

Escala Hamilton: estudo das características psicométricas em uma amostra do sul do Brasil*

*Hamilton Scale: study of the psychometric characteristics in a sample from Southern Brazil**

Manoela Ávila Freire¹, Vera Lúcia Marques de Figueiredo¹, Alina Gomide², Karen Jansen¹, Ricardo Azevedo da Silva¹, Pedro Vieira da Silva Magalhães³, Flávio Pereira Kapczinski³

RESUMO

Objetivo: Investigar as características psicométricas de uma versão traduzida da escala, propondo uma Versão Revisada que atenda aos critérios de adaptação transcultural para o contexto brasileiro. **Métodos:** Este estudo incluiu 231 sujeitos – deprimidos (45,5%), bipolares (7,8%) e saudáveis (46,7%) – que participaram de uma pesquisa epidemiológica no sul do Brasil. A avaliação de transtornos mentais foi realizada por meio da *Clinical Interview for DSM-IV* (SCID) e uma versão traduzida da Escala de Avaliação de Depressão de Hamilton (HAM-D), que habitualmente vem sendo utilizada no país sem estudos de adaptação. **Resultados:** Identificou-se o ponto de corte (9 pontos) para discriminar a presença ou não de sintomas de depressão pela análise da curva ROC, resultando em uma sensibilidade e especificidade de 90 e 91%, respectivamente. A validade interna foi investigada pela análise fatorial e consistência dos itens. Dos 17 itens originais, apenas o item que avalia a “consciência do transtorno” não apresentou carga fatorial satisfatória para avaliar depressão geral e foi eliminado; os 16 restantes agruparam-se em cinco dimensões, denominadas: Humor deprimido, Anorexia, Insônia, Somatização e Ansiedade, as quais, com exceção da última, mostraram homogeneidade nos seus construtos (coeficientes alfa entre 0,66 e 0,78). Na análise de conteúdo dos itens, cinco especialistas sugeriram alterações redacionais em sete itens. **Conclusão:** O estudo determina um ponto de corte diferente do original e evidencia características psicométricas favoráveis para a utilização da escala no Brasil.

Palavras-chave

Depressão, estudos de validação, sensibilidade, especificidade.

ABSTRACT

Objective: To investigate the psychometric characteristics of a translated version of the scale, proposing a reviewed version in order to attend the transcultural adaptation criteria. **Methods:** This study included 231 subjects – depressed (45.5%), bipolar (7.8%), and healthy (46.7%) – who participated of an epidemiological research in southern Brazil. The evaluation of mental disorders was made through Clinical Interview for DSM-IV (SCID) and a translated version of the Hamilton Scale (HAM-D), usually utilized in Brazil without adaptation studies. **Results:** The ROC curve analysis identified the cutoff (9 points) to discriminate the presence or absence of depression, resulting in a sensibility and specificity of 90 and 91%, respectively.

Recebido em
14/7/2014
Aprovado em
29/10/2014

1 Universidade Católica de Pelotas (UCPel).
2 Universidade Federal de Minas Gerais (UFMG).
3 Universidade Federal do Rio Grande do Sul (UFRGS).

* Manuscrito desenvolvido como trabalho de conclusão de curso da autora principal, em 2013/1.

Endereço para correspondência: Vera Lúcia Marques de Figueiredo
Universidade Católica de Pelotas
Rua Félix da Cunha, 412
96010-000 – Pelotas, RS, Brasil
E-mail: verafig@terra.com.br

Keywords

Depression, validation studies, sensibility, specificity.

INTRODUÇÃO

A Escala de Avaliação de Depressão de Hamilton (HAM-D), criada por Max Hamilton *et al.*¹, foi construída na década de 1960 para ser utilizada exclusivamente em pacientes previamente diagnosticados com transtorno afetivo do tipo depressivo. Em função da organização e da escolha de seus itens, eles servem para identificar a gravidade dos sintomas depressivos, e não sua existência¹. A escala original inglesa era composta por 21 itens, mas o próprio autor sugeriu, posteriormente, que os quatro últimos itens (variação diurna, despersonalização/desrealização, sintomas paranoides e sintomas obsessivo-compulsivos) fossem retirados porque eram menos frequentes e contribuíam para definir o tipo de depressão, e não sua intensidade²⁻⁴. Existe, também, uma versão com 24 itens, sendo desamparo, desesperança e desvalia os itens adicionais⁵.

Há 50 anos, a HAM-D vem sendo considerada como “padrão-ouro” de referência para estudos de validação de outras escalas¹. Além disso, foi (e continua sendo) amplamente utilizada em ensaios clínicos para testar a eficácia de determinados antidepressivos² devido à ênfase que dá aos sintomas somáticos em oposição aos sintomas cognitivos ou afetivos, que se encontram em minoria⁶. Ela é amplamente sensível às mudanças vivenciadas por pacientes acometidos por depressão grave^{5,7} e é eficiente em discriminar droga e placebo nestes estudos^{3,8,9}. Por outro lado, há também várias críticas¹⁰ sobre a escala, relacionadas a itens inadequados para avaliar a gravidade da depressão; baixos índices de confiabilidade; alternativas de respostas impróprias e poucas evidências da multidimensionalidade da escala.

A versão da HAM-D mais utilizada é composta por 17 itens¹. O tempo de aplicação varia em torno de 15 a 30 minutos⁴. O autor sugeriu que a escala fosse administrada por meio de entrevista, aplicada em parceria de dois clínicos bem treinados e experientes: um para conduzir a entrevista e o outro para fazer perguntas adicionais ao final¹. A diferenciação das pontuações em 3 e 5 pontos nos itens foi explicada por Hamilton¹⁰, que acreditava que certos aspectos seriam difíceis de serem mensurados, atribuindo, portanto, poucas opções de respostas, ao passo que o resultado de determinados itens contribuía mais para a pontuação total do que outros.

The internal validity was investigated by the factorial analysis and consistence of the items. It was observed that all 17 original items, except “Consciousness”, presented psychometric quality to evaluate general depression, and that there were five dimensions underlying those 16 items: Depressed humor, Anorexia, Insomnia, Somatization, and Anxiety, and all of them, excepting the last, showed homogeneity in their constructs (alpha coefficients between 0.66 and 0.78). On content analysis, five specialists suggested editing changes in seven items.

Conclusion: This study determinates a different cutoff and psychometric evidences favourable to the use of HAM-D in Brazil.

Não se encontram na literatura pontos de corte determinados pelo autor da escala, aceitando-se, na prática clínica, escores acima de 25 pontos como característicos de pacientes gravemente deprimidos; escores entre 18 e 24 pontos, pacientes moderadamente deprimidos; e escores entre 7 e 17 pontos, pacientes com depressão leve⁵. Por outro lado, um estudo japonês¹¹ com base em sete ensaios clínicos determinou, a partir dos instrumentos HAM-D e *Clinical Global Impression-Severity* (SCID), uma interpretação diferente, em que pontuações de 0 a 3 correspondem a pacientes normais; escores de 4 a 7, pacientes limítrofes; escores de 8 a 15, pacientes ligeiramente doentes; pontuações de 16 a 26, pacientes moderadamente doentes e escores ≥ 27 , pacientes gravemente doentes. Em uma revisão com sete artigos¹⁰, identificou que os pontos de corte na HAM-D variaram de 10 a 16, com média de 12,6/13,5. Considerando os estudos encontrados, pode-se concluir que não há consenso entre o valor do ponto de corte da escala HAM-D para diferenciar a gravidade dos sintomas de depressão.

Apesar de a HAM-D ser utilizada com muita frequência, outras críticas foram, também, direcionadas ao instrumento¹⁻³. Duas delas estão relacionadas à falta de padronização da escala com 17 itens e a falta de procedimentos de pontuação precisa³. Além disso, no Brasil, os profissionais não contam com uma versão adaptada, mas apenas com versões traduzidas^{1,3,4}.

Originalmente, a escala de 17 itens foi analisada pelo próprio autor, que identificou a presença de três fatores, primariamente nomeados como Depressão Geral, Agitação *versus* Retardo e Insônia^{6,12}. Em outro estudo⁴, ele encontrou quatro fatores, os quais não foram nomeados e tais itens mostraram-se na sua maioria complexos, não caracterizando uma estrutura satisfatória⁴. O'Brien e Glaudin¹³ identificaram a presença de quatro fatores diferentes, denominando as dimensões como: “Queixas Somáticas”, “Anorexia”, “Distúrbios do Sono” e “Agitação/Retardo” e sugerindo a exclusão de alguns itens que não se mostraram consistentes para a avaliação. Já em outra pesquisa¹⁴ os quatro fatores subjacentes aos itens da escala foram denominados como: “Ansiedade Somática/Fator de Somatização”, “Ansiedade Psíquica”, “Depressivo Puro” e “Fator de Anorexia”.

Um modelo de cinco fatores¹⁵ foi identificado, sendo quatro deles denominados como “Insônia”, “Ansiedade”,

“Melancolia” e “Sintomas Vegetativos”; o último não recebeu nenhuma classificação. Nesse estudo, a consistência interna da escala foi identificada com $\alpha = 0,72$. Um modelo similar também foi encontrado, sem apresentar o nome das cinco dimensões¹⁶, entretanto, foi identificado um fator geral.

Ainda, encontrou-se na literatura a presença de seis dimensões subjacentes aos 17 itens da HAM-D¹⁷. Um dos trabalhos não apresentou a denominação de tais fatores, apesar de divulgar a carga fatorial dos itens e o coeficiente de consistência interna da escala ($\alpha = 0,75$)¹⁷. Por outro lado, quanto à dimensionalidade da escala⁹, foram identificados seis aspectos sobre os quais se agrupavam os itens: humor, somáticos, motores, sociais, cognitivos e ansiedade¹⁹. Já em outra revisão¹⁰ que reuniu estudos de janeiro de 1980 até maio de 2003, foram identificados trabalhos que mencionaram a presença de análise fatorial e componentes principais, que se referiram à presença de dois a oito fatores, levando os autores a concluir que a escala não parece unidimensional, mas a multidimensionalidade desta também não fica evidente – ideia também compartilhada por Pancheri *et al.*¹⁴.

Bagby *et al.*¹⁰ analisaram 15 artigos que descreviam análises fatoriais da escala HAM-D. Os estudos de confiabilidade apontaram coeficientes que variaram de 0,46 a 0,97 e outros relataram estimativas $\geq 0,70$ ^{10,15,17}, sugerindo índices satisfatórios. Quanto à validade de conteúdo, encontra-se na literatura a crítica de que a HAM-D não abrange todos os critérios diagnósticos presentes no DSM-IV, tais como anedonia, indecisão e sinais vegetativos¹⁸.

A escala HAM-D é amplamente utilizada no Brasil, tanto em pesquisas quanto na prática clínica e, portanto, torna-se necessário investigar as características psicométricas e selecionar uma forma padronizada e adaptada para a sua administração¹⁹, uma vez que não se encontram informações sobre o processo de tradução da versão brasileira. Testes e escalas apenas traduzidos não são confiáveis para mensurar a patologia da nova população²⁰, tendo em vista que as características psicométricas do instrumento podem ser diferentes.

O objetivo do estudo foi examinar as características psicométricas da versão tradicionalmente utilizada no Brasil da escala HAM-D, determinando o ponto de corte adequado para uma amostra de indivíduos, moradores de uma região do sul do país. Investigou-se, também, a qualidade da tradução, buscando uma alternativa que atenda mais satisfatoriamente aos critérios de adaptação transcultural.

MÉTODOS

Para validação da escala HAM-D, foram desenvolvidos dois estudos. O primeiro estudo visou determinar a sensibilidade e a especificidade do instrumento num grupo clínico; o segundo, realizar a análise de conteúdo de uma versão habitualmente utilizada no Brasil.

Estudo 1

Amostra

Os participantes foram selecionados a partir de um estudo de base populacional²¹, que incluiu 1.560 jovens de 18 a 27 anos, do município de Pelotas, RS. Destes, 93 supostos casos de transtorno bipolar foram pareados por sexo e idade para outros dois grupos: com depressão e controles populacionais, resultando em um total de 279 sujeitos para a segunda fase. Na segunda fase, 48 sujeitos não foram localizados ou se recusaram a participar do estudo. Dessa forma, a escala HAM-D foi respondida por 231 sujeitos, dos quais 105 pontuaram positivamente os critérios de depressão da SCID (critério externo) e 126 sujeitos sem nenhum transtorno ou diagnóstico de mania, segundo os instrumentos utilizados.

Instrumentos

Escala de Avaliação de Depressão de Hamilton¹

O instrumento investiga como o paciente tem se sentido nos últimos sete dias, incluindo o dia da aplicação. Foi empregada a versão traduzida da escala, que usualmente vem sendo utilizada no país¹. Ela é composta por 17 itens, os quais podem ser pontuados numa escala *Likert* que varia entre 0 a 2 ou 0 a 4, conforme a intensidade do sintoma. O total de pontos varia entre 0 e 52 pontos. Para ser constatada a presença de depressão, os escores devem somar, no mínimo, 8 pontos, na versão original²⁰.

Structured Clinical Interview for DSM-IV (SCID)

Este instrumento consiste em uma entrevista diagnóstica semiestruturada criada a partir do DSM-IV. As respostas identificarão a presença ou não dos sintomas, sendo pontuadas de acordo com o julgamento do avaliador. Ela é composta por 10 módulos, que podem ser utilizados de maneira combinada ou independente²². No estudo, foi utilizado o módulo “A” para diagnosticar os episódios de humor (transtornos bipolar e depressivo maior). A tradução e a adaptação desta entrevista clínica para o idioma português apresentam, em geral, boa confiabilidade para os transtornos de humor²³.

Entrevista de Avaliação Socioeconômica

A entrevista traz questões que se baseiam no instrumento da Associação Brasileira de Empresas de Pesquisa (ABEP) e identificam o acúmulo de bens materiais e a escolaridade do chefe da família²⁴. Essa classificação identifica as pessoas em classes (A, B, C, D, ou E) a partir dos escores alcançados, na qual a letra “A” refere-se à classificação socioeconômica mais alta e “E”, a mais baixa.

Procedimentos para coleta de dados

O treinamento dos avaliadores para a realização das entrevistas clínicas (HAM-D e SCID) foi ministrado pela equipe do Laboratório de Psiquiatria Molecular da Universidade Federal

do Rio Grande do Sul (UFRGS), em função da pesquisa base deste estudo. As entrevistadoras foram três psicólogas treinadas para realizar a entrevista de acordo com os critérios do DSM-IV apresentados na SCID. A aplicação dos instrumentos foi feita pelas psicólogas, em forma de entrevista. As escalas foram administradas de forma individual, em uma sala de atendimento do Hospital Universitário da Universidade Católica de Pelotas (UCPel). Posteriormente, as informações foram digitadas em um banco de dados.

Procedimentos para análise de dados

Os dados foram analisados, considerando os seguintes procedimentos:

- Descrição das variáveis por meio de frequências e médias.
- Avaliação da distribuição dos escores por meio do teste de Kolmogorov-Smirnov.
- Comparação dos resultados entre os grupos clínico/controlado pelo Teste U de Mann-Whitney.
- Determinação do ponto de corte da sensibilidade e especificidade pela curva ROC.
- Análise da validade interna por duas técnicas. A primeira foi a análise fatorial exploratória, utilizando-se o método dos componentes principais com rotação Varimax, que considera as dimensões independentes, conforme estudos da literatura^{10,13}; foram mantidos os itens com carga fatorial $\geq 0,30$. Utilizaram-se os procedimentos exploratórios considerando a busca por dimensões subjacentes ao domínio avaliado pela escala; e por não haver expectativa, *a priori*, de sua composição teórica²⁵. A segunda, pela consistência interna (alfa de Cronbach), adotando-se o coeficiente 0,80 como razoável²⁶.
- A validade de critério entre os resultados da HAM-D e da SCID – dicotomizados entre presença/ausência – foi analisada pela correlação de Spearman.

Estudo 2

Amostra

Para a etapa da análise teórica dos itens da HAM-D, participaram cinco especialistas, sendo dois psicólogos e três psiquiatras. Os profissionais tinham domínio no uso da escala e na língua inglesa.

Instrumentos

Para analisar a qualidade da versão da escala HAM-D usada no Brasil, foi elaborado um protocolo para análise teórica dos itens, que consistia em um formulário impresso com instruções para preenchimento e com os 17 itens originais transcritos em inglês, juntamente com os itens da versão em português, usualmente utilizada no Brasil. Ao lado, apareciam as seguintes alternativas: (A) Tradução adequada ou inadequa-

da; (B) Conteúdo adequado ou inadequado; (C) Sugestões para a nova redação do item.

Procedimentos para coleta de dados

O protocolo com o rol dos itens da escala HAM-D foi entregue aos peritos, que o preencheram de forma individual no próprio local de trabalho. Posteriormente, os dados foram registrados numa planilha eletrônica.

Procedimentos para análise de dados

Para a análise de conteúdo dos itens, utilizou-se o percentual de concordância entre avaliadores, detendo-se itens com 80% de pareceres favoráveis²⁷ quanto à tradução e ao conteúdo.

Procedimentos éticos

O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Católica de Pelotas (protocolo 2006/96). Os jovens e os profissionais receberam informações sobre os objetivos da pesquisa e assinaram um “Termo de Consentimento Livre e Esclarecido”. Assegurou-se o direito à confidencialidade dos dados e o cuidado na utilização das informações nos trabalhos escritos. Os jovens que apresentaram sintomas relativos a algum transtorno mental na entrevista clínica do estudo principal foram encaminhados para o ambulatório de psiquiatria da Universidade.

RESULTADOS

Os participantes do estudo eram, na maioria, mulheres (68,4%). Elas tinham em média 22,04 anos ($dp = 2,18$) e as idades variaram entre 18 e 27 anos, com mediana de 23 anos. A escolaridade variou entre 1 e 18 anos de estudos, sendo a média de 9,18 anos ($dp = 3,14$ anos) e mediana, 10 anos. Quanto à classe social, a maior concentração de pessoas era da classe “C” (51,5%). De acordo com a SCID, 105 jovens tiveram diagnóstico de depressão (45,5%), sendo 83 mulheres e 22 homens. Os demais apresentavam diagnóstico de mania (4,8%), hipomania (3%) e sem transtorno de humor (46,7%). Ao analisar as discrepâncias entre os escores médios na HAM-D obtidos pelos pacientes depressivos, não se encontraram diferenças significativas relacionadas a sexo, nível socioeconômico e anos de escolaridade. O escore médio de pontos na HAM-D foi de 17 pontos ($dp = 6,63$). Os dados sociodemográficos dos participantes deprimidos mantiveram características similares aos da literatura, no que se refere à predominância do sexo feminino¹⁴⁻¹⁷ e à pontuação na escala HAM-D16; entretanto, diferenciou-se em relação à média de idade configurando um grupo mais jovem, uma vez que nos estudos descritos a média de idade dos sujeitos variou entre 35 e 49 anos¹⁴⁻¹⁷. As características dos sujeitos participantes do estudo aparecem descritas na tabela 1.

Tabela 1. Características sociodemográficas da amostra

Características	Frequência	Porcentagem
Sexo		
Feminino	158	68,4
Masculino	73	31,6
Idade		
De 18 a 22 anos	125	54,1
De 23 a 27 anos	106	45,9
Escolaridade		
De 1 a 4 anos	21	9
De 5 a 8 anos	54	23,4
De 9 a 12 anos	107	46,3
13 anos ou mais	18	7,8
Classe social		
a+b	73	31,6
c	119	51,5
d+e	39	16,9

Determinação dos pontos de corte

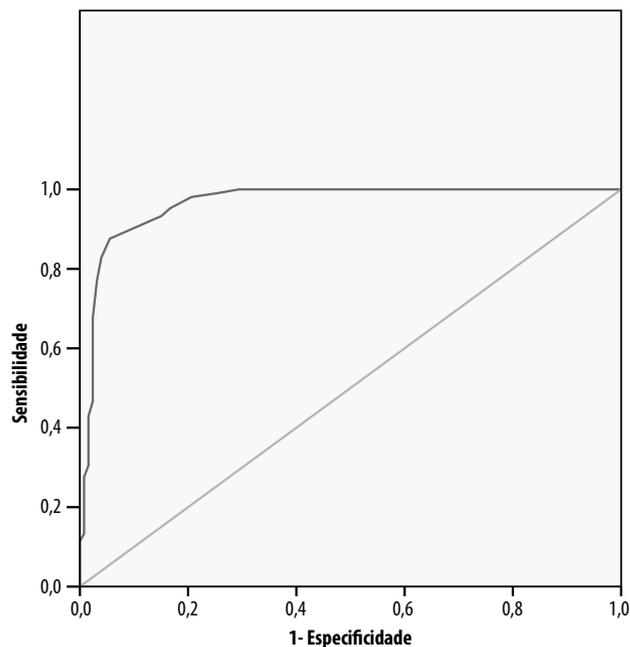
O resultado do teste de Kolmogorov-Smirnov indicou que os escores na HAM-D não apresentaram uma distribuição normal ($z = 2,5$; $p < 0,001$). Por isso, para comparação das pontuações na HAM-D entre os grupos clínico e controle classificados de acordo com a entrevista clínica (SCID), tendo como referência a variável "depressão atual", foi utilizado o teste de Mann-Whitney.

A pontuação média na HAM-D para o grupo que relatou sintomas depressivos atuais foi de 16,51 ($dp = 6,63$) pontos e, para o grupo que não relatou sintomas depressivos atuais, foi de 2,42 ($dp = 4,15$) pontos. A comparação realizada por meio do teste U de Mann-Whitney indicou a presença de diferença significativa entre os grupos ($U = 466,50$, $z = 12,32$, $p < 0,001$).

A curva ROC foi utilizada para verificar a capacidade dos escores da HAM-D e para discriminar os grupos em relação à presença de sintomas depressivos atuais, quando comparados com os obtidos na SCID. A área sob a curva foi igual a 0,97 (erro-padrão = 0,01, $p < 0,001$) com intervalo de confiança de 95% variando entre 0,94 e 0,99 (Figura 1). Esse resultado indicou que a HAM-D apresentou uma capacidade adequada para discriminar indivíduos com sintomas depressivos.

A tabela 2 apresenta os valores preditivos da HAM-D em diferentes pontos de corte que variaram entre 4 e 12 pontos e os respectivos valores de sensibilidade e especificidade. Considerou-se o escore de 9 como ponto de corte satisfatório, determinando no instrumento uma sensibilidade de 0,90 e especificidade de 0,91. Esse resultado significa 90% de chance de o teste identificar corretamente indivíduos que têm depressão. De acordo com o diagnóstico médico, realizado por meio da entrevista diagnóstica SCID (verdadeiro-

-positivo), existe 91% de chance de excluir sujeitos isentos de doença, ou seja, 9% de chance de identificar como tendo depressão sujeitos saudáveis (falso-negativo). Para a escolha do ponto de corte, considerou-se pelo valor especificidade $>$ sensibilidade porque, quando o teste é altamente específico, ele identifica com segurança os resultados falso-positivos^{28,29}.

**Figura 1.** Curva ROC dos pontos de corte da HAM-D.**Tabela 2.** Pontos de corte, sensibilidade, especificidade, valores preditivos positivos e negativos para a HAM-D

Pontos de corte	Sensibilidade	Especificidade	Valor preditivo positivo	Valor preditivo negativo
4	0,99	0,75	0,76	0,99
5	0,98	0,79	0,80	0,98
6	0,95	0,83	0,83	0,95
7	0,93	0,85	0,84	0,94
8	0,91	0,88	0,86	0,92
9	0,90	0,91	0,89	0,91
10	0,88	0,94	0,93	0,90
11	0,83	0,96	0,94	0,87
12	0,77	0,97	0,95	0,97

Evidências de validade interna

Para examinar a unidimensionalidade da escala, utilizou-se o método da Análise dos Componentes Principais, a partir do qual foi identificado um fator geral composto por 16 itens, que explicam 19,63% da variância dos resultados. As cargas fatoriais dos itens variaram entre 0,32 (ansiedade psíquica) e 0,78 (insônia tardia), sugerindo que os 16 itens avaliavam depressão geral. Constatada a fatorabilidade da matriz (Bartlett's = 485,29; KMO = 0,69), identificou-se a presença de

cinco componentes que explicaram juntos 62,29% da variância dos resultados, os quais mostraram associação entre eles.

Considerando que a escala HAM-D não apresenta um único modelo multidimensional, com base em estudos da literatura¹³⁻¹⁶, os pesquisadores denominaram os cinco componentes como: Humor deprimido (F1), Anorexia (F2), Insônia (F3), Somatização (F4) e Ansiedade (F5), conforme a tabela 3. Observaram-se alguns itens complexos, optando-se para definir o fator representante, aquele em que o item apresentou maior carga, além da interpretabilidade teórica. Entre eles, estão os itens “Trabalho e Atividades” (0,404 no F2), “Retardo” (-0,346 no F5), “Suicídio” (0,355 no F5) e “Sintomas somáticos em geral” (0,311 no F4).

Tabela 3. Estrutura fatorial da HAM-D

Itens	Fator				
	1	2	3	4	5
1 Humor deprimido	0,782				
2 Sentimentos de culpa	0,493				
3 Suicídio	0,667				
4 Insônia inicial			0,715		
5 Insônia intermediária			0,821		
6 Insônia tardia			0,859		
7 Trabalho e atividades	0,491				
8 Retardo	0,780				
9 Agitação					0,677
10 Ansiedade psíquica					0,426
11 Ansiedade somática				0,766	
12 Sintomas somáticos gastrointestinais		0,731			
13 Sintomas somáticos em geral				0,751	
14 Sintomas genitais					0,718
15 Hipocondria				0,745	
16 A Perda de peso (segundo paciente)		0,815			
16 B Perda de peso (segundo especialista)		0,863			
Número de itens	5	3	3	3	3
% de variância	19,63	14,34	11,09	10,15	7,08
Consistência interna (α)	0,71	0,78	0,74	0,66	0,38

Pela análise da consistência interna, os coeficientes demonstraram homogeneidade satisfatória entre os itens das quatro dimensões. Apenas o fator 5 (F5) demonstrou menor representatividade para o construto da depressão. A consistência interna da HAM-D, excluindo o item “Consciência”, foi de 0,72. Para os fatores, a variação é de 0,38 a 0,78. A validade de critério foi estabelecida pela correlação dos escores obtidos na escala HAM-D e a presença ou não de depressão encontrada na escala SCID, identificando-se associação entre os dois instrumentos ($r = 0,81$; $p = 0,00$).

Análise de conteúdo

A validação do conteúdo da versão da escala utilizada no país com especialistas resultou em sugestões relativas à redação. Dos 17 itens da escala, nove foram apontados com algum problema de tradução. Desses itens, três deles (humor deprimido, insônia tardia e ansiedade psíquica) não tiveram a concordância de 80% dos juízes no quesito “tradução adequada”. Quanto à análise referente à pertinência da sua representatividade ao construto depressão, todos os itens foram considerados adequados por 80% ou mais dos juízes. Na tabela 4, aparecem as sugestões mais relevantes apresentadas pelos especialistas.

DISCUSSÃO

Os resultados obtidos com os pacientes depressivos caracterizaram a capacidade da HAM-D em detectar o transtorno, uma vez que eles apresentaram características significativamente diferentes do grupo de controle. As análises para determinar a sensibilidade e a especificidade do instrumento tradicionalmente utilizado no Brasil, com o mesmo ponto de corte da versão original em inglês⁸, indicaram discrepância entre os pontos de corte das versões em português⁹ e versão original em inglês¹⁰. Em função da divergência dos pontos de corte, a escala pode estar produzindo escores equivocados, identificando casos falso-positivos ou verdadeiro-negativos.

A análise da estrutura interna da escala HAM-D evidenciou que os itens avaliam o construto geral da depressão sob cinco dimensões. Dessas cinco, apenas o fator 5 (Ansiedade) demonstrou menor confiabilidade para avaliar a depressão. Os dados mostraram resultados diferentes dos encontrados na literatura, que identificam entre três^{6,12} e seis¹ dimensões subjacentes. Quanto à consistência interna da escala como um todo, o índice encontrado foi similar aos observados na literatura, referente às versões adaptadas^{15,17}. Dos cinco fatores, quatro deles apresentaram coeficientes que podem ser classificados como satisfatórios e muito bons, entretanto, não é possível fazer alguma comparação com a versão original por não haver dados relativos à homogeneidade da escala¹⁰. No estudo, os baixos valores da fidedignidade podem estar associados ao número reduzido da amostra e à pouca variabilidade, uma vez que as análises foram realizadas apenas entre os pacientes depressivos. Da mesma forma, os baixos índices nos fatores podem ser associados ao pequeno número de itens que constituiu cada dimensão²⁷.

Quanto à validade de critério, encontraram-se na literatura correlações da HAM-D com a *Beck Depression Inventory - BDI* ($r = 0,48$) e com a *Clinical Global Impression - CGI* ($r = 0,56$)¹⁷. No presente estudo, o grau de eficácia da escala HAM-D, em prever o desempenho dos indivíduos, foi determinado pela técnica do diagnóstico psiquiátrico, utilizando-se a SCID.

Tabela 4. Itens da escala HAM-D e sugestões para a versão revisada

Item original	Versão experimental	Sugestão
1. Depressed mood (sadness, hopeless, helpless, worthless) 0. Absent 1. These feeling states indicated only on questioning 2. These feeling states spontaneously reported verbally 3. Communicates feeling states nonverbally, i.e., through facial expression, posture, voice and tendency to weep 4. Patient reports VIRTUALLY ONLY these feeling states in his spontaneous verbal and nonverbal communication	1. Humor deprimido (tristeza, desesperança, desamparo, inutilidade) 0. Ausente 1. Sentimentos relatados apenas ao ser inquirido 2. Sentimentos relatados espontaneamente com palavras 3. Comunica os sentimentos não com palavras, isto é, com a expressão facial, a postura, a voz e a tendência ao choro 4. Sentimentos deduzidos da comunicação verbal e não verbal do paciente	1. Humor deprimido (tristeza, desesperança, desamparo, inutilidade) 0. Ausente 1. Sentimentos relatados apenas ao ser inquirido 2. Sentimentos relatados espontaneamente com palavras 3. Comunica os sentimentos não com palavras, isto é, com a expressão facial, a postura, a voz e a tendência ao choro 4. Sentimentos deduzidos da comunicação verbal e não verbal espontânea do paciente
4. Insomnia – Early 0. No difficulty falling asleep 1. Complains of occasional difficulty falling asleep i.e., more than ½ hour 2. Complains of nightly difficulty falling asleep	4. Insônia inicial 0. Sem dificuldades para conciliar o sono 1. Queixa-se de dificuldade ocasional para conciliar o sono, isto é, mais de meia hora 2. Queixa-se de dificuldade para conciliar o sono todas as noites	4. Insônia inicial 0. Sem dificuldades para iniciar o sono 1. Queixa-se de dificuldade ocasional para iniciar o sono, isto é, mais de meia hora 2. Queixa-se de dificuldade para iniciar o sono todas as noites
6. Insomnia – Late 0. No difficulty 1. Waking in early hours of the morning but goes back to sleep 2. Unable to fall asleep again if gets out of bed	6. Insônia tardia 0. Sem dificuldades 1. Acorda de madrugada, mas volta a dormir 2. Incapaz de voltar a conciliar o sono se deixar a cama	6. Insônia tardia 0. Sem dificuldades 1. Acorda no início da manhã, mas volta a dormir 2. Incapaz de voltar a conciliar o sono se deixar a cama
9. Agitation 0. None 1. Fidgetiness 2. Playing with hands, hair, etc. 3. Moving about, cant't sit still 4. Hand wringing, nail biting, hair-pulling, biting of lips	9. Agitação 0. Nenhuma 1. Inquietude 2. Brinca com as mãos, com os cabelos etc. 3. Mexe-se, não consegue sentar quieto 4. Torce as mãos, rói as unhas, puxa os cabelos, morde os lábios	9. Agitação 0. Nenhuma 1. Inquietude 2. Mexe com as mãos, com os cabelos etc. 3. Mexe-se, não consegue sentar quieto 4. Torce as mãos, rói as unhas, puxa os cabelos, morde os lábios
10. Anxiety – Psychic 0. No difficulty 1. Subjective tension and irritability 2. Worrying about minor matters 3. Apprehensive attitude apparent in face or speech 4. Fears expressed without questioning	10. Ansiedade psíquica 0. Sem dificuldade 1. Tensão e irritabilidade subjetivas 2. Preocupação com trivialidades 3. Atitude apreensiva aparente no rosto ou na fala 4. Medos expressos sem serem inquiridos	10. Ansiedade psíquica 0. Sem dificuldade 1. Tensão e irritabilidade subjetivas 2. Preocupação com trivialidades 3. Atitude apreensiva aparente no rosto ou na fala 4. Medos expressos sem serem questionados
15. Hypochondriasis 0. Not present 1. Self-absorption (bodily) 2. Preoccupation with health 3. Frequent complaints, requests for help, etc. 4. Hypochondriacal delusions	15. Hipocondria 0. Ausente 1. Auto-observação aumentada (com relação ao corpo) 2. Preocupação com a saúde 3. Queixas frequentes, pedidos de ajuda etc. 4. Ideias delirantes hipocondríacas	15. Hipocondria 0. Ausente 1. Maior percepção de alterações sensoriais em seu corpo 2. Preocupação com a saúde 3. Queixas frequentes, pedidos de ajuda etc. 4. Ideias delirantes hipocondríacas
16. Loss of weight A. When Rating by History: 0. No weight loss 1. Probable weight loss associated with present illness 2. Definite (according to patient) weight loss B. On Weekly Ratings by Ward Psychiatrist, When Actual Changes are Measured: 0. Less than 1 lb. weight loss in week 1. Greater than 1 lb. weight loss in week 2. Greater than 2 lb. weight loss in week	16. Perda de peso (marcar A ou B) A. Quando avaliada pela história clínica 0. Sem perda de peso 1. Provável perda de peso associada à moléstia atual 2. Perda de peso definida (de acordo com o paciente) 3. Não avaliada B. Avaliada semanalmente pelo psiquiatra responsável, quando são medidas alterações reais de peso 0. Menos de 0,5 kg de perda por semana 1. Mais de 0,5 kg de perda por semana 2. Mais de 1 kg de perda por semana 3. Não avaliada	16. Perda de peso (marcar A ou B) A. Quando avaliada pela história clínica 0. Sem perda de peso 1. Provável perda de peso associada ao transtorno atual 2. Perda de peso definida (de acordo com o paciente) 3. Não avaliada B. Avaliada semanalmente pelo psiquiatra responsável, quando são medidas alterações reais de peso 0. Menos de 0,5 kg de perda por semana 1. Mais de 0,5 kg de perda por semana 2. Mais de 1 kg de perda por semana 3. Não avaliada

Quanto à análise de conteúdo dos itens, apesar de não haver concordância significativa dos peritos para que as sugestões fossem inseridas, diferenças em relação ao vocabulário foram apontadas. Por isso, identificou-se a necessidade de realizar alterações na redação de sete questões da versão da escala usualmente utilizada no Brasil. Logo, a análise de conteúdo evidenciou que a verificação do vocabulário de acordo com o contexto é uma etapa importante da adap-

tação transcultural³⁰. As sugestões dadas pelos especialistas entrevistados foram pertinentes, no sentido de tornar a linguagem das perguntas e alternativas de respostas mais coloquiais e, conseqüentemente, facilitar o entendimento dos pacientes.

Entre as limitações do presente estudo, pode-se destacar a impossibilidade de determinar a intensidade dos sintomas depressivos, pois na pesquisa principal não foi utilizado ne-

nhum outro instrumento que verificasse a gravidade sintomática para fins de comparação. Alguns estudos encontrados na literatura^{15,17} destacam a relevância de analisar os resultados da HAM-D segundo o sexo, o que não foi feito em função da pouca variabilidade da amostra clínica. Ainda se pode apontar a falta de variabilidade na idade dos participantes, uma vez que o estudo objetivou a população de adultos jovens (18 – 27 anos).

CONCLUSÕES

A escala HAM-D é utilizada há 50 anos no Brasil, sem evidências de validade para essa população e sem estudos de adaptação transcultural; considera-se uma adaptação ingênua, dando ênfase apenas ao processo de tradução das questões originais sem um estudo semântico da linguagem e, por essa razão, os resultados da presente pesquisa colaboraram para fornecer subsídios sobre suas qualidades psicométricas da escala em uma amostra brasileira. Os resultados sugeriram que a escala apresenta qualidades promissoras, desde que algumas modificações sejam realizadas, com base nas propostas do presente estudo. Recomendam-se outras pesquisas para revisar as características psicométricas da escala e validar a nova versão proposta.

CONTRIBUIÇÕES INDIVIDUAIS

Manoela Ávila Freire – Foi responsável pela análise e interpretação dos dados, redação do artigo e aprovação da versão final.

Vera Lúcia Marques de Figueiredo – Supervisora acadêmica responsável pela elaboração do trabalho da autora principal.

Alina Gomide – Foi responsável pela análise de dados relativa à sensibilidade/especificidade do instrumento.

Karen Jansen, Ricardo Azevedo da Silva, Pedro Vieira da Silva Magalhães e Flávio Pereira Kapczinski – Participaram da elaboração do projeto, execução do estudo de campo e aprovação da versão final.

CONFLITOS DE INTERESSE

Nenhum dos autores apresentou divergência de interesses ao longo da elaboração do projeto e execução do trabalho.

REFERÊNCIAS

- Gallucci Neto J, Campos Júnior MS, Hübner CK. Escala de Depressão de Hamilton (HAM-D): revisão dos 40 anos de sua utilização. *Rev Fac Ciênc Méd.* 2001;3(1):10-4.
- Williams JBW. Standardizing the Hamilton Depression Rating Scale: past, present, and future. *Eur Arch Psychiatry Clin Neurosci.* 2001;251 Suppl. 2:11/6-12.
- Araújo RHS. Adaptação transcultural da GRID Hamilton Rating Scale for Depression (GRID-HAMD) para o português brasileiro e avaliação do impacto de um treinamento sobre a confiabilidade interavaliadores [dissertação]. Salvador (BA): Universidade Federal da Bahia, 2011.
- Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatr.* 1960;23(56):56-62.
- Moreno RA, Moreno DH. Escalas de depressão de Montgomery & Asberg (MADRS) e de Hamilton (HAM-D). *Rev Psiquiatr Clin.* 1998;25:262-72.
- Shafer AB. Meta-analysis of the Factor Structures of Four Depression Questionnaires: Beck, CES-D, Hamilton, and Zung. *J Clin Psychol.* 2006;62(1):123-46.
- Parcias S, Rosário BP, Sakae T, Monte F, Guimarães ACA, Xavier AJ. Validação da versão em português do Inventário de Depressão Maior. *J Bras Psiquiatr.* 2011;60(3):164-70.
- Bech P. Rating scales in depression: limitations and pitfalls. *Dialog Clin Neuroscience.* 2006;8(2):207-15.
- Caill H, Pires MNL. Aspectos gerais das escalas de avaliação de depressão. *Rev Psiquiatr Clin.* 1998;25:240-4.
- Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? *Am J Psychiatry.* 2004;161:2163-77.
- Furukawa TA, Akechi T, Azuma H, Okuyama T, Higuchi T. Evidence-based guidelines for interpretation of the Hamilton Rating Scale for Depression. *J Clin Psychopharmacol.* 2007;27(5):531-3.
- Hamilton M. Development of a rating scale for primary depressive illness. *British J Social Clin Psychol.* 1967;6(4):278-96.
- O'Brien KP, Glaudin V. Factorial structure and factor reliability of the Hamilton Rating Scale for Depression. *Acta Psychiatr Scand.* 1988;78:113-20.
- Pancheri P, Picardi A, Pasquini M, Gaetano P, Biondi M. Psychopathological dimensions of depression: a factor study of the 17-item Hamilton depression rating scale in unipolar depressed outpatients. *J Affect Disord.* 2002;68:41-7.
- Ramos-Brieva JA, Cordero-Villafañila A. A new validation of the Hamilton Rating Scale for Depression. *J Psychiatr Res.* 1988;22(1):21-8.
- Gibbons RD, Clark DC, Kupfer DJ. Exactly what does the Hamilton Depression Rating Scale measure? *J Psychiatr Res.* 1993;27(3):259-73.
- Akdemir A, Türkçapar MH, Orsel SD, Demiregi N, Dag I, Ozbay MH. Reliability and validity of the Turkish version of the Hamilton Depression Rating Scale. *Compr Psychiatry.* 2001;42(2):161-5.
- Furukawa TA. Assessment of mood: guides for clinicians. *J Psychos Research.* 2010;68:581-9.
- Marques AP, Santos AMB, Assumpção A, Matsutani LA, Lage LV, Pereira CAB. Validação da Versão Brasileira do Fibromyalgia Impact Questionnaire (FIQ). *Rev Bras Reumatol.* 2006;46(1):24-31.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46(12):1417-32.
- Jansen K. Transtornos de humor: prejuízos na qualidade de vida e no funcionamento de jovens da cidade de Pelotas-RS. Universidade Católica de Pelotas. In press 2010.
- Osório FL, Crippa JAS, Loureiro SR. Aspectos cognitivos do falar em público: validação de uma escala de autoavaliação para universitários brasileiros. *Rev Psiq Clín.* 2012;39(2):48-53.
- Del-Ben CM, Rodrigues CR et al., 1996. In: Jansen K. Transtornos de humor: prejuízos na qualidade de vida e no funcionamento de jovens da cidade de Pelotas-RS. Universidade Católica de Pelotas. In press 2010.
- ABEP – Associação Brasileira de Empresas de Pesquisa. Disponível em: <http://www.ibge.gov.br>. Dados com base no Levantamento Socioeconômico (IBOPE), 2003.
- Pasquali L. O uso da análise fatorial: algumas diretrizes para pesquisadores. In: Pasquali L. Análise fatorial para pesquisadores. Brasília: LabPam. 2012:141-60.
- Pasquali L. Parâmetros psicométricos dos testes psicológicos. In: Pasquali L. Técnicas do exame psicológico – TEP. São Paulo: Casa do Psicólogo. 2001.

27. Pasquali L. Testes referentes a construto: teoria e modelo de construção. In: Pasquali L. Instrumentação psicológica – fundamentos e prática. Porto Alegre: Artmed. 2010:165-98.
28. Flechter RH, Flechter SW, Wagner EH. Diagnóstico. In: Flechter RH, Flechter SW, Wagner EH. Epidemiologia clínica: elementos essenciais. Porto Alegre: Artmed. 1996.
29. Braga ACS. Curvas ROC: aspectos funcionais e aplicações [dissertação]. Universidade do Minho, 2010.
30. Jorge MR. Adaptação transcultural de instrumentos de pesquisa em saúde mental. Rev Psiq Clin. 1998;25(5):233-9.