

Universidade Federal do Rio Grande do Sul

Faculdade de Medicina

Curso de Pós-Graduação em Medicina

Dissertação de Mestrado

**Desempenho de uma escala de rastreamento de
depressão (CES-D) em usuários de um serviço de
cuidados primários de saúde de Porto Alegre**

Gustavo Schestatsky

Orientador: Prof. Dr. Marcelo Pio de Almeida Fleck

Porto Alegre, 2002

Universidade Federal do Rio Grande do Sul

Faculdade de Medicina

Curso de Pós-Graduação em Medicina

Dissertação de Mestrado

Desempenho de uma escala de rastreamento de depressão (CES-D) em usuários de um serviço de cuidados primários de saúde de Porto Alegre

Gustavo Schestatsky

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Médicas: Psiquiatria, UFRGS, como requisito parcial para obtenção do título de Mestre em Ciências Médicas - Psiquiatria

Orientador: Prof. Dr. Marcelo Pio de Almeida Fleck

Porto Alegre, 2002

AGRADECIMENTOS

Ao meu orientador Marcelo Pio de Almeida Fleck por seu estímulo permanente, participação efetiva em todas os momentos necessários, e exemplo como pesquisador e profissional.

Aos colegas e amigos mestrandos participantes das reuniões de sexta-feira: Ana Flávia Silva de Lima, Neusa Sica da Rocha e Rogério Zimpel, por suas valiosas sugestões e companheirismo.

À equipe participante da coleta dos dados: Michel Simon, Cristina Haag e Vivian Borges. Em especial ao meu colega e amigo Sérgio Louzada pelo gerenciamento da coleta e sua incansável disponibilidade para ajudar.

Aos funcionários do Centro de Estudos Luis Guedes pelo apoio logístico.

À Vania Hirakata pelo auxílio nas análises estatísticas dos dados.

Aos meus pais Sidnei e Maria Elisa por seus exemplos éticos, profissionais, da busca do aperfeiçoamento, e pelo apoio incondicional.

À Daniella Turkienicz por sua fundamental importância em minha vida, por seu amor e companheirismo.

SUMÁRIO

1. INTRODUÇÃO	5
2. REVISÃO DA LITERATURA	10
3. REFERÊNCIAS BIBLIOGRÁFICAS	22
4. JUSTIFICATIVA	32
5. OBJETIVOS	34
6. ARTIGO EM PORTUGUÊS (1)	35
7. ARTIGO EM INGLÊS (1)	54
8. ARTIGO EM PORTUGUÊS (2)	75
9. ARTIGO EM INGLÊS (2)	89
10. ANEXO	105

INTRODUÇÃO

Os transtornos depressivos são responsáveis por um enorme impacto em termos sociais, econômicos e de saúde pública em todo o mundo. A depressão maior é o transtorno mental mais prevalente (Kessler e cols, 1994), com uma prevalência-ponto de 3 a 5% na população geral e uma prevalência ao longo da vida entre 4 e 18% (Blazer e cols, 1994; Witchenn e cols, 1994). No Brasil, dados obtidos a partir de centros urbanos apontam para uma prevalência global de depressão de 2 a 10% (Almeida-Filho e cols, 1997).

O curso da depressão maior, na maioria dos casos, tende a ser crônico e recorrente (Kupfer, 1991; Keller, 1994). Ela provoca sofrimento e prejuízos funcionais nos indivíduos acometidos comparável ou superior à maioria das doenças clínicas (Wells e cols, 1989; Hays e cols, 1995), sendo responsável por uma incapacitação social maior do que qualquer outro transtorno mental e do que a maioria de doenças médicas (Murray e Lopez, 1997a). Existe, ainda, uma projeção de que a depressão venha a ser a segunda maior causa de incapacitação no mundo e a principal causa de incapacitação em países subdesenvolvidos por volta do ano 2020 (Murray e Lopez, 1997b).

Desse modo, o custo pessoal e social dos transtornos depressivos é muito alto. Estimou-se que o custo anual da depressão nos Estados Unidos ultrapassava 44 bilhões de dólares no início da década de 1990 (Greenberg e cols, 1993). Indivíduos deprimidos apresentam uma alta utilização de serviços de cuidados primários (Johnson e cols, 1992), onde a prevalência de transtornos depressivos pode chegar a 23% (Coyne e cols, 1994).

Assim, esses indivíduos geram custos de saúde pública maior do que outros pacientes (Simon e cols, 1995), o que contribui para o custo total da depressão.

Apesar de tamanho impacto social, a maioria dos episódios depressivos não são diagnosticados, portanto não são tratados (Shapiro e cols, 1984; Pérez-Stable e cols, 1990). Ainda que um grande número de indivíduos deprimidos procurem regularmente serviços de cuidados primários devido a queixas físicas, estas muitas vezes não são vistas como indícios de depressão. A falha em estabelecer o diagnóstico acarreta, além do sofrimento direto aos indivíduos, que estes sejam submetidos a investigações e tratamentos muitas vezes desnecessários, com base nas freqüentes queixas somáticas.

Esse cenário torna os serviços de cuidados primários locais propícios para a detecção precoce dos transtornos depressivos (McQuaid e cols, 1999). Não obstante, o que se observa é uma dificuldade de se realizar o diagnóstico de transtornos depressivos nesses serviços, onde somente um terço dos indivíduos deprimidos costumam ser reconhecidos como tais (Coyne e cols, 1995). Essa falha diagnóstica é mais comum nos quadros mais leves, as chamadas depressões menores ou subsindrômicas, que são justamente os mais freqüentes na clínica (Blazer, 2000). Indivíduos com depressão subsindrômica também utilizam mais freqüentemente serviços de saúde, apresentam grau elevado de incapacitação (Broadhead e cols, 1990) e têm um risco aumentado de desenvolver depressão maior (Rapaport e cols, 2002).

A conseqüência da não detecção de depressão pode ser desastrosa, pois o diagnóstico e o tratamento precoces podem contribuir para a redução da gravidade do

episódio e a melhora do curso da doença, com menor probabilidade de futuros episódios (McQuaid e cols, 1999). Há indícios de que o tratamento da depressão em pacientes com alta utilização de serviços de saúde pode resultar em aumento da sua capacidade laborativa, melhora da qualidade de vida e, possivelmente, redução de custos com despesas médicas (Katzelnick, 1997).

Esse quadro justifica fortemente, como medida prioritária de saúde pública, o emprego de ações que propiciem uma melhora da detecção de depressão em serviços de cuidados primários (Santiago, 1993). Uma maneira de atingir essa meta é o aprimoramento da capacidade dos clínicos da rede básica de saúde em abordar mais efetivamente as questões diagnósticas e terapêuticas dos transtornos depressivos, através de programas de treinamento e intercâmbio com profissionais da área da saúde mental (Hodges e cols, 2001). Esse método, porém, ainda que potencialmente efetivo, é bastante complexo e exige empreendimentos coordenados de longa duração (Van Der Feltz-Cornelis e cols, 1997). Além disso, o custo-benefício desses programas ainda não se encontra muito bem documentado (Simon e cols, 2001).

Outra maneira, mais específica e direta, de aumentar a detecção de depressão em cuidados primários, é a utilização de instrumentos de rastreamento (Spitzer e cols, 1995). Estes são particularmente úteis quando aplicados em locais onde a prevalência da doença seja alta e exista uma maior possibilidade de tratamento (Somervell e cols, 1993), como é o caso dos serviços de cuidados primários de saúde. Ainda mais quando o instrumento é simples, de fácil aplicação e com propriedades psicométricas adequadas para esse fim. Assim, o uso desse tipo de instrumento pode auxiliar o clínico a melhorar sua capacidade

de reconhecimento de casos de depressão, podendo inclusive tratá-los quando se considerar capacitado, ou então encaminhá-los para uma avaliação mais detalhada e um tratamento apropriado por um especialista.

Entretanto, a relação entre escalas de sintomas e diagnósticos psiquiátricos não é muito clara (Myers e Weissman, 1980; Roberts e cols, 1989), sendo ainda considerada incerta a real utilidade dos instrumentos de rastreamento de depressão (Whooley e cols, 1997; Williams e cols, 1999).

Um dos instrumentos mais utilizados para rastreamento populacional de depressão é a escala *Center for Epidemiologic Studies – Depression (CES-D)* (Radloff, 1977). Por ser auto-aplicada e de fácil compreensão, apresenta bom potencial para uso em serviços de cuidados primários. Mesmo assim, não há muitos estudos que avaliaram o seu desempenho específico nesse contexto, sendo que a ênfase maior tem sido dada a populações de idosos (Callahan e cols, 1994). Existem poucos estudos no Brasil sobre a CES-D, não sendo nenhum deles em populações de serviços de cuidados primários (Silveira e Jorge, 1998; Salle, 1999).

Algumas das propriedades psicométricas da CES-D, como seu ponto de corte, variam consideravelmente de acordo com o contexto onde ela é utilizada, estando sujeitas a influências culturais. Desse modo, torna-se importante conhecer o desempenho da CES-D em nosso meio, particularmente em uma população de usuários de serviços de cuidados primários. Além da possibilidade do uso da CES-D para detectar casos de depressão em

cuidados primários, a definição de um ponto de corte adequado ao nosso meio poderá servir como parâmetro para futuros estudo locais.

REVISÃO DA LITERATURA

A escala *Center for Epidemiologic Studies – Depression (CES-D)* (Radloff, 1977) (Anexo) foi desenvolvida para o uso em estudos epidemiológicos de sintomas depressivos na população geral. Destina-se a medir níveis atuais de sintomas depressivos, com ênfase no seu componente afetivo, o humor deprimido. Pode não somente identificar sintomas depressivos como também a sua gravidade, indicada pelo número e pela frequência de sintomas (Radloff e Locke, 1986).

Os itens da CES-D foram selecionados a partir de escalas de depressão anteriormente validadas (Beck e cols, 1961; Zung, 1965, Raskin e cols, 1969). Os componentes principais da sintomatologia depressiva foram retirados a partir da literatura e de estudos de análise fatorial. Eles incluem o humor deprimido, sentimentos de culpa e desvalia, sentimentos de desamparo e desesperança, retardo psicomotor, perda de apetite e distúrbios do sono.

Por ter sido desenvolvida para o uso na população geral, a CES-D é um instrumento simples e auto-aplicado, que apresenta boa aceitabilidade e facilidade de compreensão. Também pode ser administrada por entrevistadores leigos, levando em torno de cinco minutos para ser respondida. As instruções de preenchimento enfatizam o estado atual do indivíduo, questionando sobre a frequência de sintomas nos últimos sete dias. Cada resposta é pontuada de acordo com uma ordem de frequência de sintomas: raramente ou nunca (score 0); poucas vezes (score 1); às vezes (score 2); quase sempre ou sempre

(escore 3). É constituída de 20 itens, podendo o escore total variar de 0 a 60 pontos. Quatro itens são apresentados com um sentido positivo, no intuito de bloquear a tendência a respostas repetitivas, bem como de avaliar a presença (ou ausência) de afeto positivo. Estes itens são pontuados de modo reverso, ou seja, quanto menor a freqüência dos sintomas, maior a pontuação. Escores mais altos indicam maior quantidade de sintomas. A CES-D não pretende ser um instrumento diagnóstico, não sendo possível a interpretação de escores individuais.

O ponto de corte mais utilizado é 16, ou seja, pacientes com este escore ou mais são considerados com risco para depressão ou para necessidade de tratamento. Este valor originou-se do primeiro estudo de campo com a CES-D, onde correspondia ao percentil 80 (Comstock e Helsing, 1976). Desde então este tem sido o ponto mais utilizado na literatura.

Desde o seu desenvolvimento, no início da década de 1970, a CES-D tem sido amplamente estudada e utilizada em uma variedade de contextos. Há estudos realizados em populações não-clínicas, como o estudo original de Radloff (1977) com amostras comunitárias, além de estudos com escolares (Roberts e cols, 1991; Iwata e cols, 1994; Salle, 1999) e universitários (Silveira e Jorge, 1998; Ghubash e cols, 2000). Também há vários estudos em populações clínicas, com indivíduos que procuram atendimento de saúde geral ou psiquiátrico. Como exemplos estão estudos com freqüentadores de serviços de cuidados primários (Fechner-Bates e cols, 1994; Lyness e cols, 1997; Irwin e cols, 1999), bem como pacientes deprimidos (Shrout e Yager, 1989; Fountoulakis e cols, 2001), com câncer (Hann e cols, 1999; Schroevers MJ e cols, 2000), dor crônica (Magni e cols, 1994; Turk e Okifuji, 1994), artrite reumatóide (Sheehan e cols, 1995), entre outros.

Alguns estudos restringiram-se a determinadas faixas etárias, como adolescentes (Iwata e cols, 1994; Prescott e cols, 1998; Salle, 1999), adultos jovens (Silveira e Jorge, 1998, Ghubash, 2000) e idosos (Beekman e cols, 1997; Lyness e cols, 1997; Irwin e cols, 1999). Em estudos de sintomas depressivos em populações geriátricas a CES-D é freqüentemente o instrumento de escolha (Callahan e cols, 1994).

A CES-D também tem sido utilizada e estudada em diversos países como Holanda (Beekman e cols, 1997; Schroevers e cols, 2000), Suécia (Gatz e cols, 1993), Brasil (Silveira e Jorge, 1998; Salle, 1999), Japão (Iwata e cols, 1994), Arábia Saudita (Ghubash e cols, 2000), Grécia (Fountoulakis e cols, 2001), Coréia (Cho e Kim, 1998; Cho e cols, 1998). Essa grande diversidade tem possibilitado interessantes comparações acerca das influências culturais na manifestação de síndromes depressivas.

Por último, há um número considerável de estudos em populações bastante específicas, como minorias étnicas (Kuo, 1984; Roberts e Vernon, 1989; Foley e cols, 2002), comunidades indígenas (Somervell e cols, 1993) e moradores de rua (Wong, 2000).

Se, por um lado, a multiplicidade de estudos contribuiu para consolidar a importância e a validade da CES-D como instrumento de rastreamento de depressão, por outro dificulta a comparação de seus resultados em diferentes contextos. Um exemplo é a questão do ponto de corte. Estudos em diferentes populações e contextos revelaram pontos de corte bastante variáveis. Foram descritos pontos ideais que variam desde 12 em meninos

adolescentes (Garrison e cols, 1991) até 28 em índios norte-americanos (Somervell e cols, 1993).

Apesar de algumas dificuldades de generalização, de um modo geral a CES-D tem demonstrado propriedades psicométricas satisfatórias. Seu desempenho tem sido avaliado em termos de sua confiabilidade e validade como instrumento de rastreamento de sintomas depressivos.

Confiabilidade

Os parâmetros mais utilizados são o de teste-reteste e consistência interna. A confiabilidade teste-reteste da CES-D é moderada para períodos de duas a oito semanas, com coeficientes que variam entre 0,5 e 0,7 (Hann e cols, 1999; Ghubash e cols, 2000; Fountoulakis e cols, 2001).

A consistência interna tende a ser alta e é o parâmetro que se mantém mais uniforme em diferentes populações. Geralmente situa-se em torno de 0,85 em populações não clínicas (Silveira e Jorge, 1998; Ghubash e cols, 2000; Foley e cols, 2002) e 0,90 em populações clínicas (Shrout e Yager, 1989; Fountoulakis e cols, 2001).

Validade

A CES-D tem sido avaliada em termos de validade concorrente (ou discriminante) e de conteúdo. Os estudos de validade concorrente avaliam a capacidade do instrumento discriminar pacientes e não pacientes, determinando sensibilidade, especificidade, valores

preditivos e os pontos de corte mais apropriados. Nesse sentido, a CES-D possui uma boa capacidade de discriminação, apresentando média de escores consistentemente mais elevada em indivíduos deprimidos em comparação a não deprimidos (Radloff, 1977; Garrison e cols, 1991). A validade concorrente também foi demonstrada em pacientes psiquiátricos variados, como deprimidos agudos e em remissão, dependentes de álcool e drogas, e esquizofrênicos (Weissman e cols, 1977). Existe uma boa correlação com outros instrumentos, com necessidade de ajuda, e com eventos de vida recentes, sendo sensível a mudanças do humor relacionadas a eventos de vida (Radloff, 1977, Wong, 2000).

Conforme já mencionado, os pontos de corte da CES-D podem ser bastante variáveis (Tabela 1). De uma maneira geral, embora com algumas exceções, os pontos de corte ideais tendem a ser proporcionais às taxas de prevalência dos transtornos depressivos nos diferentes contextos. Assim, os pontos de corte são mais baixos em populações não clínicas, onde as prevalências de transtornos depressivos são menores (Silveira e Jorge, 1998; Salle, 1999). Já em populações clínicas, como as de amostras de cuidados primários, pacientes clínicos e psiquiátricos, onde a prevalência de transtornos depressivos é maior, o ponto de corte também é mais elevado (Lyness e cols, 1997, Cho e Kim, 1998, Fountoulakis e cols, 2001).

Por esse motivo, Weissman e colaboradores (1977) alertou para a necessidade do ajuste do ponto de corte da CES-D, dependendo da população onde for utilizada. Assim, o ponto de corte tradicional de 16 pode não ser o mais apropriado quando utilizado em contextos diferentes daquele para o qual foi desenvolvida, ou seja, na comunidade.

Os principais estudos que avaliaram a CES-D em amostras comunitárias, utilizando o ponto de corte tradicional, encontraram altas taxas de resultados falso-negativos (entre 36 e 40%) e taxas menores de falso-positivos (entre 6 e 17%) (Myers e Weissman, 1980; Roberts e Vernon, 1983). Portanto, os autores não recomendam seu uso isoladamente para rastreamento de depressão em populações comunitárias, já que um número considerável de indivíduos deprimidos não seria detectado.

Apesar dos resultados desses estudos, a maior parte dos estudos posteriores, a despeito do contexto estudado, demonstraram resultados opostos, ou seja, taxas de resultados falso-positivos bastante elevadas (Roberts e cols, 1989; Garrison e cols, 1991; Roberts e cols, 1991; Fechner-Bates e cols, 1994). Resultados positivos na CES-D freqüentemente incluem indivíduos sem um diagnóstico definido (Roberts e cols, 1989), com traços de ansiedade proeminentes (Orme e Herz, 1986) ou outros transtornos psiquiátricos como transtornos de ansiedade (Breslau, 1985).

Essa constatação levou alguns autores a questionar a validade da CES-D para o diagnóstico de depressão maior, sugerindo que a escala não mede especificamente depressão, mas sim um sofrimento geral (Fechner-Bates e cols, 1994) ou mal-estar psicológico que pode ser melhor caracterizado como “desmoralização” (Breslau, 1985; Roberts e cols, 1989). Apesar dessa limitação, admitem a utilização da CES-D em uma primeira etapa, desde que haja correção diagnóstica num segundo momento, a fim de melhor discriminar os verdadeiros dos falso-positivos (Roberts e cols, 1991).

Com relação à validade de conteúdo, supõe-se que esteja garantida a partir do momento em que a CES-D se originou de escalas de depressão validadas (Radloff, 1977). Além disso, vários estudos de análise fatorial vieram contribuir para a validade de conteúdo e de construção da CES-D, ao extrair os domínios principais que compõem o instrumento e permitir comparações da sua estrutura em diferentes populações.

O estudo original de Radloff (1977) foi o primeiro a extrair, por meio do método de análise de componentes principais, uma solução de quatro fatores, os quais denominou: (1) afeto depressivo; (2) afeto positivo; (3) somático e retardo motor; (4) interpessoal. Estudos fatoriais subsequentes, que também utilizaram a análise de componentes principais, confirmaram esse achado (Roberts, 1980; Kuo, 1984; Shrout e Yager, 1989; Gatz, 1993). Estudos de análise fatorial confirmatória replicaram uma solução de quatro fatores, demonstrando, ainda, que estes se correlacionam fortemente entre si (Hertzog e cols, 1990; Sheehan e cols, 1995; Wong, 2000).

Outros estudos, no entanto, chegaram a soluções com três fatores, geralmente combinando itens dos fatores “afeto depressivo” e “somático e retardo motor” de Radloff (1977) em um único fator (Ying, 1988; Guarnaccia e cols, 1989; Ghubash e cols, 2000; Fountoulakis e cols, 2001). Alguns estudos, ainda que mantivessem uma estrutura com quatro fatores, também agruparam os sintomas depressivos e somáticos em um fator único (Kuo, 1984; Foley e cols, 2002). Um resumo dos principais estudos fatoriais da CES-D disponíveis na literatura encontra-se na Tabela 2.

Chama atenção que as populações dos estudos fatoriais que combinam sintomas depressivos e somáticos geralmente provêm de minorias étnicas, especialmente de cultura oriental, ou de países distintos como Arábia Saudita e Grécia. Ou seja, trata-se de populações com características socioeconômicas e/ou culturais distintas das populações caucasianas como a norte-americana e européia. Acredita-se que esse achado possa ter explicação do ponto de vista cultural. Populações não-ocidentais e/ou com menor desenvolvimento socioeconômico tendem a expressar seu sofrimento psicológico somaticamente (Kleinman, 1982; Kuo, 1984; Parker e cols, 2001). Por outro lado, acredita-se que indivíduos de culturas com maior grau de desenvolvimento, como a dos Estados Unidos e Europa Ocidental, tenham adquirido maior habilidade cognitiva para verbalizar estados afetivos, dessa forma utilizando-se menos da somatização (Katon e cols, 1982).

CES-D e Cuidados Primários

Apesar do grande número de estudos de validação da CES-D existentes, incluindo populações de cuidados primários, e da sua potencial utilidade neste contexto, a questão da sua validade para essas populações ainda não se encontra bem estabelecida.

Lyness e colaboradores (1997), estudando uma população de idosos em um serviço de cuidados primários, considerou a CES-D como tendo ótimas propriedades para o rastreamento de depressão em cuidados primários, com sensibilidade de 92% e especificidade de 87%. Irwin e colaboradores (1999), aplicando uma versão reduzida da CES-D, também encontrou bons índices de sensibilidade e especificidade em idosos (100%

e 92%, respectivamente), considerando o instrumento com excelentes propriedades para o rastreamento de depressão em idosos no contexto de cuidados primários.

Por outro lado, um estudo que avaliou uma população em uma faixa etária mais ampla observou que a CES-D não classificava corretamente boa parte dos indivíduos com depressão, apresentando especificidade e valor preditivo positivo muito baixos. Os autores concluíram que o instrumento não poderia ser usado rotineiramente como teste de rastreamento de depressão em populações de cuidados primários (Klinkman e cols, 1997). Fechner-Bates e colaboradores (1994), utilizando o ponto de corte tradicional, encontrou sensibilidade de 80%, especificidade de 71%, e valor preditivo positivo de 28%. O alto índice de classificação incorreta levou o autor a sugerir que a CES-D deveria ser considerada como uma medida de sofrimento geral, não especificamente de depressão.

Especificamente em relação ao ponto de corte da CES-D, observam-se pontos mais elevados em cuidados primários em relação ao ponto de corte original, variando desde 21 (Lyness e cols, 1997) até 27 (Schulberg e cols, 1985; Zich e cols, 1990). Diferenças culturais e metodológicas, como o uso de diferentes testes padrão-ouro, podem ser responsáveis por algumas dessas diferenças.

Existem poucos estudos no Brasil sobre a CES-D, não sendo nenhum deles em populações de usuários de serviços de cuidados primários. Eles também demonstram uma considerável variabilidade no seu desempenho, gerando diferentes pontos de corte conforme as diferentes características das populações estudadas (Silveira e Jorge, 1998; Salle, 1999). Um estudo realizado em Porto Alegre, utilizando a CES-D para detecção de

depressão em usuários de um serviço de cuidados primários, sugeriu que o ponto de corte de 16 era provavelmente muito baixo para essa população, gerando muitos resultados falso-positivos (Fleck e cols, 2002).

Dada a alta prevalência da depressão em cuidados primários, os altos índices de falha em diagnosticá-la, e suas potencialmente graves conseqüências; e levando-se em conta que fatores culturais estão implicados na variabilidade do desempenho de instrumentos diagnósticos em psiquiatria (Rubio-Stípec e cols, 2000), torna-se importante estudar o desempenho da CES-D em nosso meio. Dito de outro modo, para que um instrumento de rastreamento de depressão como a CES-D possa ser útil, é necessário que se estabeleça como se dá seu desempenho psicométrico na população de interesse, para que seus resultados sejam válidos.

Tabela 1. Estudos de validação concorrente da CES-D em diferentes populações

Estudo	Amostra (país)	Pontos de corte	Sens^a (%)	Espec^a (%)	ASC^c
Populações não clínicas					
Garrison e cols (1991)	2645 adolescentes (EUA)	12 (meninos)	85	49	0.61
		22 (meninas)	83	77	0.77
Roberts e cols (1991)	1704 adolescentes (EUA)	24	84	75	0.87
		22 (meninos)			
		24 (meninas)	0.83		
Somervell e cols (1993)	120 índios (EUA)	28	100	97	
Silveira e Jorge (1998)	50 adultos jovens (Brasil)	15	100	75	
Salle (1999)	503 adolescentes (Brasil)	14	73	76	
Ghubash e cols (2000)	350 mulheres jovens (Arábia Saudita)	21	82	83	0.84
Populações clínicas					
Lyness e cols (1997)	130 pacientes idosos de cuidados primários (EUA)	21	92	87	0.95
Cho e Kim (1998)	164 pacientes psiquiátricos (Coreia)	25	91	79	0.92
Silveira e Jorge (1998)	50 farmacodependentes (Brasil)	24	88	59	
Fountoulakis e cols (2001)	40 pacientes deprimidos e 120 controles normais (Grécia)	23	90	91	

a: sensibilidade
b: especificidade
c: área sob a curva

Tabela 2. Estudos fatoriais e de consistência interna da CES-D

Autor	Amostra	Número de fatores mantidos	Sintomas depressivos e somáticos diferenciados	Consistência interna
Radloff (1977)	5144 indivíduos da comunidade	4	Sim	0.85
	70 pacientes psiquiátricos (EUA)			0.90
Roberts (1980)	705 imigrantes latinos (EUA)	4	Sim	0.87
Kuo (1984)	499 imigrantes asiáticos (EUA)	4	Não	---
Shrout e Yager (1989)	65 pacientes deprimidos (EUA)	4	Sim	0.91
	228 indivíduos da comunidade	4	Sim	0.84
Gatz (1993)	1663 idosos (Suécia)	4	Sim	---
Iwata e cols (1994)	1500 adolescentes escolares (Japão)	5	Sim	0.81
Silveira e Jorge (1998)	50 universitários e 50 farmacodependentes (Brasil)	4	Sim	0.85
Salle (1999)	503 adolescentes escolares (Brasil)	3	Sim	---
Ghubash e cols (2000)	350 mulheres jovens (Arábia Saudita)	3	Não	0.88
Fountoulakis e cols (2001)	40 pacientes deprimidos e 120 controles normais (Grécia)	3	Não	0.95
Foley e cols (2002)	227 negros idosos (EUA)	4	Não	0,86

Nota: com exceção de Iwata e cols (1994), que utilizou rotação Oblimin, os demais estudos listados utilizaram o método de análise de componentes principais com rotação Varimax.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Almeida-Filho N, Mari JJ, Coutinho E, e cols. Brazilian multicentric study of psychiatric morbidity: methodological features and prevalence estimates. *Br J Psychiatry* 171: 524-529, 1997.
2. Beck AT, Ward CH, Mendelson, e cols. An inventory for measuring depression. *Arch Gen Psychiatry* 4: 53-63, 1961.
3. Beekman ATF, Deeg DJH, Van Limbeek J, e cols. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med* 27 (1): 231-235, 1997.
4. Blazer DG, Kessler RC, McGonagle KA, Swartz MS e cols. The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. *Am J Psychiatry* 151 (7): 979-986, 1994.
5. Blazer DG. Mood disorders: epidemiology. In: Kaplan & Sadock's *Comprehensive Textbook of Psychiatry*. 7 ed. Lippincott Williams & Wilkins. Philadelphia, 2000. pp 1298-1307.
6. Boyd JH, Weissman MM, Thompson WD, e cols. Screening for depression in a community sample: understanding the discrepancies between depression symptom and diagnostic scales. *Arch Gen Psychiatry* 39:1195-1200, 1982.
7. Breslau N. Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Res* 15: 219-229, 1985.

8. Broadhead WE, Blazer DG, George LK, Tse CK. Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA* 264 (19): 2524- 2528.
9. Callahan CM, Hui SL, Nienaber NA, e cols. Longitudinal study of depression and health services use among elderly primary care patients. *JAGS* 42:833-838, 1994.
10. Cho MJ, Nam JJ, Suh GH. Prevalence of symptoms of depression in a nationwide sample of Korean adults. *Psychiatry Res* 81: 341-352, 1998.
11. Cho MJ, Kim AKH. Use of the Center for Epidemiologic Studies Depression (CES-D) Scale in Korea. *J Nerv Ment Dis* 186: 304-310, 1998.
12. Comstock GW, Helsing KJ. Symptoms of depression in two communities. *Psychol Med* 6: 551-563, 1976.
13. Coyne JC, Fechner-Bates S, Schwenk TL. Prevalence, nature and comorbidity of depressive disorders in primary care. *Gen Hosp Psychiatry* 16: 267-276, 1994.
14. Coyne JC, Schwenk TL, Fechner-Bates S. Nondetection of depression by primary care physicians reconsidered. *Gen Hosp Psychiatry* 17: 3-12, 1995.
15. Fechner-Bates S, Coyne JC, Schwenk TL. The relationship of self-reported distress to depressive disorders and other psychopathology. *J Consult Clin Psychol* 62: 550-559, 1994.
16. Fleck MPA, Lima AF, Louzada S, e cols. Associação entre sintomas depressivos e funcionamento social em cuidados primários de saúde. *Rev Saúde Pública* 36 (4): 431-438, 2002.
17. Fletcher RH, Fletcher SW, Wagner EH. *Epidemiologia Clínica: Elementos Essenciais*. 3 ed. Porto Alegre, Artes Médicas, 1996.

18. Foley KL, Reed PS, Mutran EJ, De Vems RF. Measurement adequacy of the CES-D among a sample of older African-Americans. *Psychiatry Res* 109: 61-69, 2002.
19. Fountoulakis K, Iacovides A, Kleanthous S, e cols. Reliability, validity and psychometric properties of the Greek translation of the Center fo Epidemiological Studies – Depression (CES-D) Scale. *BMC Psychiatry* 1: 3-13, 2001.
20. Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. *J Affect Disord* 46 (1): 1-13, 1997.
21. Garrison CZ, Addy CL, Jackson KL, e cols. The CES-D as a screen for depression and other psychiatric disorders in adolescents. *J Am Acad Child Adolesc Psychiatry* 30 (4): 636-41, 1991.
22. Gatz M, Johansson B, Pedersen N, e cols. A cross-national self-report measure of depressive symptomatology. *International Psychogeriatrics* 5 (2): 147-156, 1993.
23. Ghubash R, Daradkeh TK, Al Naseri KS, e cols. The performance of the Center for Epidemiologic Study Depression Scale (CES-D) in an Arab female community. *International Journal of Social Psychiatry* 46 (4): 241-249, 2000.
24. Greenberg PE, Stiglin LE, Finkelstein SN, Berndt ER. The economic burden of depression in 1990. *J Clin Psychiatry* 54 (11): 405-418, 1993.
25. Guarnaccia PJ, Angel R, Worobey JL. The factor structure of the CES-D in the Hispanic health and nutrition examination survey: the influences of ethnicity, gender and language. *Soc Sci Medicine* 29: 85-94, 1989.
26. Hann D, Winter K, Jacobsen P. Measurement of depressive symptoms in cancer patients: evaluation of the Center for Epidemiological Studies Depression Scale (CES-D). *Journal of Psychosomatic Research* 46 (5): 437-443, 1999.

27. Hays RD, Wells KB, Sherbourne CD, et al. Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Arch Gen Psychiatry* 52: 11-19, 1995.
28. Hertzog C, Alstine JV, Usala PD, et al. Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychological Assessment* 2 (1): 64-72, 1990.
29. Hodges B, Inch C, Silver I. Improving the psychiatric knowledge, skills, and attitudes of primary care physicians, 1950-2000: a review. *Am J Psychiatry* 158: 1579-1586, 2001.
30. Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). *Arch Intern Med* 159:1701-1704, 1999.
31. Iwata N, Saito K, Roberts RE. Responses to a self-administered depression scale among younger adolescents in Japan. *Psychiatry Res* 53: 275-287, 1994.
32. Johnson J, Weissman MM, Klerman GL. Service utilization and social morbidity associated with depressive symptoms in the community. *JAMA* 267 (11): 1478-1483, 1992.
33. Katon W, Kleinman A, Rosen G. Depression and somatization: a review. *The American Journal of Medicine* 72: 127-135, 1982.
34. Katzelnick. Effect of primary care treatment of depression on service use by patients with high medical expenditures. *Psychiatric Services* 48 (1): 59-64, 1997.
35. Keller MB. Depression: a long-term illness. *Br J Psychiatry* 165 (suppl 26): 9-15, 1994.

36. Kessler RC, McGonagle, Zhao S, e cols. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Study. *Arch Gen Psychiatry* 51: 8-19, 1994.
37. Kleinman A. Neurasthenia and depression: a study of somatization and culture in China *Culture, Medicine and Psychiatry* 6: 117-190, 1982.
38. Klinkman MS, Coyne JC, Gallo S, Schwenk TL. Can case-finding instruments be used to improve physician detection of depression in primary care? *Arch Fam Med* 6 (6): 567-73, 1997.
39. Kuo WH. Prevalence of depression among Asian-Americans. *J Nerv Ment Dis* 172 (8): 449-457, 1984.
40. Kupfer DJ. Long-term treatment of depression. *J Clin Psychiatry* 52 (suppl. 5): 28-34, 1991.
41. Lyness JM, Noel TK, Cox C, e cols. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Dpression Scale. *Arch Intern Med* 157 (4): 449-454, 1997.
42. Magni G, Moreschi C, Rigatti-Luchini S, Merskey H. Prospective study on the relationship between depressive symptoms and chronic musculoskeletal pain. *Pain* 56: 289-297, 1994.
43. McQuaid JR, Stein MB, Laffaye C, McCahill. Depression in a primary care clinic: the prevalence and impact of an unrecognized disorder. *J Affect Disord* 55: 1-10, 1999.
44. Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 349: 1436-1442, 1997.
45. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 349: 1498-1504, 1997.

46. Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry* 137 (9): 1081-1084, 1980.
47. Orme JG, Reis J, Herz EJ. Factorial and discriminant validity of the Center for Epidemiological Studies Depression (CES-D) Scale. *Journal of Clinical Psychology* 42 (1): 28-33, 1986.
48. Parikh RM. The sensitivity and specificity of the Center for Epidemiologic Studies Depression Scale in screening for post-stroke depression. *Int J Psychiatry Med* 18 (2):169-81, 1988.
49. Parker G, Gladstone G, Chee KT. Depression in the planet's largest ethnic group: the Chinese. *Am J Psychiatry* 158 (6): 857-864, 2001.
50. Pérez-Stable EJ, Miranda J, Muñoz RF, Ying YW. Depression in medical outpatients: underrecognition and misdiagnosis. *Arch Intern Med* 150 (5): 1083-1088, 1990.
51. Prescott CA, McArdle JJ, Hishinuma ES, e cols. Prediction of major depression and dysthymia from CES-D scores among ethnic minority adolescents. *J Am Acad Child Adolesc Psychiatry* 37 (5): 495-503, 1998.
52. Radloff LS. The CES-D Scale: a self-report depression scale for research in general population. *Applied Psychol Measurement*, 1 (3), 385-401, 1997.
53. Radloff LS, Locke BZ. The community mental health survey and CES-D scale. In: Weissman MM, Myers JK, Ross CE (eds) *Community survey of psychiatric disorders*. New Brunswick, NJ. Rutgers University Press, 1986.
54. Rapaport MH, Judd LL, Schettler PJ, e cols. A descriptive analysis of minor depression. *Am J Psychiatry* 159: 637-643, 2002.

55. Raskin A, Schulterbrandt MS, Reating N, McKeon JJ. Replication of factors of psychopathology in interview, ward behavior and self-report ratings of hospitalized depressives. *J Nerv Ment Dis* 148 (1): 87-98, 1969.
56. Roberts RE. Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Res* 2: 125-134, 1980.
57. Roberts RE, Vernon SW. The Center for Epidemiologic Studies Depression Scale: its use in a community sample. *Am J Psychiatry* 140 (1): 41-6, 1983.
58. Roberts RE, Vernon SW, Rhoades HM. Effects of language and ethnic status on reliability and validity of the Center for Epidemiologic Studies-Depression Scale with psychiatric patients. *J Nerv Ment Dis* 177 (10): 581-92, 1989.
59. Roberts RE, Lewinsohn PM, Seeley JR. Screening for adolescent depression: a comparison of depression scales. *J Am Acad Child Adolesc Psychiatry* 30 (1): 58-66, 1991.
60. Rubio-Stipec M, Hicks MH, Tsuang MT. Cultural factors influencing the selection, use, and interpretation of psychiatric measures. In: *Handbook of Psychiatric Measures*. American Psychiatric Association. Washington, 2000. pp 33-42.
61. Salle E. Estudo da sintomatologia depressiva em adolescentes de 15 a 17 anos de uma escola de 2º grau de Porto Alegre através das escalas auto-aplicáveis BDI, CRS e CES-D. Dissertação de Mestrado em Clínica Médica/UFRGS. Porto Alegre, abril de 1999.
62. Santiago JM. The costs of treating depression. *J Clin Psychiatry* 54 (11): 425-426, 1993.
63. Schroevers MJ, Sanderman R, Sonderen E, Ranchor AV. The evaluation of the Center for Epidemiologic Studies Depression (CES-D) Scale: depressed and positive affect in

- cancer patients and healthy reference subjects. *Quality of Life Research* 9:1015-1029, 2000.
64. Schulberg HC, Saul M, McClelland M, e cols. Assessing depression in primary medical and psychiatric practices. *Arch Gen Psychiatry* 42: 1164-1170, 1985.
 65. Shapiro S, Skinner EA, Kessler LG, e cols. Utilization of health and mental health services: three epidemiological catchment area sites. *Arch Gen Psychiatry* 41: 971-978, 1984.
 66. Sheehan TJ, Fifiield J, Reisine S, Tennen H. The measurement structure of the Center for Epidemiologic Studies Depression Scale. *Journal of Personality Assessment* 64 (3): 507-521, 1995.
 67. Shrout PE, Yager TJ. Reliability and validity of screening scales: effect fo reducing scale length. *J Clin Epidemiol* 42 (1): 69-78, 1989.
 68. Silveira DX, Jorge MR. Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens. *Rev Psiq Clín* 25 (5) Edição Especial: 251-261, 1998.
 69. Simon GE, VonKorff M, Barlow W. Health care costs of primary care patients with recognized depression. *Arch Gen Psychiatry* 52, 850-856, 1995.
 70. Simon GE, Katon WJ, VonKorff M, e cols. Cost-effectiveness of a collaborative care program for primary care patients with persistent depression. *Am J Psychiatry* 158 (10): 1638-1644, 2001.
 71. Somervell PD, Beals J, Kinzie JD, e cols. Criterion validity of the Center for Epidemiologic Studies Depression Scale in a population sample from an American indian village. *Psychiatry Res* 47 (3): 255-66, 1993.

72. Spitzer RL, Kroenke K, Linzer M, et al. Health-related quality of life in primary care patients with mental disorders: results from the PRIME-MD 1000 Study. *JAMA* 274 (19): 1511-1517, 1995.
73. Turk DC, Okifuji A. Detecting depression in chronic pain patients: adequacy of self-reports. *Behav Res Ther* 32 (1): 9-16, 1994.
74. Van Der Feltz-Cornelis CM, Lyons JS, Huyse FJ, et al. Health services research on mental health in primary care. *Int'l J Psychiatry in Medicine* 27 (1): 1-21, 1997.
75. Weissman MM, Sholomskas D, Pottenger M, et al. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol* 106: 203-214, 1977.
76. Wells KB, Stuart A, Hays RD, et al. The functioning and well-being of depressed patients. *JAMA* 262 (7): 914-919, 1989.
77. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *J Gen Intern Med* 12: 439-445, 1997.
78. Williams JW, Mulrow CD, Kroenke K, et al. Case-finding for depression in primary care: a randomized trial. *Am J Med* 106:36-43, 1999.
79. Wittchen H, Knaüper B, Kessler RC. Lifetime risk of depression. *Br J Psychiatry* 165 (suppl 26): 16-22, 1994.
80. Wong YI. Measurement properties of the Center for Epidemiologic Studies – Depression Scale in a homeless population. *Psychological Assessment* 12 (1): 69-76, 2000.
81. Ying YW. Depressive symptomatology among Chinese-Americans as measured by the CES-D. *Journal of Clinical Psychology* 44: 739-746, 1988.
82. Zich JM, Attkisson CC, Greenfield TK. Screening for depression in primary care clinics: the CES-D and the BDI. *Int J Psychiatry Med* 20: 259-277, 1990.

83. Zung WK. A self-rating depression scale. Arch Gen Psychiatry 12: 63-70, 1965.

JUSTIFICATIVA

A depressão maior é um transtorno mental altamente prevalente, de alto impacto em termos econômicos e de saúde pública em todo o mundo. Sua prevalência é particularmente elevada em serviços de cuidados primários de saúde, os quais se constituem em locais propícios para sua detecção. Não obstante, a depressão é muito pouco detectada nesses serviços, deste modo privando os indivíduos acometidos do seu direito ao tratamento apropriado.

O uso de instrumentos de rastreamento de depressão como a CES-D pode auxiliar o clínico a aprimorar sua capacidade diagnóstica para esse transtorno, permitindo intervenções terapêuticas apropriadas que possam modificar positivamente o curso da doença.

Algumas das propriedades psicométricas da CES-D, como seu ponto de corte, variam de acordo com o contexto onde ela é utilizada, estando sujeito a influências culturais. Deste modo, torna-se importante conhecer o desempenho da CES-D em nosso meio, particularmente em uma população de usuários de serviços de cuidados primários. Existem poucos estudos no Brasil sobre a CES-D, não sendo nenhum deles em populações de serviços de cuidados primários.

Além da possibilidade do uso da CES-D para detectar casos de depressão em cuidados primários no nosso meio, a definição de um ponto de corte adequado ao nosso meio poderá servir como parâmetro para futuros estudo locais.

OBJETIVOS

OBJETIVO GERAL

Avaliar o desempenho da CES-D como instrumento de rastreamento de depressão em uma amostra brasileira de pacientes que procuram serviços de cuidados primários de saúde, a fim de permitir seu uso adequado em nosso meio.

OBJETIVOS ESPECÍFICOS

- Avaliar a validade concorrente da CES-D em usuários de um serviço de cuidados primários de saúde de Porto Alegre, estabelecendo o ponto de corte mais apropriado para essa população, e verificando possíveis diferenças de gênero.
- Avaliar a confiabilidade da CES-D em uma população de usuários de um serviço de cuidados primários de saúde de Porto Alegre.
- Avaliar a estrutura fatorial da CES-D em uma população de usuários de um serviço de cuidados primários de saúde de Porto Alegre.

ARTIGO EM PORTUGUÊS (1)

**VALIDADE DE CRITÉRIO DA CES-D EM USUÁRIOS DE UM
SERVIÇO DE CUIDADOS PRIMÁRIOS DE SAÚDE DE PORTO
ALEGRE**

Gustavo Schestatsky¹, Marcelo P. A. Fleck²

1) Médico Psiquiatra. Mestrando do Programa de Pós-Graduação em Ciências Médicas: Psiquiatria da Universidade Federal do Rio Grande do Sul.

2) Médico Psiquiatra. Professor Adjunto do Departamento de Psiquiatria e Medicina Legal da Universidade Federal do Rio Grande do Sul.

SUMÁRIO

Objetivo: Avaliar a validade concorrente de uma versão brasileira da CES-D como instrumento de rastreamento de depressão em usuários de um serviço de cuidados primários de saúde, estabelecendo um ponto de corte adequado para essa população. **Método:** 436 indivíduos responderam ao instrumento de rastreamento populacional de depressão CES-D, seguido de uma entrevista diagnóstica através do instrumento CIDI. Procedeu-se a uma análise ROC para localização dos pontos de corte na curva. Foram calculadas sensibilidade e especificidade para os escores da CES-D. **Resultados:** O ponto de corte de melhor desempenho global da CES-D foi 24, tanto para a amostra total (sensibilidade 74%, especificidade 73%) como para o grupo das mulheres, e 22 para o grupo de homens. Do ponto de vista de rastreamento, privilegiando maior sensibilidade, os melhores pontos foram 18 para a amostra total, 11 para homens e 19 para mulheres. **Discussão:** a CES-D apresentou desempenho global satisfatório, porém com excesso de resultados falso-positivos. Houve diferença marcante do desempenho da CES-D entre os gêneros quando utilizada com a finalidade de rastreamento populacional. Pode haver necessidade de ajuste do ponto de corte de acordo com a finalidade do uso da CES-D e com o gênero.

INTRODUÇÃO

Os transtornos depressivos são responsáveis por um enorme impacto em termos sociais, econômicos e de saúde pública em todo o mundo (Murray e Lopez, 1997), sendo a depressão maior o transtorno mental mais prevalente (Kessler e cols, 1994). Indivíduos deprimidos têm uma alta utilização de serviços de saúde (Johnson e cols, 1992), particularmente os de cuidados primários, onde a prevalência de transtornos depressivos é muito elevada (Coyne e cols, 1994). Esse cenário torna os serviços de cuidados primários locais propícios para a detecção precoce dos transtornos depressivos (McQuaid e cols, 1999). Não obstante, a maioria dos episódios depressivos não é diagnosticada (Pérez-Stable e cols, 1990; Coyne e cols, 1995), deste modo privando os indivíduos acometidos do tratamento apropriado.

Uma das estratégias possíveis no sentido de diminuir tamanho impacto pode ser a melhora da detecção de casos de depressão em serviços de cuidados primários, através do uso de instrumentos de rastreamento auto-aplicados. Um dos instrumentos mais utilizados em estudos epidemiológicos para detecção de depressão é a escala *Center for Epidemiologic Studies – Depression (CES-D)* (Radloff, 1977), a qual foi desenvolvida para o estudo de sintomas depressivos na população geral. A CES-D tem sido amplamente estudada e utilizada nos mais variados contextos, em diversos países, incluindo populações de cuidados primários (Lyness e cols, 1997; Irwin e cols, 1999).

Desde os primeiros estudos com a CES-D tem sido demonstrada a sua validade e confiabilidade satisfatórias para o rastreamento de sintomas depressivos em diversas

populações (Radloff, 1977; Weissman e cols, 1977; Myers e Weissman, 1980; Roberts, 1980; Roberts e Vernon, 1983). No entanto, para o uso no contexto de cuidados primários, sua utilidade ainda é alvo de controvérsias (Fechner-Bates e cols, 1994; Klinkman e cols, 1997).

O desempenho da CES-D pode variar bastante conforme o contexto estudado, estando sujeito a influências culturais. Em particular, o ponto de corte mais apropriado para determinada população pode diferir bastante do ponto de corte originalmente proposto. De um modo geral, os pontos de corte tendem a ser mais baixos em populações não clínicas (Silveira e Jorge, 1998; Salle, 1999), e mais altos em populações clínicas, incluindo populações de cuidados primários (Lyness e cols, 1997, Cho e Kim, 1998, Fountoulakis e cols, 2001). Por esse motivo torna-se importante a definição de um ponto de corte apropriado ao contexto em que a escala está sendo estudada, a fim de usufruir o máximo de seu desempenho na população de interesse.

No Brasil existem poucos estudos sobre a CES-D, não sendo nenhum deles em populações de usuários de serviços de cuidados primários (Silveira e Jorge, 1998; Salle, 1999). Um estudo que utilizou a CES-D para detectar depressão em uma população brasileira de cuidados primários sugeriu que o ponto de corte tradicional 16 provavelmente é muito baixo para essa população, gerando muitos resultados falso-positivos (Fleck e cols, 2002).

Assim, o presente estudo objetivou avaliar a validade concorrente de uma versão brasileira da CES-D como instrumento de rastreamento de depressão em usuários de

serviços de cuidados primários de saúde, estabelecendo um ponto de corte adequado. Outro objetivo foi o de verificar possíveis diferenças de gênero no desempenho da CES-D nessa população.

MÉTODOS

Amostra

Participaram do estudo 436 indivíduos que procuraram atendimento em três postos de um serviço municipal de cuidados primários de saúde ligados a um hospital de referência, em uma área delimitada da cidade de Porto Alegre. Os motivos da procura incluíam atendimentos típicos de um serviço de cuidados primários, como consultas médicas, odontológicas e de enfermagem.

Os dados foram coletados em duas etapas, constituindo duas amostras complementares, oriundas da mesma população. A primeira amostra foi composta por indivíduos participantes do estudo LIDO (Fleck e cols, 2002), um estudo longitudinal observacional de pacientes depressivos em cuidados primários. Nesta etapa foram selecionados 300 indivíduos cujo escore na escala CES-D fosse 16 ou mais (ponto de corte tradicional). A amostra complementar (segunda etapa da coleta) foi constituída de 136 indivíduos que apresentavam escore inferior a 16 na CES-D, a fim de que obtivéssemos toda a gama de escores distribuídos nessa população (não somente aqueles com escores de 16 ou mais da primeira etapa de coleta).

Foram convidados a participar todos os sujeitos com idades entre 18 e 75 anos, pertencentes à área geográfica dos postos, que estivessem buscando atendimento nos mesmos. Não foram incluídos aqueles que não demonstravam capacidade de entendimento para completar o questionário auto-respondido, os que se encontravam em tratamento para depressão e os que tinham diagnóstico de algum transtorno psiquiátrico maior.

Coleta de dados

Os sujeitos que buscaram atendimento nos referidos postos nos períodos de coleta, e que satisfizeram os critérios de inclusão e exclusão, foram convidados de modo consecutivo a participar do estudo e assinaram um consentimento informado. Havia uma equipe de entrevistadores que cobria diversos turnos de atendimento nos referidos postos. Os sujeitos que concordaram em participar do estudo responderam ao instrumento de rastreamento auto-aplicável CES-D nas salas de espera dos postos. Com base nos escores de obtidos nesta fase de recrutamento (CES-D \geq 16 como critério de interesse na primeira etapa da coleta, e CES-D $<$ 16 como interesse na segunda etapa), os sujeitos participavam de uma segunda fase de avaliação, dentro de um período máximo de 14 dias. Nesta avaliação, que podia ser realizada no próprio posto ou no domicílio dos sujeitos, estes respondiam novamente à CES-D, e eram então submetidos à entrevista diagnóstica por entrevistadores treinados, utilizando-se o instrumento diagnóstico *Composite International Diagnostic Interview (CIDI)*.

Instrumentos diagnósticos

- **CES-D:** Foi desenvolvida para medir sintomas depressivos na população geral. Não fornece o diagnóstico de depressão, apenas identifica indivíduos com maior probabilidade de estarem deprimidos. É constituída de 20 itens considerados representativos dos sintomas principais da depressão. Os indivíduos classificam cada item utilizando uma escala de frequência (de 0=raramente ou nenhum período a 3=a maior parte ou todo o tempo) referente à última semana. O escore é uma soma simples, podendo variar de 0 a 60. O ponto de corte de 16 pontos ou acima é geralmente usado para distinguir indivíduos com maior possibilidade de apresentar depressão. É uma escala auto-aplicada, de fácil administração e compreensão. Para o presente estudo foi utilizada uma versão na língua portuguesa, a qual havia sido traduzida e adaptada por um painel bilíngüe com a supervisão de uma pesquisadora estrangeira.
- **CIDI:** Foi desenvolvida para uso em cuidados primários de saúde como instrumento diagnóstico de uma variedade de problemas de saúde mental, incluindo a depressão. No presente estudo a CIDI foi administrada por entrevistadores devidamente treinados, utilizado-se o módulo de depressão da sua versão revisada (2.1).

Análise dos dados

Foram estimadas sensibilidade e especificidade dos vários pontos de corte da CES-D, utilizado-se o diagnóstico de depressão maior fornecido pela CIDI como padrão-ouro. A variação de sensibilidade e especificidade ao longo do espectro de pontos de corte permitiu

determinar os pontos mais apropriados. Optou-se por analisar separadamente o grupo de homens e de mulheres para verificar diferenças de gênero no desempenho da escala.

A fim de obter um perfil do poder de discriminação da escala dentro de todas as possibilidades de pontos de corte, e auxiliar na determinação dos melhores pontos de corte, procedeu-se a uma análise ROC (*receiver operating characteristic*) e à construção de uma curva ROC (Murphy e cols, 1987). Os pontos de corte ideais foram identificados como os pontos da curva com a maior distância perpendicular da linha diagonal que parte da coordenada 0,0 até a coordenada 1,1 (linha preditiva do acaso). Estes foram os pontos que produziram a menor soma das taxas de falso-positivos e de falso-negativos (Garrison e cols, 1991).

Como parâmetro de desempenho global da CES-D utilizou-se a área que se encontra sob a curva ROC (ASC). A ASC é um índice global do poder de discriminação ou acurácia que a escala oferece, resumizando o desempenho do instrumento como um todo (Fletcher e cols, 1996; Silveira e cols, 2000). Quanto maior a ASC (quanto mais alta a curva), melhor é o instrumento (Garrison e cols, 1991).

Foi utilizado o programa estatístico SPSS versão 8.0.

RESULTADOS

A amostra total constituiu-se de 436 indivíduos (Tabela 1). A média de idade foi de 38,9 anos (desvio-padrão 13,8), com predomínio do sexo feminino (76,8%). A maior parte dos indivíduos eram casados (53,7%), seguida de solteiros (25,5%), separados (14,7%) e viúvos (6,2%). O nível de instrução foi, em média, de 9,5 anos de estudo (desvio-padrão 3,3). A média dos escores da CES-D foi de 22,1 pontos, embora esta não seja uma média representativa desta população (conforme explicado acima, a parte da amostra com CES-D ≥ 16 esteve super-representada).

A ASC foi de 0,80 na amostra total, demonstrando um desempenho global satisfatório da CES-D (Figura 1). De acordo com o sexo, a ASC foi de 0,77 para os homens e 0,82 para as mulheres. O formato irregular da curva ROC no grupo dos homens provavelmente reflete o tamanho de amostra mais reduzido deste subgrupo.

A Tabela 2 mostra o balanço entre sensibilidade e especificidade dos vários pontos de corte da CES-D. À medida que o ponto de corte aumenta, ocorre uma redução da sensibilidade e um aumento da especificidade.

Utilizando-se o ponto de corte 16, habitualmente preconizado para rastreamento de prováveis casos de depressão, obtivemos uma sensibilidade de 93% e uma especificidade de 46%. O melhor desempenho da CES-D, definido como o ponto de melhor equilíbrio entre sensibilidade e especificidade – e também o ponto mais alto da curva ROC –, ocorreu

quando utilizamos o ponto de corte 24, produzindo uma sensibilidade de 74% e uma especificidade de 73%.

Quando o desempenho foi analisado de acordo com o sexo, obteve-se o mesmo ponto de corte 24 para mulheres, com sensibilidade de 79% e especificidade de 72%. No grupo dos homens o melhor ponto de corte foi 22, com sensibilidade de 67% e especificidade de 73%.

Devido à baixa sensibilidade obtida no ponto de corte de melhor desempenho global, optou-se por definir um ponto de corte alternativo mais apropriado à função original da CES-D, a de rastreamento de sintomas depressivos na população geral. Nesse sentido, consideramos adequado definir um ponto que privilegiasse maior sensibilidade, mesmo que às custas de uma redução da especificidade. Desse modo, a falha em detectar indivíduos deprimidos (resultados falso-negativos) é reduzida para um nível mais aceitável, inferior a 10%.

Assim, analisando as mudanças de desempenho ao longo da faixa de pontos de corte, consideramos o ponto de corte 18 como o mais apropriado para o rastreamento de sintomas depressivos na população estudada. Neste ponto a perda de sensibilidade é mínima, mantendo-se elevada (91%), havendo um pequeno ganho de especificidade, atingindo 52%. Acima desse ponto a sensibilidade cai de modo mais acentuado a níveis inferiores a 90%, ocorrendo uma perda progressiva da capacidade de identificação de indivíduos acometidos, o que deixa de ser interessante do ponto de vista do rastreamento de depressão.

Para as mulheres, o melhor ponto de corte para rastreamento foi 20, com sensibilidade de 93% e especificidade de 55%. Para os homens, o melhor ponto foi 11, com sensibilidade de 93% e uma especificidade de 36%.

DISCUSSÃO

Em uma população de usuários de um serviço de cuidados primários de saúde, a CES-D apresentou um desempenho compatível com sua finalidade de rastreamento. Com o ponto de corte 16 demonstrou ser um instrumento bastante sensível, dificilmente deixando de identificar possíveis casos de depressão. A contrapartida, entretanto, foi a sua baixa especificidade, gerando uma elevada taxa de resultados falso-positivos.

Esse achado está de acordo com estudos que também encontraram taxas excessivamente altas de falso-positivos (Roberts e cols, 1989; Garrison e cols, 1991; Roberts e cols, 1991), inclusive no contexto de cuidados primários (Fechner-Bates e cols, 1994). Um estudo brasileiro que utilizou a CES-D para avaliar a presença de depressão em uma população de cuidados primários encontrou uma taxa de indivíduos deprimidos excessivamente elevada com o uso do ponto de corte 16, sugerindo que este ponto também estivesse gerando muitos resultados falso-positivos (Fleck e cols, 2002).

É possível que a presença de outros diagnósticos psiquiátricos não detectados, que são muito frequentes em cuidados primários, como transtornos de ansiedade, transtornos de ajustamento, distímia e transtornos somatoformes, tenha contribuído para escores elevados

da CES-D que não configuraram diagnóstico de depressão maior, desse modo gerando o excesso de resultados falso-positivos (Boyd e cols, 1982; Roberts e cols, 1989). Também é possível que, por tratar-se de uma população com nível socioeconômico menos favorecido, a ocorrência de sintomas depressivos temporários e subsindrômicos seja maior, o que também contribuiria para maiores taxas de falso-positivos (Cho e Kim, 1998).

A ASC de 0,80 indicou que a CES-D é um instrumento de acurácia satisfatória para o rastreamento de indivíduos com depressão em serviços de cuidados primários de saúde. Esse valor é comparável aos encontrados em outros instrumentos utilizados em psiquiatria, como o BDI (Roberts e cols, 1991).

O ponto de corte 24 foi o mais apropriado do ponto de vista psicométrico, baseado em critérios objetivos para sua definição. No entanto, esse escore produziu valores de sensibilidade e especificidade relativamente baixos (respectivamente 74 e 73%). Por isso, na prática, esse ponto não pode ser considerado satisfatório para o rastreamento de depressão em cuidados primários, situação em que seria recomendável um nível mais alto de sensibilidade. Esse fato confirma que a CES-D deixa a desejar se o objetivo for a correta classificação de indivíduos com diagnóstico de depressão. Por esse motivo, tem sido proposto que a CES-D na realidade não mede sintomas específicos de depressão, mas sim um sofrimento geral (Fechner-Bates e cols, 1994) ou um mal-estar psicológico que pode ser melhor caracterizado como “desmoralização” (Breslau, 1985; Roberts e cols, 1989).

Um fato que chama atenção é a discrepância entre o ponto de corte encontrado, 24, com o ponto de corte tradicional, 16. Vários estudos também encontraram pontos de corte

superiores ao ponto tradicional (Roberts e cols, 1991; Somervell e cols, 1993; Cho e Kim, 1998; Ghubash e cols, 2000; Fountoulakis e cols, 2001). Uma possível explicação é a diferença de critério utilizado para definição do ponto de corte neste e em outros estudos em relação ao estudo original de Radloff (1977). Enquanto o ponto de corte tradicional foi estabelecido através de um critério arbitrário baseado no percentil 80 (Comstock e Helsing, 1976), nosso estudo utilizou um método de validação concorrente com o uso de um teste padrão-ouro, conforme tem sido utilizado nos mais recentes estudos de validação da CES-D. Outra possível explicação para a discrepância entre nosso ponto de corte e o tradicional 16 é a tendência de que o ponto seja mais elevado em populações onde a prevalência de sintomas depressivos é usualmente maior, como em cuidados primários (Coyne e cols, 1994). Outra possibilidade é de que populações como esta, com características socioculturais diferentes do que em países mais desenvolvidos, apresentem maior frequência de sintomas depressivos subsindrômicos que não atingem intensidade suficiente para um diagnóstico clínico (Cho e Kim, 1998).

Por outro lado, quando optamos por definir um ponto de corte alternativo mais compatível com a intenção de rastrear indivíduos com sintomas depressivos, sem uma preocupação inicial com o diagnóstico definitivo, o ponto de corte sugerido de 18 aproxima-se bastante do ponto tradicional. O ajuste do ponto de corte para 18 permitiu um pequeno ganho de especificidade, de 46 para 52%, sem uma grande perda da sensibilidade; esta se manteve alta, em 91%. Ainda assim, a CES-D permaneceu apresentando uma taxa elevada de falso-positivos.

É importante levar em conta que, ainda que estejamos buscando um ponto de corte ideal para a CES-D nessa população de cuidados primários de saúde, a determinação de um ponto de corte é relativa, pois depende de uma série de fatores. Além da análise do desempenho do instrumento propriamente dito, há que se levar em conta fatores como a prevalência da doença e as conseqüências do diagnóstico e do não diagnóstico, de acordo com a gravidade e as possibilidades de tratamento da doença (Fletcher e cols, 1996; Furukawa e cols, 1997). Portanto, a definição do melhor ponto de corte é um processo arbitrário que depende da finalidade para a qual se pretende utilizar o instrumento (Prescott e cols, 1998).

Se o objetivo do rastreamento for referenciar indivíduos para uma avaliação diagnóstica mais detalhada, uma certa taxa de falso-positivos é aceitável, pois haverá correção num segundo momento (Radloff e Locke, 1982; Prescott e cols, 1998). Pode ser preferível, em cuidados primários, diagnosticar em excesso depressão no rastreamento, e confirmá-lo ou refutá-lo posteriormente, do que perdê-lo completamente com o uso de pontos de corte mais altos (Cho e Kim, 1998). Nesse sentido, o ponto de corte 18 foi o melhor de acordo com a finalidade original da CES-D, como instrumento de rastreamento primário para identificar indivíduos com sintomas depressivos a despeito de seu diagnóstico. Evidentemente haverá a necessidade de uma segunda avaliação mais detalhada a fim de confirmar (ou excluir) a suspeita diagnóstica inicial.

Caso a finalidade do uso da CES-D seja para seleção de indivíduos para fins de pesquisa, o ponto de corte 24 é o mais apropriado, pois produz uma taxa menor de

resultados falso-positivos. Desse modo, há uma redução do número de indivíduos a serem rastreados numa primeira etapa, conseqüentemente reduzindo custos na segunda etapa.

Outro achado interessante diz respeito ao desempenho da CES-D de acordo com o sexo. Do ponto de vista do ponto de corte de melhor desempenho geral, houve uma pequena diferença que não parece ser clinicamente relevante: 22 para homens e 24 para mulheres. Curiosamente, esse resultado é idêntico ao encontrado por Roberts e cols (1991), ainda que em uma população distinta, de adolescentes escolares.

Por outro lado, os pontos de corte de rastreamento para homens e mulheres são bastante diferentes (respectivamente 11 e 19), corroborando a idéia de que os sintomas depressivos podem se manifestar de modo distinto entre eles. Essa discrepância, com ponto de corte menor para indivíduos do sexo masculino, também foi encontrada por Garrison e cols (1991) em adolescentes escolares. É possível que as mulheres expressem maior quantidade de sintomas subsindrômicos, principalmente sintomas somáticos/ansiosos (Silverstein, 2002), que não chegam a configurar quadros de depressão maior. Já os homens, que habitualmente expressam menos seus sintomas, quando o fazem é porque provavelmente já atingiram um nível maior de depressão. Roberts e colaboradores (1991) chegou a sugerir que fossem utilizados pontos de corte distintos para adolescentes masculinos e femininos.

Uma limitação de nosso estudo foi o fato da coleta de dados ter ocorrido em momentos distintos, com cerca de seis meses de diferença entre as duas coletas. Teoricamente poderia haver diferenças nas características dos indivíduos que procuram

atendimento de saúde em diferentes épocas do ano. No entanto, as duas amostras apresentaram características demográficas similares (dados não mostrados), e não há indícios de que o horário da coleta, o dia da semana ou o mês do ano provoquem diferenças nos escores da CES-D (Radloff, 1977).

Outra limitação foi não termos avaliado a presença de outros diagnósticos psiquiátricos que freqüentemente coexistem com depressão maior, como transtornos de ansiedade, transtorno somatoforme e outros transtornos do humor. Conforme já comentado, é possível que essas freqüentes comorbidades tenham sido em parte responsáveis pela alta taxa de resultados falso-positivos encontrados.

Apesar dessas limitações, consideramos que nossos resultados confirmam a utilidade da CES-D como instrumento de rastreamento de depressão em um contexto de cuidados primários. É possível obter um melhor aproveitamento do instrumento se forem feitos os ajustes necessários dos pontos de corte de acordo com o sexo e com a finalidade de uso desejada.

Tabela 1. Características da amostra

	N (%)	Média	Desvio Padrão	Mínimo	Máximo
Sexo					
Masculino	100 (22.9)	---	---	---	---
Feminino	335 (76.8)				
Idade	---	38.9	13.8	18	75
Estado marital					
Casado	234 (53.7)				
Viúvo	27 (6.2)	---	---	---	---
Separado	64 (14.7)				
Solteiro	111 (25.5)				
Anos de estudo	---	9.5	3.3	1	20
CES-D		22.1	12.2	0	52
≥16 (amostra 1)	300 (68.8)				
<16 (amostra 2)	136 (31.2)				
CIDI					
Positivo	169 (38.8)	---	---	---	---
Negativo	267 (61.2)				

Tabela 2. Sensibilidade, especificidade e área sob a curva da CES-D para vários pontos de corte

Ponto de corte	Amostra total		Homens		Mulheres	
	Sens ^a (%)	Espec ^b (%)	Sens (%)	Espec (%)	Sens (%)	Espec (%)
10	98	29	93	33	99	27
11^c	98	30	93	36	99	29
12	96	34	89	38	99	33
13	96	38	87	38	99	38
14	94	42	84	42	98	42
15	94	44	84	47	98	43
16^d	93	46	82	49	97	46
17	92	49	82	51	96	49
18^c	91	52	80	53	94	51
19^c	87	55	71	55	93	55
20	84	60	71	62	89	59
21	83	62	71	66	87	61
22^e	79	65	67	73	83	63
23	76	69	60	76	81	67
24^e	74	73	60	80	79	72
25	70	74	58	82	75	72
ASC ^f (IC 95%)	0.80 (0.76-0.85)		0.77 (0.68-0.86)		0.82 (0.78-0.87)	

a: sensibilidade

b: especificidade

c: pontos de corte para rastreamento

d: ponto de corte original

e: pontos de melhor desempenho global

f: área sob a curva

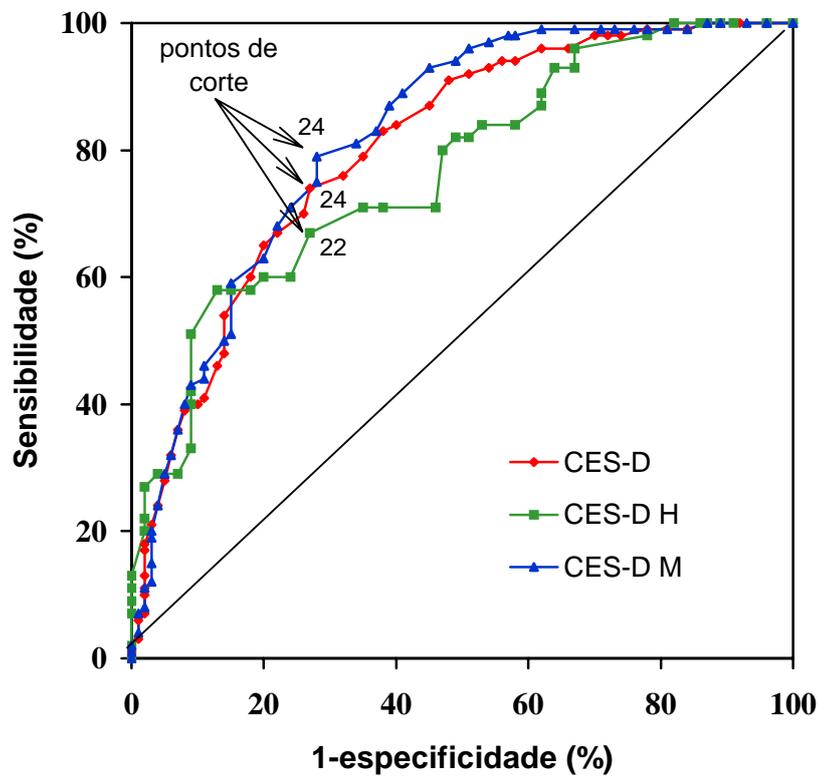


Figura 1. Curvas ROC da CES-D na amostra total, em homens e mulheres, com respectivos pontos de corte

ARTIGO EM INGLÊS (1)

**CRITERION VALIDITY OF THE CES-D IN PRIMARY CARE
PATIENTS IN BRAZIL**

Gustavo Schestatsky, MD¹ and Marcelo P. A. Fleck, MD²

1) Psychiatrist. Student from the Post-Graduate Program of the Department of Psychiatry, Federal University of Rio Grande do Sul, Brazil.

2) Psychiatrist. Associate Professor of the Department of Psychiatry, Federal University of Rio Grande do Sul, Brazil.

ABSTRACT

Objective: To assess the concurrent validity of a Brazilian version of the CES-D for screening for depression in primary care users, setting a suitable cut point for that population. **Method:** 436 individuals responded to the CES-D followed by a diagnostic interview with the CIDI instrument. A ROC analysis was performed to locate the cut points on the curve. Sensitivity and specificity of the CES-D scores were assessed. **Results:** The CES-D cut point that produced the best performance was 24 for both the total sample (sensitivity 74%, specificity 73%) and the female group. For males the best cut point was 22. As for screening, putting emphasis in sensitivity, the best scores were 18 for the total sample, 11 for male and 19 for female. **Discussion:** The CES-D demonstrated satisfactory global performance even though there were too many false-positive results. There was a striking difference in the CES-D performance within the genders when utilized with the purpose of screening. The CES-D cut point may be adjusted according to the purpose of its use and also according to the gender.

INTRODUCTION

Depressive disorders account for an enormous impact in terms of social, economic and public health issues throughout the world (Murray e Lopez, 1997). Also, major depression itself is the most prevalent mental disorder (Kessler *et al*, 1994). Depressed individuals are high utilizers of health services (Johnson *et al*, 1992), particularly in primary care settings, where the prevalence of depressive disorders is very high (Coyne *et al*, 1994). Such a scenery makes primary care settings prone to early recognition of depressive disorders (McQuaid *et al*, 1999). Nevertheless, most of the depressive episodes are not diagnosed (Pérez-Stable *et al*, 1990; Coyne *et al*, 1995), thus depriving the ill individuals from adequate treatment.

One of the possible strategies with views to decrease such an impact is improving recognition of depression in primary care services, through self-administered screening instruments. One of the most frequently used instruments in epidemiologic studies for the screening of depression is the *Center for Epidemiologic Studies – Depression (CES-D) Scale* (Radloff, 1977), which was developed for the study of depressive symptoms in the general population. The CES-D has been largely used within the most different contexts, in several countries, including primary care populations (Lyness *et al*, 1997; Irwin *et al*, 1999).

Since the early studies of the CES-D has it validity and reliability been demonstrated as satisfactory for the screening of depression symptoms across various populations (Radloff, 1977; Weissman *et al*, 1977; Myers and Weissman, 1980;

Roberts, 1980; Roberts and Vernon, 1983). Nevertheless, its usefulness in primary care setting is still controverse (Fechner-Bates *et al*, 1994; Klinkman *et al*, 1997).

The performance of the CES-D may vary according to the context in which it is used, being subject to cultural influences. In particular, the most appropriate cut point for a given population may highly differ from the originally proposed cut point. In general the cut points tend to be lower in non clinical populations (Silveira and Jorge, 1998; Salle, 1999), and higher in clinical populations, including primary care populations (Lyness *et al*, 1997; Cho an Kim, 1998; Fountoulakis *et al*, 2001). For this reason it becomes important to define a cut point that is adequate for the context in which the scale is being studied, in order to maximize its performance within a particular population.

In Brazil, there are scarce studies on the CES-D, and none of them are related to primary care populations (Silveira and Jorge, 1998; Salle, 1999). One of the studies that used the CES-D to detect depression in a Brazilian primary care population suggested that the traditional cut point 16 is probably too low for that population, yelding too many false-positive results (Fleck *et al*, 2002).

Thus, the present study aimed at evaluating the concurrent validity of a Brazilian version of the CES-D as a screening instrument for depression in primary care users, establishing the most appropriate cut point for this population.

METHOD

Sampling procedures

Four hundred and thirty-six individuals who were attended in three primary care clinics from a delimited area of Porto Alegre, Brazil, were included in the study.

The data were obtained in two phases, constituting two additional samples taken from the same population. The first sample included individuals who had participated in the LIDO study, a longitudinal observational study of depressed patients in primary care (Fleck *et al*, 2002). In this phase 300 individuals whose CES-D scale score was 16 or higher (traditional cut point) were selected. The additional sample (second phase of data collection) included 136 participants whose score was lower than 16, in order to obtain the whole CES-D score spectrum distributed in that population (not only the ones with scores 16 or higher from the first collection phase).

Subjects aged 18 through 75 who lived within the nearby area of the clinics and were searching for care were invited to participate. Individuals who were unable to understand how to fill in the self-responder questionnaire, those who were in treatment for depression and those with prior diagnosis of any major psychiatric disorders were not included.

Data procedures

Subjects who searched for help in the mentioned clinics during the data collection procedures period and who fulfilled the inclusion or exclusion criteria were consecutively invited to participate in the study and signed an informed consent document. There was a team of interviewers covering most of the days of the week in the clinics. The subjects who agreed to participate in the study answered the CES-D self-respondered screening instrument in the clinics' waiting-rooms.

Based on the scores obtained in this recruitment phase (being the $CES-D \geq 16$ as criterion of interest in the first phase of collection and $CES-D < 16$ as criterion of interest in the second phase), the subjects participated in a second phase assessment, within a maximum period of 14 days. In this assessment, which could be performed either within the clinic itself or in the subjects' homes, they responded the CES-D again and were then submitted to a diagnostic interview using the *Composite International Diagnostic Interview (CIDI)* instrument.

Diagnostic instruments

- **CES-D:** It was developed to assess depressive symptoms in the general population. It does not provide depression diagnosis but rather identifies individuals more likely to be depressed. It comprises 20 items which are considered to be representative of the main depression symptoms. The individuals rate each item using a frequency scale (from

0=rarely or never to 3=most of the time or all the time) in relation to their last week. The final score is set by the simple addition, possibly ranging from 0 to 60. The cut point 16 or higher is usually used to distinguish individuals more likely to be depressed. The CES-D is a self-responed scale that is simple to be applied and easily understood. In this study we used a Portuguese language version which had been translated and adapted by a bilingual panel under the supervision of a foreigner researcher.

- **CIDI:** It was developed to be used in primary health care as a diagnosis instrument for a wide range of mental health problems including depression. In the present study the CIDI was applied by trained interviewers, being utilized the depression module from its revised version (2.1).

Data analysis

Sensitivity and specificity of the various CES-D cut points were estimated using the major depression diagnosis provided by the CIDI as the gold-standard. The sensitivity and the specificity variation along the cut points spectrum made possible to determine the more suitable points. We decided to analyse separately the male and female groups in order to check possible gender differences in scale performance.

In order to obtain the the scale discrimination ability within all possible cut points and also to determine the most appropriate cut points, a ROC (*receiver operating characteristic*) analysis was performed, producing a ROC curve (Murphy *et al*, 1987). The ideal cut points were identified as being the curve points with the longest perpendicular

distance from the diagonal line that departs from the 0,0 coordinate to the 1,1 coordinate (random line). These were the points that accounted for the lowest sum of the false-positive and false-negative rates (Garrison *et al*, 1991).

As a global performance parameter of the CES-D, the area under the ROC curve (AUC) was used. The AUC is a global index of the discrimination ability or accuracy of the scale, summarizing the instrument performance as a whole (Fletcher *et al*, 1996; Silveira and Jorge, 2000). The larger the AUC (the higher the curve), the better the instrument is (Garrison *et al*, 1991).

SPSS version 8.0 was used for data analysis.

RESULTS

The sample was composed of 436 individuals (Table 1). The average age was 38,9, being the female the predominant gender (76,8%). Most of the individuals were married (53,7%), following the singles (25,5%), separated (14,7%) and widowed (6,2%). As for schooling, the average was 9,5 years of study. The CES-D average score was 22,1, although this is not a representative average for this population (as was mentioned before, the part of the sample with CES-D \geq 16 has been over-represented).

The AUC was 0,80 in the total sample, demonstrating a satisfactory global performance of the CES-D (Figure 1). That performance was slightly better in the female group (ASC=0,82) compared to the male group (ASC=0,77). The irregular format of the ROC curve in the male group probably reflects the most reduced sample size in this subgroup.

Table 2 shows the balance between sensitivity and specificity in the various CES-D cut points. As the cut point increases, the sensitivity decreases and the specificity increases.

Using the cut point 16, usually indicated for screening procedures of possible cases of depression, sensitivity was 93% and specificity was 46%. The best CES-D performance, considered the point with the most balanced values of sensitivity and specificity (also the ROC curve highest point) occurred at the cut point 24, producing 74% of sensitivity and 73% of specificity.

When the performance was analysed according to gender, the same cut point 24 was found for females, with 79% of sensitivity and 72% of specificity. In the male group the best cut point was 22, with 67% percent of sensitivity and 73% of specificity.

Due to the low sensitivity obtained with the best global performance cut point, the option was to define an alternative cut point to be more appropriate to the CES-D original function, that is, the screening for depressive symptoms in the general population. For this reason, we considered adequate to define a point which could grant more importance to sensitivity, even at the cost of lowering the specificity. Thus the failure rate for the

detection of depressed individuals (false-negative results) is reduced to a more acceptable level, below 10 percent.

Thus, after analysing the performance changes along the cut points spectrum, we considered the cut point 18 as the most appropriate for the screening of depressive symptoms in the studied population. At this point the sensitivity loss is minimal, still remaining high (91%), and there is a small increase in specificity, which reaches 52%. Above that point, the sensitivity dramatically falls down to levels below 90 percent, with a progressive loss of the ability to identify depressed individuals, and this fact would not be desirable as far as depression screening is concerned.

As for females, the best screening cut point was 20, with 93% of sensitivity and 55% of specificity. As for males, the best cut point was 11, with 93% percent of sensitivity and 36% of specificity.

DISCUSSION

In a population of users of primary health care services, the CES-D presented a performance compatible with the screening purpose. At the cut point 16, it showed to be very sensitive, hardly never failing in identifying cases of depression. The counterpart, however, was its low specificity levels, accounting for a high rate of false-positive results.

This finding is according with studies that also found very high rates of false-positive results (Roberts *et al*, 1989; Garrison *et al*, 1991; Roberts *et al*, 1991), including in the primary care context (Fechner-Bates *et al*, 1994). A study developed in Brazil, using the CES-D at the cut point 16 to assess the presence of depression in a primary care population, found an excessively high rate of depressed individuals. The authors have suggested that that cut point was also accounting for too many false-positive results (Fleck *et al*, 2002).

It is possible that the presence of non-detected psychiatric comorbidities, such as anxiety disorders, adjustment disorders, dysthymia and somatoform disorder, which are quite frequent in primary care, have contributed for the high CES-D scores which have not been configured as diagnosis of major depression, this way contributing to a surplus of false-positive results (Boyd *et al*, 1982; Roberts *et al*, 1989). Accordingly it is possible that, since the studied population belongs from low socioeconomical patterns, the occurrence of temporary and subsyndromal depressive symptoms may be higher, fact that would also contribute for higher levels of false-positive results. (Cho and Kim, 1998).

The AUC of 0,80 demonstrates that the CES-D is moderately accurate as a screening instrument in primary care services. This value is comparable to those found for other instruments used in psychiatry such as the BDI (Roberts *et al*, 1991).

The cut point 24 was the most appropriate from the psychometric point of view, based on objective criteria for its definition. However, that score accounted for quite low values of sensitivity and specificity (74% and 73%, respectively). For that reason, that point cannot be considered as satisfactory in terms of depression screening in primary care,

where a higher level of sensibility should be preferred. That fact confirms that the CES-D is not absolutely suitable if the objective be a correct classification of individuals with depression diagnosis. For this reason, it has been assumed that actually the CES-D does not measure specific depression symptoms, but surely can measure a general distress state (Fechner-Bates *et al*, 1994) or a psychological distress that may be better described as “demoralization” (Breslau, 1985; Roberts *et al*, 1989).

A striking fact is the discrepancy between the found cut point 24 and the traditional cut point 16. Several studies also found higher cut points than the traditional one (Roberts *et al*, 1991; Somervell *et al*, 1993; Cho and Kim, 1998; Ghubash *et al*, 2000; Fountoulakis *et al*, 2001). A possible explanation is the criterion difference used for the definition of the cut point in the present study as well as in other ones in relation to the original study by Radloff (1977). While the traditional cut point was established by an arbitrary criterion based on percentil 80 (Comstock e Helsing, 1976), our study used a concurrent validation method using a gold-standard test, as used in CES-D validation studies. Another possible explanation for the discrepancy between our cut point and the traditional cut point 16 one is the tendency that the point may be higher in populations where the prevalence of depressive symptoms is usually higher, as in primary care (Coyne *et al*, 1994). Another possibility is that such populations, with sociocultural features distinct from more developed countries present higher frequency rates of subsyndromal depressive symptoms which do not reach enough intensity for a clinical diagnosis (Cho and Kim, 1998).

On the other hand, when our option was centered in the definition of an alternative cut point more compatible with the intention of screening individuals with symptoms of

depression, without any preliminary concern with the definite diagnosis, the suggested cut point 18 becomes quite closer to the traditional point. The cut point arrangement to 18 permitted some gain in specificity, from 46% to 52%, being small the loss of sensitivity; having the sensitivity remained high, in 91 percent. Even though the CES-D remained producing a high rate of false-positive results.

It is important to consider that, although we are searching for an ideal cut point for the CES-D within that primary health care population, the setting of a cut point is relative since it depends on various factors. Besides the instrument performance analysis itself, one has to take into consideration some factors such as the illness prevalence and the consequences of the diagnosis and the non-diagnosis as well, according to the seriousness and the possibilities of the disease treatment. (Fletcher *et al*, 1996; Furukawa *et al*, 1997). Thus the definition of the best cut point is an arbitrary process that depends on the aim of the instrument utilization (Prescott *et al*, 1998).

Whether the screening objective is the referral of individuals for a more detailed diagnostic assessment, a certain amount of false-positive results is acceptable, for there will be a correction later on (Radloff and Locke, 1982; Prescott *et al*, 1998). In primary care it is likely to be preferable to diagnose excessive cases of depression, and/or later on confirm or exclude them, otherwise one can lose them completely when using higher cut points (Cho and Kim, 1998). In this case the cut point 18 was the most suitable according to the original CES-D aim as a primary screening instrument used to identify individuals with depressive symptoms whatever their diagnosis is. Of course it will be necessary a second more detailed assessment in order to confirm or to exclude the preliminary diagnostic suspicion.

In case the CES-D be utilized to select individuals for research purposes, the cut point 24 is the most suitable as it accounts for a low rate of false-positive results. In that case there will be a reduction of the number of individuals to be screened in the first phase, thus reducing the costs in the second phase.

Another outstanding finding refers to the CES-D performance according to gender. As for the cut point best general performance, there was a quite small difference which does not seem to be clinically relevant: 22 for males and 24 for females. Curiously, that result is identical to the one found by Roberts *et al* (1991), even though that study used a school adolescent population.

In contrast the screening cut points for males and females are quite different (11 and 19, respectively), strengthening the idea that depressive symptoms may be manifested in distinct ways among men and women. That discrepancy, with lower cut point for male individuals, was also found by Garrison *et al* (1991) in school adolescents. In one way, there is a possibility that females express a greater amount of subsyndromal symptoms, especially anxious/somatic symptoms (Silverstein, 2002) not intense enough to reach a clinical picture of depression. While in males, who usually express less often their symptoms, will do it only when they reach a higher level of depression. Roberts *et al* (1991) even suggested distinct cut points for male and female adolescents.

One limitation of the present study was the fact that the data collection occurred in distinct moments, with about six months interval between the two periods. In theory there could have had differences in the characteristics of the individuals who searched for health

help in different periods of the year. However, the two samples showed demographic characteristics quite similar (data not shown), and there is no data to suggest that the hour, the day of the week or the month of the year can interfere in the CES-D scores (Radloff, 1977).

Another limitation was the lack of assessment for other psychiatric diagnoses which are usually comorbid with major depression, such as anxiety disorders, somatoform disorders and other mood disorders. As we mentioned before, it is possible that those frequent comorbidities have been partially responsible for the high rates of false-positive results found.

Despite the limitations cited above, we consider that the obtained results confirm the CES-D usefulness as a screening instrument for depression in a primary care context. It is possible to take more significant advantage of the instrument if necessary adjustments in the cut points be made in terms of gender and according to the objective of its utilization.

References

Boyd JH, Weissman MM, Thompson WD, *et al.* Screening for depression in a community sample: understanding the discrepancies between depression symptom and diagnostic scales. *Arch Gen Psychiatry* 39:1195-1200, 1982.

Cho MJ, Kim AKH. Use of the Center for Epidemiologic Studies Depression (CES-D) Scale in Korea. *J Nerv Ment Dis* 186: 304-310, 1998.

Coyne JC, Fechner-Bates S, Schwenk TL. Prevalence, nature and comorbidity of depressive disorders in primary care. *Gen Hosp Psychiatry* 16: 267-276, 1994.

Coyne JC, Schwenk TL, Fechner-Bates S. Nondetection of depression by primary care physicians reconsidered. *Gen Hosp Psychiatry* 17: 3-12, 1995.

Fechner-Bates S, Coyne JC, Schwenk TL. The relationship of self-reported distress to depressive disorders and other psychopathology. *J Consult Clin Psychol* 62: 550-559, 1994.

Fletcher RH, Fletcher SW, Wagner EH. *Epidemiologia Clínica: Elementos Essenciais*. 3 ed. Porto Alegre, Artes Médicas, 1996.

Fountoulakis K, Iacovides A, Kleanthous S, e cols. Reliability, validity and psychometric properties of the Greek translation of the Center for Epidemiological Studies – Depression (CES-D) Scale. *BMC Psychiatry* 1: 3-13, 2001.

Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. *J Affect Disord* 46 (1): 1-13, 1997.

Garrison CZ, Addy CL, Jackson KL, e cols. The CES-D as a screen for depression and other psychiatric disorders in adolescents. *J Am Acad Child Adolesc Psychiatry* 30 (4): 636-41, 1991.

Ghubash R, Daradkeh TK, Al Naseri KS, e cols. The performance of the Center for Epidemiologic Study Depression Scale (CES-D) in an Arab female community. *International Journal of Social Psychiatry* 46 (4): 241-249, 2000.

Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). *Arch Intern Med* 159:1701-1704, 1999.

Johnson J, Weissman MM, Klerman GL. Service utilization and social morbidity associated with depressive symptoms in the community. *JAMA* 267 (11): 1478-1483, 1992.

Kessler RC, McGonagle, Zhao S, e cols. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Study. *Arch Gen Psychiatry* 51: 8-19, 1994.

Klinkman MS, Coyne JC, Gallo S, Schwenk TL. Can case-finding instruments be used to improve physician detection of depression in primary care? *Arch Fam Med* 6 (6): 567-73, 1997.

Lyness JM, Noel TK, Cox C, e cols. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Depression Scale. *Arch Intern Med* 157 (4): 449-454, 1997.

McQuaid JR, Stein MB, Laffaye C, McCahill. Depression in a primary care clinic: the prevalence and impact of an unrecognized disorder. *J Affect Disord* 55: 1-10, 1999.

Murphy JM, Berwick DM, Weinstein MC, e cols. Performance of screening and diagnostic tests: application of Receiver Operating Characteristic Analysis. *Arch Gen Psychiatry* 44: 550-555, 1987.

Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 349: 1436-1442, 1997.

Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry* 137 (9): 1081-1084, 1980.

Pérez-Stable EJ, Miranda J, Muñoz RF, Ying YW. Depression in medical outpatients: underrecognition and misdiagnosis. *Arch Intern Med* 150 (5): 1083-1088, 1990.

Prescott CA, McArdle JJ, Hishinuma ES, e cols. Prediction of major depression and dysthymia from CES-D scores among ethnic minority adolescents. *J Am Acad Child Adolesc Psychiatry* 37 (5): 495-503, 1998.

Radloff LS. The CES-D Scale: a self-report depression scale for research in general population. *Applied Psychol Measurement*, 1 (3), 385-401, 1997.

Radloff LS, Locke BZ. The community mental health survey and CES-D scale. In: Weissman MM, Myers JK, Ross CE (eds) *Community survey of psychiatric disorders*. New Brunswick, NJ. Rutgers University Press, 1986.

Roberts RE. Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Res* 2: 125-134, 1980.

Roberts RE, Vernon SW. The Center for Epidemiologic Studies Depression Scale: its use in a community sample. *Am J Psychiatry* 140 (1): 41-6, 1983.

Roberts RE, Vernon SW, Rhoades HM. Effects of language and ethnic status on reliability and validity of the Center for Epidemiologic Studies-Depression Scale with psychiatric patients. *J Nerv Ment Dis* 177 (10): 581-92, 1989.

Roberts RE, Lewinsohn PM, Seeley JR. Screening for adolescent depression: a comparison of depression scales. *J Am Acad Child Adolesc Psychiatry* 30 (1): 58-66, 1991.

Salle E. Estudo da sintomatologia depressiva em adolescentes de 15 a 17 anos de uma escola de 2º grau de Porto Alegre através das escalas auto-aplicáveis BDI, CRS e CES-D. Dissertação de Mestrado em Clínica Médica/UFRGS. Porto Alegre, abril de 1999.

Silveira DX, Jorge MR. Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens. *Rev Psiq Clín* 25 (5) Edição Especial: 251-261, 1998.

Silveira DX, Vieira A, Palomo VR, Doering Silveira E. Validade de critério e confiabilidade da versão brasileira de uma escala de rastreamento para dependência de sexo. *Revista Brasileira de Psiquiatria* 22 (1): 4-10, 2000.

Silverstein B. Gender differences in the prevalence of somatic versus pure depression: a replication. *Am J Psychiatry* 159 (6): 1051-1052, 2002.

Somervell PD, Beals J, Kinzie JD, e cols. Criterion validity of the Center for Epidemiologic Studies Depression Scale in a population sample from an American indian village. *Psychiatry Res* 47 (3): 255-66, 1993.

Weissman MM, Sholomskas D, Pottenger M, e cols. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol* 106: 203-214, 1977.

Table 1. Characteristics of the sample

	N (%)	Mean	Standard Deviation	Minimum	Maximum
Gender					
Male	100 (22.9)	---	---	---	---
Female	335 (76.8)				
Age	---	38.9	13.8	18	75
Marital status					
Married	234 (53.7)				
Widow	27 (6.2)	---	---	---	---
Separated	64 (14.7)				
Single	111 (25.5)				
Years of study	---	9.5	3.3	1	20
CES-D		22.1	12.2	0	52
≥16 (sample 1)	300 (68.8)				
<16 (sample 2)	136 (31.2)				
CIDI					
Positive	169 (38.8)	---	---	---	---
Negative	267 (61.2)				

Table 2. Sensitivity, specificity and area under curve of the CES-D

Cut points	Total sample		Male		Female	
	Sens ^a (%)	Spec ^b (%)	Sens (%)	Spec (%)	Sens (%)	Spec (%)
10	98	29	93	33	99	27
11^c	98	30	93	36	99	29
12	96	34	89	38	99	33
13	96	38	87	38	99	38
14	94	42	84	42	98	42
15	94	44	84	47	98	43
16^d	93	46	82	49	97	46
17	92	49	82	51	96	49
18^c	91	52	80	53	94	51
19^c	87	55	71	55	93	55
20	84	60	71	62	89	59
21	83	62	71	66	87	61
22^e	79	65	67	73	83	63
23	76	69	60	76	81	67
24^e	74	73	60	80	79	72
25	70	74	58	82	75	72
AUC ^f (IC 95%)	0.80 (0.76-0.85)		0.77 (0.68-0.86)		0.82 (0.78-0.87)	

a: sensitivity

b: specificity

c: cut points for screening

d: original cut point

e: best performance cut points

f: area under curve

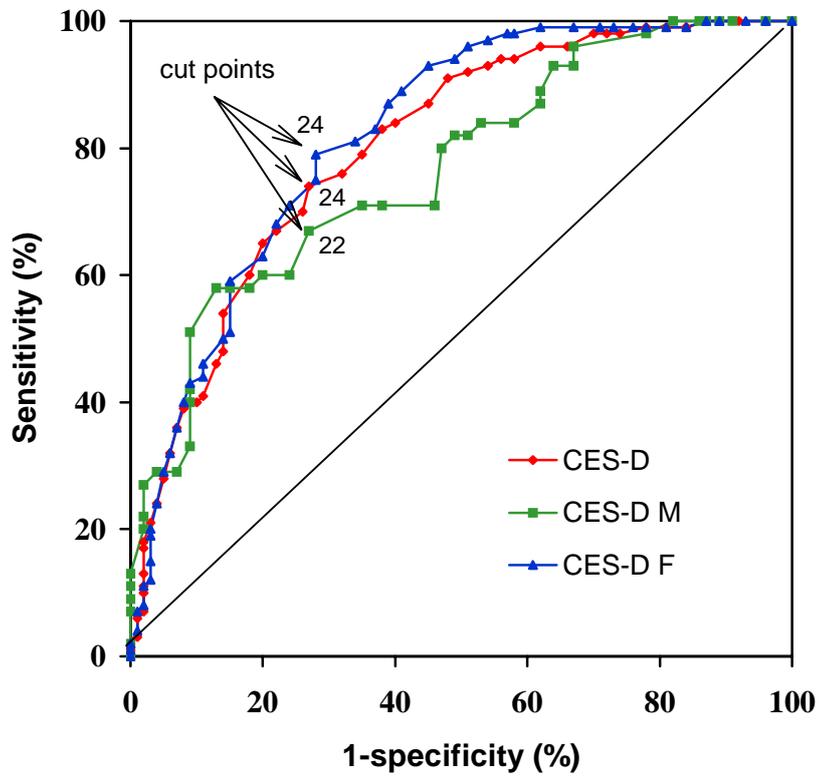


Figure 1. ROC curves of the CES-D for the total sample, males and females, and their cut points

ARTIGO EM PORTUGUÊS (2)

**ESTUDO DA CONFIABILIDADE E ESTRUTURA FATORIAL DA
CES-D EM CUIDADOS PRIMÁRIOS**

Gustavo Schestatsky¹, Marcelo P. A. Fleck²

1) Médico Psiquiatra. Mestrando do Programa de Pós-Graduação em Ciências Médicas: Psiquiatria da Universidade Federal do Rio Grande do Sul.

2) Médico Psiquiatra. Professor Adjunto do Departamento de Psiquiatria e Medicina Legal da Universidade Federal do Rio Grande do Sul.

SUMÁRIO

Objetivo: Estudar a confiabilidade e estrutura fatorial de uma versão brasileira da CES-D em uma população de usuários de um serviço de cuidados primários de saúde. **Método:** 436 indivíduos responderam ao instrumento de rastreamento populacional de depressão CES-D. Foram avaliadas a confiabilidade em termos de consistência interna (coeficiente alfa de Cronbach) e a estrutura fatorial através do método de análise de componentes principais. **Resultados:** A CES-D apresentou consistência interna de 0,90. A análise fatorial revelou uma solução com três fatores: (1) afeto depressivo e somático/retardo motor; (2) afeto positivo; (3) interpessoal. **Discussão:** A CES-D apresentou boa confiabilidade para o uso em uma população de cuidados primários. A estrutura fatorial mostrou-se semelhante a do estudo original de Radloff (1977), porém agrupando os itens relativos ao afeto depressivo e os somáticos e de retardo motor em um único fator. Essa solução fatorial pode ter explicação do ponto de vista cultural, reforçando a boa validade de conteúdo e de construção da CES-D em uma população brasileira de cuidados primários.

INTRODUÇÃO

A *Center for Epidemiologic Studies – Depression Scale* (CES-D) é uma escala desenvolvida para medir sintomas depressivos na população geral (Radloff, 1977), sendo constituída de 20 itens considerados representativos dos sintomas principais da depressão.

Sua importância como instrumento de rastreamento de depressão tem sido consolidada através de inúmeros estudos se propuseram a analisar suas propriedades psicométricas nos mais variados contextos. Há estudos em populações não clínicas (Silveira e Jorge, 1998; Ghubash e cols, 2000), em populações clínicas (Shrout e Yager, 1989; Fountoulakis e cols, 2001), em determinadas faixas etárias (Beekman e cols, 1997, Salle, 1999), em diferentes países (Furukawa e cols, 1997; Cho e Kim, 1998; Schroevers e cols, 2000), além de algumas populações bastante específicas (Somervell, 1993, Wong, 2000; Foley e cols, 2002).

Por tratar-se de um instrumento com potencial de aplicabilidade para rastreamento de transtornos depressivos, também tem sido estudada em populações de usuários de serviços de atenção primária de saúde (Lyness e cols, 1997), locais onde os transtornos do humor são bastante prevalentes (Coyne e cols, 1994).

Alguns dos aspectos avaliados que contribuem para estabelecer sua validade de construção e de conteúdo da CES-D são a sua confiabilidade e estrutura fatorial. A confiabilidade da CES-D tem sido descrita em termos de sua consistência interna, a qual mostra-se geralmente elevada (acima de 0,85 na maioria dos estudos). Os estudos de

análise fatorial vieram contribuir para a validade de conteúdo e de construção da CES-D, ao extrair os domínios principais que compõem o instrumento e permitir comparações da sua estrutura em diferentes populações.

O estudo original de Radloff (1977) foi o primeiro a extrair, por meio do método de análise de componentes principais, uma solução de quatro fatores, os quais denominou: (1) afeto depressivo; (2) afeto positivo; (3) somático e retardo motor; (4) interpessoal. Estudos fatoriais subseqüentes, que também utilizaram a análise de componentes principais, confirmaram esse achado (Roberts, 1980; Kuo, 1984; Shrout e Yager, 1989; Gatz, 1993). Estudos de análise fatorial confirmatória replicaram uma solução de quatro fatores, demonstrando, ainda, que estes se correlacionam fortemente entre si (Hertzog e cols, 1990; Sheehan e cols, 1995; Wong, 2000).

Outros estudos, no entanto, chegaram a soluções com três fatores, geralmente combinando os fatores “afeto depressivo” e “somático e retardo motor” de Radloff (1977) em um único fator (Ying, 1988; Guarnaccia e cols, 1989; Ghubash e cols, 2000; Fountoulakis e cols, 2001). Por isso, acredita-se que a estrutura original da CES-D com quatro fatores possa não ser aplicável a populações diversificadas, podendo ser influenciada por características socioculturais ou de linguagem próprias de diferentes populações (Foley e cols, 2002). A Tabela 1 resume alguns dos principais estudos fatoriais da CES-D.

O presente estudo teve por objetivo estudar a validade de construção e de conteúdo da CES-D em uma população de usuários de um serviço de cuidados primários brasileiro, utilizando como parâmetros a sua confiabilidade e estrutura fatorial.

MÉTODOS

Amostra

Participaram do estudo 436 indivíduos que procuraram atendimento em três postos de um serviço municipal de cuidados primários de saúde ligados a um hospital de referência, em uma área delimitada da cidade de Porto Alegre. Os motivos da procura incluíam atendimentos típicos de um serviço de cuidados primários, como consultas médicas, odontológicas e de enfermagem.

Os dados foram coletados em duas etapas, constituindo duas amostras complementares, oriundas da mesma população. A primeira amostra foi composta por indivíduos participantes do estudo LIDO (Fleck e cols, 2002), um estudo longitudinal observacional de pacientes depressivos em cuidados primários. Nesta etapa foram selecionados 300 indivíduos cujo escore na escala CES-D fosse de 16 ou mais (ponto de corte tradicional). A amostra complementar (segunda etapa da coleta) foi constituída de 136 indivíduos que apresentavam escore inferior a 16 na CES-D, a fim de que se obtivesse toda a gama de escores da CES-D distribuídos nessa população (não somente aqueles com escores de 16 ou mais da primeira etapa de coleta).

Foram convidados a participar todos os sujeitos com idades entre 18 e 75 anos, pertencentes à área geográfica dos postos, que estivessem buscando atendimento nos mesmos. Não foram incluídos aqueles que não demonstravam capacidade de entendimento

para completar o questionário auto-aplicado, os que se encontravam em tratamento para depressão, e os que tinham diagnóstico de algum transtorno psiquiátrico maior.

Coleta de dados

Os sujeitos que buscaram atendimento nos referidos postos nos períodos de coleta, e que satisfizeram os critérios de inclusão e exclusão, foram convidados de modo consecutivo a participar do estudo e assinaram um consentimento informado. Havia uma equipe de entrevistadores que cobria diversos turnos de atendimento nos referidos postos. Os sujeitos que concordaram em participar do estudo responderam ao instrumento de rastreamento auto-aplicado CES-D nas salas de espera dos postos. Para este estudo foi utilizada uma versão na língua portuguesa, a qual havia sido traduzida e adaptada por um painel bilíngüe com a supervisão de uma pesquisadora estrangeira.

Análise dos dados

A confiabilidade da CES-D foi estimada a partir da determinação de sua consistência interna. Para tanto, foi utilizado o coeficiente alfa de Cronbach. Esse coeficiente estima o grau de correlação entre os itens da escala para verificar se eles estão medindo o mesmo construto. Quanto maior a correlação dos itens da escala no sentido de medir o mesmo construto (consistência interna), maior a sua confiabilidade (Silveira e Jorge, 1998).

A análise da estrutura fatorial da CES-D foi realizada utilizando-se a técnica de análise de componentes principais (ACP). Esta técnica fornece o número de componentes (fatores) da escala, a variância explicada de cada componente, e as cargas fatoriais de cada questão (Silveira e Jorge, 1998). Para a determinação do número de componentes da escala, consideramos válidos aqueles com auto-valor maior do que a unidade (critério de Kaiser) (Fleck e Bourdel, 1998). A fim de melhorar o potencial de interpretação dos fatores, procedemos ao método de rotação ortogonal Varimax (Artes e Barroso, 1998). Foram incluídos na solução de fatores somente os itens com carga fatorial acima de 0,5.

Foi utilizado o programa estatístico SPSS versão 8.0.

RESULTADOS

A amostra total constituiu-se de 436 indivíduos. A média de idade foi de 38,9 anos (desvio-padrão 13,8), com predomínio do sexo feminino (76,8%). A maior parte dos indivíduos eram casados (53,7%), seguida de solteiros (25,5%), separados (14,7%) e viúvos (6,2%). O nível de instrução foi, em média, de 9,5 (desvio-padrão 3,3) anos de estudo. A média dos escores da CES-D foi de 22,1 pontos, embora esta não seja uma média representativa desta população (conforme explicado acima, a parte da amostra com CES-D ≥ 16 esteve super-representada).

Com relação à consistência interna, a CES-D apresentou coeficiente alfa de Cronbach, de 0,90.

A ACP antes da rotação ortogonal mostrou que 14 dos 20 itens apresentaram uma correlação significativa (acima de 0,6) com o fator principal (**Fator 1**). Após a rotação Varimax obtivemos uma solução de três fatores, os quais explicaram conjuntamente 55,2% da variabilidade total dos dados. Nenhum item ficou de fora dos principais componentes pelo critério de carga fatorial acima de 0,5, sugerindo que todos os itens da escala foram importantes para a composição dos fatores.

A composição final dos três fatores obtidos pela ACP, bem como as cargas fatoriais de cada questão, encontram-se na Tabela 2. Nessa composição, o **Fator 1** (o mais importante em termos da explicação da variância) foi composto tanto por questões relativas ao afeto depressivo propriamente dito (“humor não reativo”, “depressão”, “fracasso”, “medo”, “solidão”, “choro”, “tristeza”), como por questões referentes a sintomas somáticos e retardo motor (“aborrecimento”, “falta de apetite”, “baixa concentração”, “lassidão”, “sono agitado”, “baixa comunicação”, “desânimo”). O **Fator 2** agrupou questões relativas a um afeto positivo (“bem-estar”, “esperança”, “felicidade”, “prazer”), e o **Fator 3** relacionou-se a questões da esfera interpessoal (“pessoas não amigáveis”, “rejeição”).

Tomando como referência o estudo original de Radloff (1977), observou-se que, das 20 questões da escala original, 11 se mantiveram em agrupamentos fatoriais idênticos àqueles encontrados por aquela autora: “humor não reativo”, “depressão”, “solidão”, “choro” e “tristeza” no **Fator 1**; “bem-estar”, “esperança”, “felicidade” e “prazer” no **Fator**

2; “pessoas não amigáveis” e “rejeição” no **Fator 3** (que corresponde ao fator 4 de Radloff). Curiosamente, todas as questões que no estudo de Radloff compunham o fator somático (“aborrecimento”, “falta de apetite”, “lassidão”, “sono agitado”, “desânimo”), juntamente com outras quatro que não estiveram fortemente associadas a nenhum fator no estudo original (“baixa concentração”, “fracasso”, “medo”, “baixa comunicação”) agruparam-se no **Fator 1**, junto às questões relativas ao afeto depressivo.

A fim de comparar de modo mais direto nossos dados com o estudo original de Radloff (1977) e outros que revelaram uma solução de 4 fatores, também avaliamos uma solução com este número de fatores (dados não mostrados). Nesta hipótese, a questão 2, relativa à falta de apetite, comporia isoladamente o quarto fator. Esse fato confirma que a solução de 3 fatores é a mais adequada para nossa amostra, já que não faria sentido do ponto de vista clínico um fator composto por somente uma questão.

DISCUSSÃO

A consistência interna de 0,90 (alfa de Cronbach) foi comparável, sendo geralmente superior, ao valor encontrado na maioria dos estudos (Tabela 1). Consideramos este um indicativo de que a CES-D manteve sua confiabilidade para o uso em nossa população de usuários de serviços de cuidados primários. O grau elevado da consistência interna sugere um desempenho unidimensional da CES-D nessa população, presumivelmente relacionado ao construto da depressão.

Com relação ao número de fatores extraídos através da ACP (três fatores), nosso estudo difere do estudo original de Radloff (1977) e de outros que encontraram soluções de quatro fatores em populações norte-americanas e européias (Shrout e Yager, 1989; Gatz e cols, 1993).

Por outro lado, nossa solução de três fatores, com a combinação de itens relativos ao afeto depressivo e somático/retardo motor está de acordo com estudos realizados em populações distintas da norte-americana e européia (Ying, 1988; Guarnaccia e cols, 1989; Ghubash e cols, 2000; Fountoulakis e cols, 2001). Chama atenção que as populações desses estudos provêm de minorias étnicas, especialmente de cultura oriental, ou de países como Arábia Saudita e Grécia. No caso das duas últimas, trata-se de populações com características teoricamente semelhantes à nossa do ponto de vista do desenvolvimento socioeconômico.

Consideramos que nosso achado de confluência de sintomas essencialmente depressivos com sintomas somáticos e de retardo motor possa ter explicação do ponto de vista cultural. Populações não-ocidentais e/ou com menor desenvolvimento socioeconômico, como é o caso da nossa amostra, tendem a expressar seu sofrimento psicológico somaticamente (Kuo, 1984; Parker e cols, 2001). Ao contrário, acredita-se que indivíduos de culturas com maior grau de desenvolvimento, como a dos Estados Unidos e Europa Ocidental, tenham adquirido maior capacidade cognitiva para verbalização de estados afetivos, dessa forma utilizando-se menos da somatização (Katon e cols, 1982).

Esse fato pode ser particularmente verdadeiro no caso de usuários de serviços de cuidados primários. Isto porque, nestes locais, não somente a prevalência de transtornos depressivos é bastante elevada (Coyne, 1994), como esses indivíduos procuram tais serviços devido a queixas somáticas que muitas vezes estão ligadas a sintomas depressivos (Fava, 2002).

Por fim, consideramos um achado importante o fato de o **Fator 1** ter sido composto por itens relativos ao afeto depressivo e a sintomas somáticos e de retardo motor. Essa composição está de acordo com outros instrumentos consagrados como a Escala de Hamilton (Hamilton, 1960), que também identificaram um fator depressivo ou melancólico composto tanto por humor deprimido como por retardo motor. Além disso, os itens que constituíram o **Fator 1** em nosso estudo fazem sentido do ponto de vista clínico, pois são exatamente os chamados sintomas “nucleares” da depressão (Fleck, 1997).

Uma limitação de nosso estudo foi não termos avaliado a presença de outros diagnósticos psiquiátricos que freqüentemente coexistem com depressão maior, principalmente no contexto de cuidados primários, como transtornos de ansiedade, transtorno somatoforme e outros transtornos do humor (referências). Essas comorbidades poderiam estar compondo uma amostra “híbrida” que não representaria apenas os indivíduos com depressão maior.

Apesar dessa limitação, consideramos que a composição da CES-D obtida realmente representa os principais sintomas da depressão, o que reforça sua validade de conteúdo e de construção para o uso em uma população de cuidados primários brasileira.

Cabe lembrar que a análise fatorial exploratória, baseada na análise de componentes principais, é um método desenvolvido para gerar hipóteses, não testá-las (Fleck e cols, 1995). O uso de métodos de análise fatorial confirmatória e análise de Rasch (modelo logístico) (Custers e cols, 2000) poderiam reforçar e melhorar nossa compreensão acerca da validade de conteúdo da CES-D em nosso meio.

Tabela 1. Estudos fatoriais e de consistência interna da CES-D.

Autor	Amostra	Número de fatores mantidos	Sintomas depressivos e somáticos diferenciados	Consistência interna
Radloff (1977)	5144 indivíduos da comunidade	4	Sim	0.85
	70 pacientes psiquiátricos (EUA)			0.90
Roberts (1980)	705 imigrantes latinos (EUA)	4	Sim	0.87
Kuo (1984)	499 imigrantes asiáticos (EUA)	4	Não	---
Shrout e Yager (1989)	65 pacientes deprimidos (EUA)	4	Sim	0.91
	228 indivíduos da comunidade	4	Sim	0.84
Gatz (1993)	1663 idosos (Suécia)	4	Sim	---
Iwata e cols (1994)	1500 adolescentes escolares (Japão)	5	Sim	0.81
Silveira e Jorge (1998)	50 universitários e 50 farmacodependentes (Brasil)	4	Sim	0.85
Salle (1999)	503 adolescentes escolares (Brasil)	3	Sim	---
Ghubash e cols (2000)	350 mulheres jovens (Arábia Saudita)	3	Não	0.88
Fountoulakis e cols (2001)	40 pacientes deprimidos e 120 controles normais (Grécia)	3	Não	0.95
Foley e cols (2002)	227 negros idosos (EUA)	4	Não	0,86

Nota: com exceção de Iwata e cols (1994), que utilizou rotação Oblimin, os demais estudos listados utilizaram o método de análise de componentes principais com rotação Varimax.

Tabela 2. Fatores resultantes da análise de componentes principais após rotação Varimax com as cargas fatoriais das questões da CES-D

Questões	Fator 1	Fator 2	Fator 3
Afeto depressivo e somático/retardo motor			
06. Depressão	0.82		
18. Tristeza	0.78		0.32
03. Humor não reativo	0.76		
20. Desânimo	0.73		
01. Aborrecimento	0.72		
05. Baixa concentração	0.67		
14. Solidão	0.63		0.31
10. Medo	0.63		0.28
17. Crises de choro	0.63		0.27
07. Lassidão	0.63		
09. Sensação de fracasso	0.61		0.39
11. Sono agitado	0.58		
02. Falta de apetite	0.54		
13. Baixa comunicação	0.53		0.35
Afeto positivo			
12. Felicidade		0.84	
08. Esperança		0.79	
04. Bem-estar		0.75	
16. Prazer		0.70	
Interpessoal			
15. Pessoas não amigáveis	0.26		0.77
19. Rejeição	0.33		0.74
VE / VEA (%)¹	32.3	12.7 / 44.9	10.3 / 55.2

1. Variância explicada / variância explicada acumulada

Nota: os itens foram agrupados em ordem decrescente de intensidade de correlação (carga fatorial) com determinado fator; valores de carga fatorial acima de 0,5 (negrito) foram considerados significativamente determinantes para um fator; valores abaixo de 0,25 foram omitidos.

ARTIGO EM INGLÊS (2)

**RELIABILITY AND FACTOR STRUCTURE OF THE CES-D IN
PRIMARY CARE**

Gustavo Schestatsky, MD¹ and Marcelo P. A. Fleck, MD²

1) Psychiatrist. Student from the Post-Graduate Program of the Department of Psychiatry, Federal University of Rio Grande do Sul, Brazil.

2) Psychiatrist. Associate Professor of the Department of Psychiatry, Federal University of Rio Grande do Sul, Brazil.

ABSTRACT

Objective: To study the reliability and factorial structure of a CES-D Brazilian version in a population of primary care patients. **Method:** 436 individuals responded to the CES-D populational depression screening instrument. The reliability in terms of internal consistency (Cronbach alfa coeficient) and the factorial structure through principal component analysis were assessed. **Results:** The CES-D showed an internal consistency of 0,90. The factorial analysis extracted a three-factor solution: (1) depressive affect and somatic/motor retardation; (2) positive affect; (3) interpersonal. **Discussion:** The CES-D showed significant reliability for use in a primary care population. The factorial structure was similar to the Radloff original study (1977), however it clustered the items related to depressive affect as well as the somatic and motor retardation ones in a single factor. That factorial solution can be explained from cultural point of view, giving supporting to a significant CES-D content and construct validity in a primary care population in Brazil.

INTRODUCTION

The *Center for Epidemiologic Studies – Depression (CES-D) Scale* was developed to assess depressive symptoms in the general population (Raloff, 1977). It is composed by 20 items considered to be representative of the main depression symptoms.

The CES-D importance as a depression screening instrument has been consolidated after several studies intended to evaluate its psychometric features in various contexts. There are studies in non clinical populations (Silveira and Jorge, 1998; Ghubash *et al*, 2000), in clinical populations (Shrout and Yager, 1989; Fountoulakis *et al*, 2001), within different age ranges (Beekman *et al*, 1997; Salle, 1999), in diverse countries (Furukawa *et al*, 1997; Cho and Kim, 1998; Schroevers *et al*, 2000), and in very specific populations (Somervell *et al*, 1993; Wong, 2000; Foley *et al*, 2002).

As an instrument with potential application for the screening of depressive disorders, it has also being studied in populations of primary care users (Lyness *et al*, 1997; Irwin *et al*, 1999), where mood disorders are quite prevalent (Coyne *et al*, 1994).

Reliability and factor structure are the assessed aspects that contribute to confirm the CES-D construct and content validity. The CES-D reliability has been described based on its internal consistency, which is usually high (above 0,85 in most of the studies). The studies on factorial analysis have contributed to the CES-D content and construction validity since they extract the main domains which are part of the instrument and permit to make comparisons of its structure in different populations.

The original study by Radloff (1977) was the first to extract, by means of the principal component analysis method, a four-factor solution named as follows: (1) depressive affect; (2) positive affect; (3) somatic and motor retardation; (4) interpersonal. Later factorial studies, which also used the main components analysis confirmed that finding (Roberts, 1980; Kuo, 1984; Shrout and Yager, 1989; Gatz *et al*, 1993). Confirmatory factor analysis studies replicated the four-factor solution also demonstrating that there was a strong correlation among them. (Hertzog *et al*, 1990; Sheehan *et al*, 1995; Wong, 2000).

Other studies, however, indicated three-factor solutions, in general rendering “depressive affect” and Radloff’s “somatic and motor retardation” issues (1977) in one single factor (Ying, 1988; Guarnaccia *et al*, 1989; Ghubash *et al*, 2000; Fountoulakis *et al*, 2001). For this reason, the original four-factor structure of the CES-D may not be suitable for every population, as it can be influenced by sociocultural or language characteristics of different populations (Foley *et al*, 2002). Table 1 shows some factorial studies of the CES-D.

The present study aimed at studying the CES-D construct and content validity among users of a primary care population, in Brazil, using as parameters its reliability and factorial structure.

METHOD

Sample

Four hundred and thirty-six individuals who were attended in three primary care clinics from a delimited area of Porto Alegre, Brazil, were included in the study.

The data were obtained in two phases, constituting two additional samples taken from the same population. The first sample included individuals who had participated in the LIDO study, a longitudinal observational study of depressed patients in primary care (Fleck *et al*, 2002). In this phase 300 individuals whose CES-D scale score was 16 or higher (traditional cut point) were selected. The additional sample (second phase of data collection) included 136 participants whose score was lower than 16, in order to obtain the whole CES-D score spectrum distributed in that population (not only the ones with scores 16 or higher from the first collection phase).

Subjects aged 18 through 75 who lived within the nearby area of the clinics and were searching for care were invited to participate. Individuals who were unable to understand how to fill in the self-responder questionnaire, those who were in treatment for depression and those with prior diagnosis of any major psychiatric disorders were not included.

Data procedures

Subjects who searched for help in the mentioned clinics during the data collection procedures period and who fulfilled the inclusion or exclusion criteria were consecutively invited to participate in the study and signed an informed consent document. There was a team of interviewers covering most of the days of the week in the clinics. The subjects who agreed to participate in the study answered the CES-D self-responder screening instrument in the clinics' waiting-rooms. In the present study a Portuguese language version was used which was next translated and adapted by a bilingual panel under the supervision of a researcher from abroad.

Data analysis

The CES-D reliability was measured from the determination of its internal consistency through the Cronbach alpha coefficient. That coefficient measures the correlation degree among the scale items in order to check whether the same construct is being measured. The higher the scale items correlation, with views to measuring the same construct (internal consistency), the higher its reliability is (Silveira and Jorge, 1998).

The CES-D factorial structure analysis was performed by means of the principal component analysis technique (PCA). This technique provides the number of the scale components (factors), the explained variance of each component as well as the factor loadings of each item (Silveira and Jorge, 1998). In order to determine the number of components of the scale components, we considered valid those with eigenvalue higher

than the unit (Kaiser's criterion) (Fleck e Bourdel, 1998). In order to improve the potential for interpretation of the factors, we performed the Varimax orthogonal rotation method (Artes and Barroso, 1999). Only the items with factor loadings higher than 0,5 were included in the factors solution.

SPSS version 8.0 was used for data analysis.

RESULTS

The total sample was composed of 436 individuals. Mean age was 38,9, being females the predominant gender (76,8%). Most of the individuals were married (53,7%), following the singles (25,5%), separated (14,7%) and widowed (6,2%). As for schooling, the mean was 9,5 years of study. The CES-D mean scores was 22.

As for the internal consistency, the CES-D presented a 0,90 Cronbach alpha coefficient.

The APC before the ortogonal rotation demonstrated that 14 out of 20 items showed a significant correlation (above 0,6) with the principal factor (factor 1). After the Varimax rotation, we obtained a three-factor solution, which explained 55,2% of the total variability of the data. None of the items were excluded of the main components by the factor loading criterion of 0,5, wich suggests that all scale items were important for the factors composition.

The final composition of the three factors obtained by the APC, as well as the factor loadings of each item are demonstrated in Table 2. In that composition, the **Factor 1** (the most important in terms of the explained variance explanation) included not only items related to depressive affect itself (“blues”, “depresses”, “failure”, “fearful”, “lonely”, “crying”, “sad”) but also items related to somatic symptoms and motor retardation (“bothered”, “appetite”, “mind”, “effort”, “sleep”, “talk”, “get going”). The **Factor 2** grouped items related to a positive affect (“happy”, “hopeful”, “good”, “enjoy”), and the **Factor 3** had items related to the interpersonal sphere (“unfriendly”, “dislike”).

Taking as reference the original study by Radloff (1977) it was observed that out of the 20 questions from the original scale, 11 of them remained in identical factorial clusters to those found by that author: “blues”, “depressed”, “lonely”, “crying” and “sad” in **Factor 1**; “happy”, “hopeful”, “good” and “enjoy” in **Factor 2**; “unfriendly” and “dislike” in **Factor 3** (which corresponds to Radloff’s Factor 4). Surprisingly all items that in Radloff study included the somatic factor (“bothered”, “appetite”, “effort”, “sleep”, “get going”), together with other four which were not strongly associated to any factor in the original study (“concentration”, “failure”, “fearful”, “talk”) were grouped in factor 1, together with the items related to depressive affect.

In order to make a more straight comparison of our data with the Radloff’s original study (1977) and also with others which pointed out a 4-factor solution, we also assessed a solution with the same number of factors (data not showed). In such an hypothesis, the item 2, related to lack of appetite, would compose, isolatedly, the fourth factor. That fact

confirms that the solution of 3 factors is more adequate to our sample since it would not make sense, from a clinical point of view, a factor composed by only one item.

DISCUSSION

The 0,90 internal consistency (Cronbach's alpha) was compared, being generally higher, to the value of the majority of the studies (Table 1). We consider this an indicator that the CES-D maintained its reliability for its use in this population of primary care users. The high internal consistency suggests an unidimensional performance of the CES-D in this population which is probably related to the construct of depression.

As for the number of factors extrated by the APC (three factors), our study differs from Radloff's original study (1977) e from other ones which have found four-factor solutions in North American and European populations (Schrouf and Yager, 1989; Gatz *et al*, 1993).

On the other hand, our three-fator solution, combining items related to depressive affect and somatic/motor retardation is in accordance to studies in populations different from North American and European ones (Ying, 1988; Guarnaccia *et al*, 1989; Ghubash *et al*, 2000; Fountoulakis *et al*, 2001). What calls special attention is that the populations included in these studies come from ethnic minorities, especially from eastern cultures or from countries such as Saudi Arabia and Greece. As for the last two ones, they are

populations with characteristics in theory similar to ours in regard to lower socioeconomical development.

We consider that this finding may be explained from a cultural perspective. Non-western and/or with lower socioeconomic development populations tend to somatize their psychological suffering (Kuo, 1984; Parker *et al*, 2001). In contrast, it is a belief that individuals from high developed cultures such as the United States and Western Europe, have acquired higher capacity for verbalizing affective states, thus making less use of somatization (Katon *e cols*, 1982).

This may be particularly true in the case of primary care users. Not only is the prevalence of depressive disorders very high in this setting (Coyne *et al*, 1994) but also these individuals usually search for help exactly because of their somatic complains which are often linked to depressive symptoms (Fava, 2002).

Lastly, we consider an important finding that factor 1 was formed by items related to depressive affect and to somatic and motor retardation symptoms. That composition is accordingly with other acclaimed instruments as the Hamilton scale (1960) that have also identified a depressive or melancholic factor formed not only by depressive humor but also by motor retardation. In addition, the items included in factor 1 make sense from the clinical point of view since they are exactly the so-called “nuclear” symptoms of depression.

Thus we may conclude that the CES-D composition really represents the principal symptoms of depression, what reinforces its content and construct validity in a Brazilian primary care population.

It is worth reminding that the factorial analysis based on the principal components analysis is a basic method developed to generate hypothesis, not to test them (Fleck *et al*, 1995). The use of confirmatory factor analysis methods as well Rasch’s analysis (logistic model) (Custers *et al*, 2000) could reinforce and improve our comprehension of the CES-D content validity in our culture.

References

Artes R, Barroso LP. Aspectos estatísticos da análise fatorial de escalas de avaliação. *Rev Psiquiatr Clín* 25 (5): 223-228, 1998.

Beekman ATF, Deeg DJH, Van Limbeek J, e cols. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med* 27 (1): 231-235, 1997.

Cho MJ, Kim AKH. Use of the Center for Epidemiologic Studies Depression (CES-D) Scale in Korea. *J Nerv Ment Dis* 186: 304-310, 1998.

Coyne JC, Fechner-Bates S, Schwenk TL. Prevalence, nature and comorbidity of depressive disorders in primary care. *Gen Hosp Psychiatry* 16: 267-276, 1994.

Custers JWH, Hoijtink H, Net J, Helders PJM. Cultural differences in functional status measurement: analyses of person fit according to the Rasch model. *Quality of Life Research* 9: 571-578, 2000.

Fava M. Somatic symptoms, depression, and antidepressant treatment. *J Clin Psychiatry* 63 (4): 305-307, 2002.

Fleck MPA, Poirier-Littre MF, Guelfi J, e cols. Factorial structure of the 17-item Hamilton Depression Rating Scale. *Acta Psychiatr Scand* 92:168-172, 1995.

Fleck MPA. Estudo transcultural de avaliação da sintomatologia depressiva em pacientes hospitalizados. Dissertação de Doutorado em Clínica Médica/UFRGS. Porto Alegre, 1997.

Fleck MPA, Bourdel MC. Método de simulação e escolha de fatores na análise dos principais componentes. *Rev Saúde Pública* 32 (3): 267-272, 1998.

Foley KL, Reed PS, Mutran EJ, De Vems RF. Measurement adequacy of the CES-D among a sample of older African-Americans. *Psychiatry Res* 109: 61-69, 2002.

Fountoulakis K, Iacovides A, Kleanthous S, e cols. Reliability, validity and psychometric properties of the Greek translation of the Center for Epidemiological Studies – Depression (CES-D) Scale. *BMC Psychiatry* 1: 3-13, 2001.

Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. *J Affect Disord* 46 (1): 1-13, 1997.

Gatz M, Johansson B, Pedersen N, e cols. A cross-national self-report measure of depressive symptomatology. *International Psychogeriatrics* 5 (2): 147-156, 1993.

Ghubash R, Daradkeh TK, Al Naseri KS, e cols. The performance of the Center for Epidemiologic Study Depression Scale (CES-D) in an Arab female community. *International Journal of Social Psychiatry* 46 (4): 241-249, 2000.

Guarnaccia PJ, Angel R, Worobey JL. The factor structure of the CES-D in the Hispanic health and nutrition examination survey: the influences of ethnicity, gender and language. *Soc Sci Medicine* 29: 85-94, 1989.

Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* 26: 56-62, 1960.

Hertzog C, Alstine JV, Usala PD, e cols. Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychological Assessment* 2 (1): 64-72, 1990.

Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). *Arch Intern Med* 159:1701-1704, 1999.

Katon W, Kleinman A, Rosen G. Depression and somatization: a review. *The American Journal of Medicine* 72: 127-135, 1982.

Kuo WH. Prevalence of depression among Asian-Americans. *J Nerv Ment Dis* 172 (8): 449-457, 1984.

Lyness JM, Noel TK, Cox C, e cols. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Dpression Scale. *Arch Intern Med* 157 (4): 449-454, 1997.

Parker G, Gladstone G, Chee KT. Depression in the planet's largest ethnic group: the Chinese. *Am J Psychiatry* 158 (6): 857-864, 2001.

Radloff LS. The CES-D Scale: a self-report depression scale for research in general population. *Applied Psychol Measurement*, 1 (3), 385-401, 1997.

Roberts RE. Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Res* 2: 125-134, 1980.

Schroevers MJ, Sanderman R, Sonderen E, Ranchor AV. The evaluation of the Center for Epidemiologic Studies Depression (CES-D) Scale: depressed and positive affect in cancer patients and healthy reference subjects. *Quality of Life Research* 9:1015-1029, 2000.

Sheehan TJ, Fifield J, Reisine S, Tennen H. The measurement structure of the Center for Epidemiologic Studies Depression Scale. *Journal of Personality Assessment* 64 (3): 507-521, 1995.

Shrout PE, Yager TJ. Reliability and validity of screening scales: effect fo reducing scale length. *J Clin Epidemol* 42 (1): 69-78, 1989.

Silveira DX, Jorge MR. Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens. *Rev Psiq Clín* 25 (5) Edição Especial: 251-261, 1998.

Somervell PD, Beals J, Kinzie JD, e cols. Criterion validity of the Center for Epidemiologic Studies Depression Scale in a population sample from an American indian village. *Psychiatry Res* 47 (3): 255-66, 1993.

Wong YI. Measurement properties of the Center for Epidemiologic Studies – Depression Scale in a homeless population. *Psychological Assessment* 12 (1): 69-76, 2000.

Ying YW. Depressive symptomatology among Chinese-Americans as measured by the CES-D. *Journal of Clinical Psychology* 44: 739-746, 1988.

Table 1. Factor studies and internal consistency of the CES-D

Study	Sample	Number of factors	Depressive and somatic symptoms differentiated	Internal consistency
Radloff (1977)	5144 individuals from the community	4	Yes	0.85
	70 psychiatric patients (USA)			0.90
Roberts (1980)	705 Latin immigrants (USA)	4	Yes	0.87
Kuo (1984)	499 Asian immigrants (USA)	4	No	---
Shrout and Yager (1989)	65 depressed patients (USA)	4	Yes	0.91
	228 individuals from the community	4	Yes	0.84
Gatz <i>et al</i> (1993)	1663 elderly (Sweden)	4	Yes	---
Iwata <i>et al</i> (1994)	1500 school adolescents (Japan)	5	Yes	0.81
Silveira and Jorge (1998)	50 college students and 50 addicted patients (Brazil)	4	Yes	0.85
Salle (1999)	503 school adolescents (Brazil)	3	Yes	---
Ghubash <i>et al</i> (2000)	350 young women (Saudi Arabia)	3	No	0.88
Fountoulakis <i>et al</i> (2001)	40 depressed patients and 120 normal controls (Greece)	3	No	0.95
Foley <i>et al</i> (2002)	227 elderly black (USA)	4	No	0.86

Note: with exception of Iwata *et al* (1994), who used the Oblimin rotation, the other studies used the principal component analysis method with Varimax rotation.

Table 2. Factors resulting from the principal component analysis after Varimax rotation and factorial loading of the CES-D items

Items	Factor 1	Factor 2	Factor 3
Depressive affect and somatic/motor retardation			
06. Depressed	0.82		
18. Sad	0.78		0.32
03. Blues	0.76		
20. Get going	0.73		
01. Bothered	0.72		
05. Mind	0.67		
14. Lonely	0.63		0.31
10. Fearful	0.63		0.28
17. Crying	0.63		0.27
07. Effort	0.63		
09. Failure	0.61		0.39
11. Sleep	0.58		
02. Appetite	0.54		
13. Talk	0.53		0.35
Positive affect			
12. Happy		0.84	
08. Hopeful		0.79	
04. Good		0.75	
16. Enjoy		0.70	
Interpersonal			
15. Unfriendly	0.26		0.77
19. Dislike	0.33		0.74
EV / CEV (%)¹	32.3	12.7 / 44.9	10.3 / 55.2

1. Explained variance / cumulative explained variance.

Note: Items were clustered in a decreasing order of correlation intensity (factor loading) with the given factor; factor loadings higher than 0,5 were considered to be significant determinants for the factor; factor loadings lower than 0,25 were omitted.

ANEXO

CENTER FOR EPIDEMIOLOGIC STUDIES – DEPRESSION SCALE

(CES-D)

Instruções: Abaixo há uma lista de sentimentos e comportamentos. Por favor, assinale a frequência com que lhe ocorreram na última semana.

Raramente ou nunca (menos que 1 dia)

Poucas vezes (1-2 dias)

Às vezes (3-4 dias)

Quase sempre ou sempre (5-7 dias)

Durante a última semana:

1. Eu me chateei por coisas que normalmente não me chateavam.
 2. Não tive vontade de comer; estava sem apetite.
 3. Sinto que não consegui me livrar da tristeza mesmo com a ajuda da minha família ou dos meus amigos.
 4. Eu me senti tão bem quanto as outras pessoas.
 5. Eu tive problemas para manter a concentração (prestar atenção) no que estava fazendo.
 6. Eu me senti deprimido.
 7. Sinto que tudo que eu fiz foi muito custoso.
 8. Eu me senti com esperança em relação ao futuro.
 9. Eu pensei que minha vida tem sido um fracasso.
 10. Eu me senti com medo.
 11. Meu sono esteve agitado.
 12. Eu estive feliz.
 13. Eu conversei menos que o meu normal.
 14. Eu me senti sozinho.
 15. As pessoas não foram amigáveis.
 16. Eu me diverti.
 17. Eu tive crises de choro.
 18. Eu me senti triste.
 19. Eu senti que as pessoas não gostam de mim.
 20. Eu me sinto desanimado.
-