade funcional em uma abordagem longitudinal, por exemplo, devemos ser capazes de medir, com precisão, a capacidade funcional em todo o espectro da escala e não apenas em indivíduos com médio a alto comprometimento funcional. Considerando que o escore obtido por meio da TRI independe das características do instrumento, desde que as suposições do modelo sejam verificadas²⁸, a remodelagem do instrumento propiciaria uma estimativa mais precisa do escore de funcionalidade para os indivíduos com pouco comprometimento, sem prejuízo dos demais.

A análise dos itens do instrumento utilizado no *Inquérito de Saúde da Região Metropolitana de Belo Horizonte* sinaliza para a necessidade de se repensar a abrangência e o formato das questões utilizadas. Uma maior abrangência pode ser buscada por meio da inclusão de itens que cubram outros domínios contemplados na *Classificação Internacional de Funcionalidade, Incapacidade e Saúde* (CIF)¹. Quanto ao formato, se a redução do número de categorias de respostas não é desejável, é preciso ter em mente que manter categorias que os entrevistados não conseguem diferenciar pode causar uma falsa impressão de precisão, além de gerar uma carga maior para o respondente. Nossos dados sugerem que quatro

categorias, da maneira como estas foram apresentadas, é um número elevado para a população idosa. Talvez a redução para três categorias seja uma decisão de compromisso entre a perda de informação e a capacidade de discernimento dos entrevistados.

A existência de grande número de itens associados a um intervalo da escala coexistindo com extensão equivalente da escala associada a um pequeno número de itens, ou a nenhum, indica

a necessidade de se repensar o objetivo desta escala. Em especial, é importante considerar essa questão na perspectiva dos estudos longitudinais, em que se pretende avaliar mudanças. Nesse contexto é necessário dispor de instrumentos mais sensíveis, capazes de quantificar pequenos níveis de limitações. Somente assim será possível obter informações que subsidiem políticas de prevenção e promoção de saúde, que busquem atrasar o início da incapacidade em idosos.

Resumen

Este artículo tiene como objetivo la evaluación de las preguntas acerca de la capacidad funcional de los ancianos, presentes en la Encuesta de Salud de la Región Metropolitana de Belo Horizonte, Minas Gerais, Brasil, realizada en 2010, utilizando la Teoría de Respuesta al Ítem. Hemos analizado la dimensionalidad de la escala; la posición de los ítems en la misma; el poder de discriminación de los ítems y la precisión de la estimación de la capacidad funcional. El estudio se basó en una muestra de 2.174 ancianos de 60 a 99 años, 61% mujeres, con una educación media de 4 años de estudios. El instrumento de la capacidad funcional contenía 21

ítems (4 opciones de respuesta). El coeficiente alfa de Cronbach fue de 0,98 y el análisis factorial señaló uno de los factores, explicando el 92% de la variabilidad de las preguntas. Los resultados indicaron: redundancia en las preguntas; precisión de la estimación sólo para los ancianos que están por debajo de la media en la escala e incapacidad de los ancianos para diferenciar entre las opciones de respuesta. El análisis sugiere la necesidad de discutir las preguntas del instrumento.

Salud del Anciano; Evaluación de Programas e Instrumentos de Investigación; Análisis Multivariante

Colaboradores

C. C. César, J. V. M. Mambrini e M. F. Lima-Costa contribuíram com a concepção do artigo, análise dos dados, redação e revisão crítica do artigo e aprovação da versão final. F. R. Ferreira contribuiu com a concepção do artigo, revisão crítica e aprovação da versão final.

Agradecimentos

Agradecemos ao CNPq por apoiar as pesquisadoras C. C. César e M. F. Lima-Costa e financiar parcialmente este estudo por meio do projeto 474973/2011-5. À Capes pelo apoio à pesquisadora F. R. Ferreira.

Referências

1. World Health Organization. Towards a common language for functioning, disability and health. <http://www.who.int/classification/icf> (acessado em 15/Jul/2011).
2. Ramos LR. Fatores determinantes do envelhecimento saudável em idosos residentes em centro urbano: Projeto Epidoso, São Paulo. *Cad Saúde Pública* 2003; 19:793-8.
3. Lima-Costa MF, Oliveira CO, Macinko J, Marmot M. Socioeconomic inequalities in health in older adults in Brazil and England. *Am J Public Health* 2012; 102:1535-41.
4. Fillenbaum GG. The well-being of the elderly: approaches to multidimensional assessment. Geneva: World Health Organization; 1984.
5. Kane RA, Kane RL. Assessing the elderly: a practical guide to measurement. Lexington: Lexington Books; 1981.
6. Paixão Jr. CM, Reichenheim ME. Uma revisão sobre instrumentos de avaliação do estado funcional do idoso. *Cad Saúde Pública* 2005; 21:7-19.
7. Caldas CP. Envelhecimento com dependência: responsabilidades e demandas da família. *Cad Saúde Pública* 2003; 19:773-81.
8. Hébert 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.
9. Hébert R, Veil A, Raïche M, Dubois M-F, Dubuc N, Tousignant M. Evaluation of the implementation of PRISMA, a coordination-type integrated service delivery system for frail older people in Québec. *Journal of Integrated Care* 2008; 16:4-14.
10. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; 9:179-86.
11. Millán-Calenti JC, Tubío J, Pita-Fernández S, González-Abraldes I, Lorenzo T, Fernández-Arruty T, et al. Prevalence of functional disability in activities of daily living (ADL), instrumental activities of daily living (IADL) and associated factors, as predictors of morbidity and mortality. *Arch Gerontol Geriatr* 2010; 50:306-10.
12. Liao W, Li C, Lin Y, Wang C, Chen Y, Yen C, et al. Healthy behaviors and onset of functional disability in older adults: results of a national longitudinal study. *J Am Geriatr Soc* 2011; 59:200-6.

13. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA* 2010; 304:1787-94.
14. Barnato AE, Albert SM, Angus DC, Lave JR, De-genholtz HB. Disability among elderly survivors of mechanical ventilation. *Am J Respir Crit Care Med* 2011; 183:1037-42.
15. Santos CA, Dantas EEM, Moreira MHR. Correlation of physical aptitude; functional capacity, corporal balance and quality of life (QoL) among elderly women submitted to a post-menopausal physical activities program. *Arch Gerontol Geriatr* 2011; 53:344-9.
16. Griffith L, Raina P, Wu H, Zhu B, Stathoskostas L. Population attributable risk for functional disability associated with chronic conditions in Canadian older adults. *Age Ageing* 2010; 39:738-45.
17. Ferreira FR, César CC, Camargos VP, Lima-Costa ME, Proietti FA. Aging and urbanization: the neighborhood perception and functional performance of elderly persons in Belo Horizonte Metropolitan Area-Brazil. *J Urban Health* 2010; 87:54-66.
18. Clarke P, Smith J. Aging in a cultural context: cross-national differences in disability and the moderating role of personal control among older adults in the United States and England. *J Gerontol B Psychol Sci Soc Sci* 2011; 66:457-67.
19. Embreston SE. The continued search for nonarbitrary metrics in psychology. *Am Psychol* 2006; 61:50-5.
20. DeMars C. Item response theory: understanding statistics measurement. New York: Oxford University Press; 2010.
21. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000; 38(9 Suppl):II28-42.
22. Lima-Costa ME, Turci M, Macinko J. Saúde dos adultos em Belo Horizonte. Belo Horizonte: Núcleo de Estudos em Saúde Pública e Envelhecimento, Fundação Oswaldo Cruz/Universidade Federal de Minas Gerais; 2012.
23. Lin TH, Yao G. Evaluating item discrimination power of WHOQOL-BREF from an item response model perspectives. *Soc Indic Res* 2009; 91:141-53.
24. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th Ed. Oxford: Oxford University Press; 2008.
25. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institute for Educational Research; 1960.
26. Samejima F. Comprehensive latent trait theory. *Behaviormetrika* 1988; 24:1-24.
27. Andrade DF, Tavares HR, Valle RC. Teoria de resposta ao item: conceitos e aplicações. <http://www.inf.ufsc.br/~dandrade/TRI/> (acessado em 10/Ago/2011).
28. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah: Erlbaum; 2000.
29. Rizopoulos D. Irm: an R package for latent variable modelling and item response theory analyses. *J Stat Softw* 2006; 17:1-25.
30. Ramos LR, Andreoni S, Coelho-Filho JM, Lima-Costa ME, Matos DL, Rebouças M, et al. Screening for dependence in activities of daily living in the elderly: minimum set of questions. *Rev Saúde Pública* 2013; 47:506-13.
31. Spector WD, Fleishman JA. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J Gerontol* 1998; 53:S46-57.
32. DeVellis RE. Scale development: theory and applications. Thousand Oaks: Sage Publications; 2012.
33. Wolfe F, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II. A revised version of the Health Assessment Questionnaire. *Arthritis Rheum* 2004; 50:3296-305.
34. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reports health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010; 63:1179-94.

Recebido em 13/Jun/2014

Versão final reapresentada em 26/Jan/2015

Aprovado em 03/Mar/2015