

**AVALIAÇÃO DE UM INSTRUMENTO DE MENSURAÇÃO DE MORBIDADE PSÍQUICA:
ESTUDO DE VALIDAÇÃO DO SELF-REPORTING QUESTIONNAIRE (SRQ-20)**

Kionna Oliveira Bernardes Santos^a

Tânia Maria de Araújo^a

Paloma de Sousa Pinho^{a,b}

Ana Cláudia Conceição Silva^{a,c}

Resumo

O *Self-Reporting Questionnaire* (SRQ-20), desenvolvido pela Organização Mundial de Saúde, tem sido utilizado para mensuração de nível de suspeição de transtornos mentais em estudos brasileiros, especialmente em grupos de trabalhadores. O objetivo deste estudo foi avaliar o desempenho do SRQ-20, com base em indicadores de validade (sensibilidade, especificidade, taxa de classificação incorreta e valores preditivos), e determinar o melhor ponto de corte para classificação dos transtornos mentais comuns na população estudada. O estudo incluiu 91 indivíduos selecionados aleatoriamente de um estudo de corte transversal realizado com população residente em áreas urbanas de Feira de Santana (BA). Entrevistas clínicas, realizadas por psicólogas, utilizando o *Revised Clinical Interview Schedule* (CIS-R), foi adotada como padrão-ouro. Na avaliação do desempenho do SRQ-20 foram estimados indicadores de validade (sensibilidade e especificidade). A curva *Receiver Operator Characteristic Curve* (ROC) foi utilizada para determinar o melhor ponto de corte para classificação de suspeitos/não suspeitos. O ponto de corte de melhor desempenho foi de 6/7 para a população investigada, revelando desempenho razoável com área sob a curva de 0,789. Os resultados indicam que o SRQ-20 possui característica discriminante regular.

Palavras-chave: SRQ-20. Indicadores de validade. Transtornos mentais. Sensibilidade. Especificidade.

^a Núcleo de Epidemiologia. Departamento de Saúde. Universidade Estadual de Feira de Santana.

^b Universidade Federal do Recôncavo da Bahia.

^c Universidade Estadual do Sudoeste da Bahia.

Endereço para correspondência: Av. Transnordestina, s/n, Departamento de Saúde, Núcleo de Epidemiologia, Sexto Módulo, Feira de Santana, Bahia, Brasil. CEP: 44031-460. araujo.tania@uefs.br

EVALUATION OF AN INSTRUMENT FOR MEASURING PSYCHIATRIC MORBIDITY: A VALIDITY STUDY OF THE SELF-REPORTING QUESTIONNAIRE (SRQ-20)

Abstract

The *Self-Reporting Questionnaire (SRQ-20)*, developed by the World Health Organization, has been used for measuring mental disorder in Brazilian studies, specially among occupational groups. This study aimed to evaluate the performance of the SRQ-20 for measuring common mental disorders (CMD) and to determine the best cutoff point to classify CMD case/ no case CMD. The study included 91 individuals randomly selected from a cross-sectional study carried out in a population living in urban areas in Feira de Santana, Bahia. Clinical interviews, conducted by psychologist, using the Revised Clinical Interview Schedule (CIS-R), were adopted as the gold standard. Sensibility and specificity indicators were estimated to evaluate the SRQ-20 performance. The ROC curve (*Receiver Operator Characteristic Curve*) was used to determine the best cutoff point. The cutoff point of best performance was at 6/7 for the population investigated, showing moderate performance with the area under the curve at 0,789 indicating that the SRQ-20 has a regular discriminating characteristic.

Key words: SRQ-20. Validity indicators. Mental Disorders. Sensibility. Specificity.

EVALUACIÓN DE UN INSTRUMENTO PARA LA MEDICIÓN DE MORBILIDAD PSÍQUICA: ESTUDIO DE VALIDACIÓN DEL SELF-REPORTING QUESTIONNAIRE (SRQ-20)

Resumen

O *Self-Reporting Questionnaire (SRQ-20)*, desarrollado por la Organización Mundial de la Salud, se ha utilizado, en los estudios en Brasil, para medir el nivel de sospecha de trastornos mentales, especialmente en los grupos de trabajadores. El objetivo de este estudio fue evaluar el desempeño del SRQ-20 sobre la base de indicadores de validez (sensibilidad, especificidad, la tasa de errores de clasificación, y los valores predictivos), y determinar el mejor punto de corte para la clasificación de los trastornos mentales comunes en la población. En el estudio participaron 91 sujetos seleccionados al azar de un estudio transversal realizado con personas que viven en las zonas urbanas de Feira de Santana (BA). Entrevistas clínicas, llevada a cabo por psicólogos, utilizando la versión revisada del *Clinical Interview Schedule (CIS-R)*, fue adoptado como el estándar de oro. Al evaluar el desempeño de SRQ-20 se estimaron indicadores de validez (sensibilidad y especificidad). Receiver Operator Characteristic Curve (ROC) se utilizó para determinar el mejor punto de corte para la clasificación de sospecha / no sospecha. El punto de corte para un mejor rendimiento fue de 6 / 7 para la población investigada, mostrando un rendimiento razonable con un área bajo la curva de 0,789. Los resultados indican que el SRQ-20 tiene característica regular discriminante.

Palabras-clave: Los indicadores de validez. Trastornos Mentales. Sensibilidad. Especificidad.

INTRODUÇÃO

Com o objetivo de avaliar os transtornos mentais em países em desenvolvimento, a Organização Mundial de Saúde (OMS) patrocinou a elaboração do instrumento de pesquisa *Self-Reporting Questionnaire* (SRQ). A preocupação com os impactos que os problemas de saúde mental poderiam ter em países periféricos e a necessidade de uma adequada avaliação desses transtornos mentais, considerando o contexto cultural desses países, orientaram o processo de construção do SRQ.¹

O SRQ derivou de instrumentos de *screening* para morbidade psíquica utilizados em pesquisa psiquiátrica:² *Patient Symptom Self-Report* (PASSR), um instrumento desenvolvido na Colômbia; *Post Graduate Institute Health Questionnaire N2*, desenvolvido na Índia; *General Health Questionnaire*, na sua versão de 60 itens, usado em países desenvolvidos e em desenvolvimento; e os itens de “sintomas” da versão reduzida do *Present State Examination* (PSE).

Em sua versão original, o SRQ incluía 24 itens, sendo os primeiros 20 itens para triagem de distúrbios não psicóticos e os quatro últimos para detecção de distúrbios psicóticos. O SRQ é um instrumento autoaplicável, contendo escala dicotômica (sim/não) para cada uma das suas questões.

O SRQ destina-se à detecção de *sintomas*, ou seja, sugere nível de suspeição (presença/ ausência) de algum transtorno mental, mas não discrimina um diagnóstico específico; assim, avalia se há algum transtorno, mas não oferece diagnóstico do tipo de transtorno existente. Por este caráter de triagem, é bastante adequado para estudos de populações, sendo muito útil para uma primeira classificação de possíveis casos e não casos.

A versão em português do SRQ adotou os 20 primeiros itens para investigar morbidade não psicótica. São considerados aspectos positivos na utilização do SRQ-20 o fato de ser de fácil compreensão, de rápida aplicação, diminuindo os custos operacionais, e ser um instrumento padronizado internacionalmente, alcançando níveis de desempenho aceitáveis no tocante à sensibilidade, especificidade e valores preditivos.³

Os sintomas neuróticos avaliados pela versão de 20 itens do SRQ (o SRQ-20) aproximam-se dos transtornos mentais comuns (TMC),⁴ que se caracterizam por sintomas não psicóticos, como insônia, fadiga, irritabilidade, esquecimento, dificuldade de concentração e queixas somáticas.

Na literatura internacional há vários estudos de avaliação do desempenho do SRQ. Em geral, os estudos apontam desempenho satisfatório deste instrumento.⁵⁻¹⁰ Além da análise de indicadores de validade e de confiabilidade, têm sido também analisados os fatores capazes de interferir no desempenho do SRQ.

A situação conjugal influenciou erros de classificação do SRQ; divorciados, separados e viúvos foram mais propensos a serem classificados como falsos positivos.¹¹ Entretanto, o fator de maior influência para a classificação errônea foi o nível de instrução (escolaridade). Pessoas com nível educacional mais baixo eram mais plausíveis de serem classificadas erroneamente como falsos positivos e pessoas com nível educacional mais elevado como falsos negativos. A questão de gênero também foi importante: a taxa de falsos negativos foi mais elevada entre homens do que entre mulheres.¹¹

O SRQ-20 tem sido intensivamente usado no Brasil nas últimas duas décadas, contudo poucos estudos preocuparam-se em avaliar o desempenho deste instrumento, de modo a precisar o quanto a mensuração realizada foi, de fato, satisfatória ou não.^{1,12-14}

Este estudo objetivou avaliar o desempenho do SRQ-20 com base em indicadores de validade (sensibilidade, especificidade, taxa de classificação incorreta e valores preditivos) e determinar o melhor ponto de corte para classificação dos transtornos mentais comuns na população estudada. Adicionalmente, o estudo buscou examinar possíveis fatores de influência no desempenho do SRQ-20. Foram avaliados fatores sociodemográficos (gênero e escolaridade).

INDICADORES DE VALIDADE DE INSTRUMENTOS DE MENSURAÇÃO: ASPECTOS GERAIS RELEVANTES PARA A ANÁLISE DE DESEMPENHO

Quando se aplica um instrumento para mensuração de um evento de ampla abrangência e sem dimensões físicas a serem avaliadas, podem ser utilizadas estratégias metodológicas baseadas em questionários e entrevistas, que são estruturados para investigar uma determinada realidade com o apoio de questões agrupadas cujo objetivo é aferir fenômenos específicos. A avaliação criteriosa do desempenho dos instrumentos de mensuração, ou seja, a validação desse instrumento para verificar se a mensuração realizada mede consistentemente o que o instrumento propõe-se a medir (validade) deve ser realizada, tornando-se uma condição importante para o seu uso.¹⁵

Diversos indicadores são utilizados na análise do desempenho de um instrumento, dentre os quais se podem destacar a validade de critério e a validade discriminante.

A validade de critério descreve uma relação entre a medida realizada e um determinado instrumento de uso padrão com critérios confiáveis (*gold standard*). Pode ser classificada em dois tipos: concorrente – que demonstra a aplicação de dois instrumentos simultaneamente, numa mesma população e a posterior análise dos resultados obtidos; e preditiva – que se relaciona com a predição de um evento no futuro.¹⁶

Outro tipo de análise de desempenho utilizada é a validade discriminante, em que o produto dos escores obtidos pelo instrumento classifica os sujeitos pesquisados, sendo comum a presença de erros do tipo “falso positivo” e “falso negativo”. No erro do tipo “falso negativo”, é atestada a incapacidade de identificar um evento em quem de fato o apresenta; já no erro do tipo “falso positivo”, é atestado o fenômeno em indivíduos que, de fato, não o apresentam.¹⁵

Dessa forma, os conceitos de sensibilidade e especificidade do instrumento são incluídos neste tipo de validação, em que a sensibilidade confere a capacidade que o instrumento tem em identificar os reais fenômenos apresentados por indivíduos que os têm, e a especificidade apresenta a capacidade em discriminar indivíduos que não possuem o fenômeno estudado.

Quando o instrumento avalia o evento de interesse em uma escala contínua, é necessário o estabelecimento de um ponto de corte, abaixo e acima do qual seja possível definir-se ausência ou presença de um dado desfecho, respectivamente.

Para definir um bom ponto de corte, deve-se levar em conta os dados que mostram os resultados dos testes e a condição de cada indivíduo em um determinado estudo.¹⁷ A análise da curva ROC (*Receiver Operating Characteristic*) é um procedimento utilizado na análise do ponto de corte de melhor desempenho. A curva pode ser traçada calculando-se a sensibilidade e especificidade para diversos pontos de corte da variável que contém os resultados do teste. No gráfico da curva ROC, o eixo y corresponde à sensibilidade e o eixo x à taxa de erro falso-positivo, ou seja, ao complemento da especificidade. A área sob a curva ROC representa a capacidade do teste de discriminar corretamente os indivíduos que têm a característica que está sendo medida daqueles que não a têm. Quanto maior a área, maior a capacidade de discriminação do teste. Um bom critério para a determinação do ponto de corte “ótimo” é considerar aquele que maximiza, ao mesmo tempo, a sensibilidade e a especificidade. Isto pode ser observado num gráfico de sensibilidade *versus* especificidade do teste considerando-se vários pontos de corte.

A análise do desempenho de um instrumento deve, portanto, levar em consideração um conjunto de indicadores, analisando-se a sua capacidade de discriminar caso/não caso, a definição do melhor ponto de corte, bem como fatores que podem interferir no seu desempenho em contextos específicos.

Indicadores de desempenho do SRQ-20

Na década de 1990, o SRQ-20 destacou-se como um dos instrumentos de triagem de morbidade psíquica mais utilizado em países desenvolvidos e em desenvolvimento.²

No que se refere aos pontos de corte para classificação de suspeitos, destaca-se a existência de variação em pontos de corte adotados para suspeição diagnóstica da morbidade psíquica, podendo variar entre 3/4 (até 3 respostas positivas=não caso/ 4 ou mais questões positivas=caso) e 11/12 (até 11 respostas positivas=não caso/ 12 ou mais respostas positivas=caso). A mensuração dos sintomas realizada pelo SRQ é feita por agrupamento de itens, o que pode facilitar fatores de classificação errônea para esse instrumento.¹⁸

A avaliação do desempenho do SRQ foi conduzida em vários estudos. O **Quadro 1** apresenta os indicadores dos estudos de validação realizados no Brasil e em outros países. Como se pode observar, os achados das pesquisas, em geral, apontaram um desempenho razoável do instrumento. A variabilidade da especificidade em pesquisas internacionais foi de 44%¹⁹ a 95,2%,²⁰ enquanto entre estudos no Brasil foi de 74,6%¹³ a 89,3%;²¹ já a sensibilidade variou de 73%⁵ a 89,7%,²⁰ e de 57%²² a 86,3%,²¹ respectivamente para pesquisas fora do país e no país.

Autores/ano	Ponto de corte*	População estudada	Sensibilidade (%)	Especificidade (%)
Harding et al. ⁵ - 1980**	5-11	Usuários de serviços de atenção primária	73- 83	72-85
Dhadphale et al. ²⁰ - 1982	7/8	Pacientes de hospital geral	89,7	95,2
Sen et al. ²⁹ - 1987	11/12	Usuários de serviços de atenção primária	79,0	75,0
Kortmann e Ten Horn ¹⁹ - 1988	8/9	Usuários de clínica psiquiátrica	77,0	44,0
Penayo ⁷ - 1990	9/10	População geral Usuários de atenção primária	81,0	58,0
Araya, Wynn e Lewis ¹¹ - 1992	9/10	Usuários de serviços de atenção primária	77,0	76,0
El-Rufaie e Absood ⁸ - 1994	5/6	Usuários de serviços de atenção primária	78,3	75,2
Glubash et al. ³⁴ - 2001	3/4	População geral (mulheres)	83,0	80,0

Quadro 1. Ponto de Corte, sensibilidade e especificidade encontrados em estudos de validação do SRQ-20

(Continua)

(Conclusão)

Autores/ano	Ponto de corte*	População estudada	Sensibilidade (%)	Especificidade (%)
Husain et al. ⁹ - 2004	8/9	Pacientes de um ambulatório	80,0	85,4
Giang et al. ¹⁰ - 2006	5/6	Usuários hospital	85,0	46,0
	6/7	População Geral	85,0	61,0
Brasil				
Mari e Williams ²⁷ - 1985	7/8	Usuários de serviços de atenção primária	83,0	80,0
Fernandes e Almeida Filho ²² - 1998	6/7	Trabalhadores em informática	57,0	79,0
Palácios et al. ²² - 1998	4/5	Bancários	75,6	78,4
	5/6	Bancárias	77,8	79,7
Ludermir e Melo Filho ¹² - 2002	5/6	População geral	62,0	78,0
Gonçalves et al. ²¹ - 2008	7/8	População geral	86,3	89,3
Scazufca et al. ¹³ - 2009	4/5	População de idosos	76,1	74,6

* Número de respostas positivas às questões do SRQ e expressam pontos para classificação de não caso/caso: valor utilizado para classificar suspeição de não caso/ valor utilizado para suspeição de caso.

** Neste estudo foram investigadas populações distintas (na Colômbia, Índia, Sudão e Filipinas). Os pontos de corte de melhor desempenho variaram de 5 a 11 respostas positivas e os indicadores de validade segundo os dados expressos no Quadro.

Quadro 1. Ponto de Corte, sensibilidade e especificidade encontrados em estudos de validação do SRQ-20

Alguns autores avaliaram o desempenho do SRQ-20 comparando-o com outros instrumentos também usados para mensuração de transtornos mentais comuns. Uma investigação em uma clínica de atenção primária em Santiago, Chile, comparou o desempenho do *General Health Questionnaire* (GHQ-12) e do *Self-Reporting Questionnaire* (SRQ-20), validados simultaneamente, tomando como padrão o *Clinical Interview Schedule* (CIS-R).¹¹ A curva ROC foi utilizada para determinar o melhor ponto de corte para definição de casos e comparar os indicadores de desempenho desses dois instrumentos. Os coeficientes de sensibilidade e especificidade obtidos nessa validação para o GHQ-12 e o SRQ-20 respectivamente foram: especificidade 73% e 77% e sensibilidade 76% e 74%; a taxa de classificação incorreta foi estimada em 26% para o GHQ-12 e 25% para o SRQ-20. Esse estudo demonstrou um desempenho similar, dos dois instrumentos investigados, para identificar os transtornos mentais comuns (TMC) entre os usuários de serviços de atenção primária no Chile.

No Brasil, a avaliação do desempenho do SRQ-20 ainda é restrita. A avaliação do desempenho do SRQ-20 em uma população de usuários de serviços primários de saúde revelou que o instrumento é um bom indicador de morbidade psiquiátrica.¹ Estudo envolvendo empregados de 15 agências de um banco estatal no Rio de Janeiro, avaliou o desempenho do SRQ-20 considerando diferentes pontos de corte.²³ O estudo conduziu análise, em separado, para homens e mulheres. Os resultados apontaram ponto de corte com melhor equilíbrio entre sensibilidade e especificidade de cinco ou mais respostas positivas para os homens e de seis ou mais respostas positivas para as mulheres.

Outro estudo realizado no Brasil entre trabalhadores de informática encontrou resultados que revelaram ponto de corte de sete ou mais respostas positivas como melhor ponto para suspeição de TMC, apresentando, para este ponto, sensibilidade de 57% e especificidade de 79%.²²

Na avaliação do desempenho do SRQ-20 em um estudo de base populacional envolvendo uma amostra representativa da população de Olinda, composta por 620 sujeitos de 15 anos ou mais de idade, o ponto de corte de melhor desempenho para suspeição diagnóstica foi de 5/6 (não caso/caso) com indicadores de sensibilidade de 62% e especificidade de 78%.¹²

METODOLOGIA

A análise do desempenho do SRQ-20 foi feita com base em um estudo em amostra representativa da população urbana do município de Feira de Santana (BA) com idade de 15 anos ou mais, selecionada por amostragem aleatória por conglomerado, estratificada por subdistrito. O estudo incluiu duas etapas: uma etapa de triagem de morbidade psiquiátrica e uma segunda etapa de confirmação diagnóstica. Os resultados aqui descritos referem-se à segunda etapa.

A zona urbana do município de Feira de Santana, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE),²⁴ está subdividida em cinco subdistritos (Pampalona, Santo Antônio, Subaé, Mangabeira e Santana).

Para o estudo de triagem inicialmente foi realizado levantamento populacional de cada subdistrito e delimitação geográfica de cada área. Em seguida, por procedimento aleatório, foi feita a seleção dos setores censitários de cada subdistrito. Em cada setor censitário selecionaram-se as ruas a serem incluídas no estudo. Todos os domicílios das ruas sorteadas foram visitados e todos os indivíduos com 15 anos ou mais de idade foram considerados elegíveis para a pesquisa.

Na etapa de triagem, o tamanho da amostra foi definido considerando-se uma prevalência estimada de transtornos mentais menores de 25%, segundo estimativas da OMS,²⁵ erro amostral de 3%, com 95% de confiança. Assumindo-se esses parâmetros, estimou-se uma amostra de 800 indivíduos, acrescidos de 20% em função de perdas e recusas, totalizando uma amostra de 960 indivíduos.

A avaliação do desempenho do SRQ-20 foi realizada na segunda etapa de confirmação diagnóstica, conduzida com entrevistas clínicas. Para esta etapa, a capacidade operacional da equipe de pesquisa norteou a delimitação do total de 96 entrevistas (cerca de 10% do tamanho da amostra estabelecida para a etapa de triagem do estudo), numa proporção de 60% dos suspeitos de portar um transtorno e 40% dos não suspeitos de transtorno mental. Adotou-se procedimento cego para as entrevistas (os responsáveis pela avaliação clínica desconheciam o escore obtido no SRQ-20).

Definido o número total de entrevistas a ser realizado, procedeu-se à seleção da amostra com base nos escores obtidos no SRQ-20 na primeira etapa do estudo (suspeito de não ser um caso: até sete respostas positivas; de ser um caso: oito ou mais respostas positivas),^{1,21} sorteando-se uma subamostra de suspeitos de transtornos mentais comuns e uma subamostra de não suspeitos. Dos 96 selecionados para a entrevista, cinco indivíduos não foram encontrados. Assim, a subamostra estudada incluiu 91 indivíduos, participantes da etapa de triagem.

Nas entrevistas clínicas conduzidas por psicólogas, foi utilizado um instrumento de entrevista estruturada – o *Revised Clinical Interview Schedule* (CIS-R),²⁶ – por meio do qual foi feito o diagnóstico clínico. Os indivíduos foram diagnosticados como portadores de transtornos mentais comuns (caso) ou sem transtornos mentais (não casos).

Foram calculados os indicadores de sensibilidade e de especificidade. Análise da Curva ROC foi realizada para avaliação do melhor ponto de corte para classificação de suspeição de transtornos mentais. Também foram avaliados os melhores pontos de corte por gênero e grau de instrução. A área inferior da curva ROC foi estimada como indicador de maior habilidade discriminante para o teste, conforme indicado por Mari e Williams.¹ Esta área pode apresentar valores entre 0,5, traduzindo uma habilidade discriminativa inferior, cuja avaliação pode ser considerada ao acaso, e 1,0 significando perfeita predição. As significâncias estatísticas das áreas sob a curva ROC foram comparadas por seus respectivos intervalos de confiança. Considerou-se significância estatística de 5% ($p \leq 0,05$).

Para análise dos dados foram empregados o Programa Estatístico *Statistical Package for the Social Science* (SPSS) na versão 10.0 e o Programa Estatístico “R” *The R Foundation for Statistical Computing* versão 2.2.1.

Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos do Hospital São Rafael, sob protocolo 17/01, atendendo aos requisitos da Resolução Nº 196/96 do Conselho Nacional de Saúde.

RESULTADOS

Como já mencionado, no estudo de validação, foram incluídos 91 indivíduos. A média de idade apresentada por esta amostra foi de 43,05 ($\pm 20,7$) anos, sendo 68,4% dos entrevistados do sexo feminino e 31,6% do sexo masculino. Apenas 25,5% da amostra apresentavam nível baixo de escolaridade (até o ensino fundamental I incompleto – 1ª à 4ª série), enquanto 74,5% relataram um mínimo de cinco anos ou mais de estudo, ou seja, havia concluído o ensino fundamental I e II (1ª à 8ª), médio ou superior.

A amostra avaliada na etapa de confirmação diagnóstica, em comparação com a população geral estudada, manteve as mesmas proporções para gênero e mostrou-se ser mais velha e com maior nível de escolaridade.

O escore médio obtido pelo SRQ-20 foi de 6,46 respostas positivas com desvio padrão de 4,29. O diagnóstico clínico, obtido com base no CIS-R, classificou 41 indivíduos (45,1%) como normais e 50 indivíduos (54,9%) como portadores de algum grau de transtorno. No SRQ-20, 45 indivíduos foram classificados como normais (49,5%) e 46 indivíduos como suspeitos de transtornos mentais comuns (50,5%). Portanto, a prevalência de transtorno foi mais elevada no diagnóstico clínico do que na avaliação da triagem feita pelo SRQ-20.

Considerando pontos de corte entre 5 e 9, observou-se que a sensibilidade variou de 82,0% (ponto 5) a 46,0% (ponto 9) e a especificidade variou de 63% (ponto 5) a 85,0% (ponto 9). (**Tabela 1**).

Tabela 1. Estimativa de sensibilidade e especificidade para diferentes pontos de corte do SRQ-20 na população estudada – Município de Feira de Santana (BA) – 2007

Ponto de Corte	Sensibilidade (%)	Especificidade (%)
5	82,0	63,0
6	76,0	63,0
7	68,0	70,7
8	64,0	78,0
9	46,0	85,0

A estimativa do ponto de corte para o SRQ-20 foi verificada na curva ROC, utilizando a entrevista clínica como padrão-ouro (**Gráfico 1**). O ponto de corte de melhor desempenho obtido foi 6/7 com sensibilidade de 68% e especificidade de 70,7% (**Tabela 1**), taxas de falso positivo de 29,3% e falso negativo de 32%, o que estabelece um desempenho razoável para o teste. O valor preditivo positivo (VPP) para o melhor ponto de corte obtido (6/7) foi de 73,9%; o valor preditivo negativo (VPN) foi de 64,4%; e a taxa de classificação incorreta (TCI) foi de 30,7%.

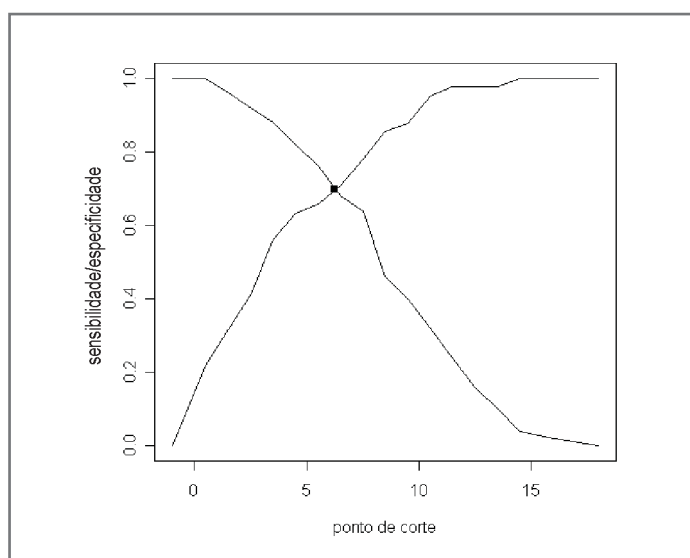


Gráfico 1. Estimativa de ponto de corte do SRQ-20 para a população estudada – Município de Feira de Santana (BA) – 2007

A área sob a curva ROC foi analisada apresentando um valor de 0,789 com um desvio padrão de 0,48 e intervalo de 95% de confiança de 0,696 a 0,882 apontando, assim, um nível razoável de discriminação entre casos e não casos (**Gráfico 2**).

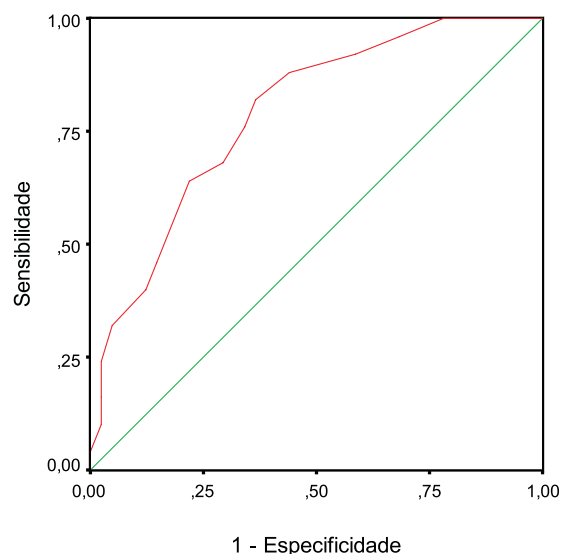


Gráfico 2. Relação entre sensibilidade e especificidade do SRQ-20 em diferentes pontos de cortes para a população estudada – Município de Feira de Santana (BA) – 2007

Foi realizada também a análise de pontos de corte do SRQ-20 por sexo e escolaridade. Foram observadas diferenças entre os melhores pontos para suspeição dos TMC entre homens e mulheres (**Tabela 2**). Para as mulheres foi estimado o ponto de corte “ótimo” entre sensibilidade (64,5%) e especificidade (64,5%), no escore 7, apresentando uma área abaixo da curva de 0,708, com desvio padrão de 0,069 e intervalo de confiança de 95% de 0,572 a 0,843. Entre os homens foi estimado um melhor ponto de corte no escore 5, apresentando sensibilidade de 80%, especificidade de 83,4% e uma área abaixo da curva de 0,919, estimada com imprecisão em virtude do pequeno tamanho da amostra masculina. Observa-se, assim, que o desempenho dos indicadores de validade do SRQ-20 variou segundo o sexo, sendo mais satisfatório para a avaliação de morbidade psíquica entre os homens (sensibilidade e especificidade mais elevadas).

Outro aspecto avaliado foi a interferência do grau de escolaridade referido pela população na estimativa do ponto de corte. Para os indivíduos que apresentaram escolaridade baixa (<5 anos) foi estabelecido o escore 8 como valor de referência para suspeição dos TMC, com uma sensibilidade de 58,0% e especificidade de 66,0%; para aqueles que apresentaram

um maior grau de instrução (≥ 5 anos) foi estimado, como melhor ponto, o escore 5 com sensibilidade de 76,8% e especificidade de 74% (**Tabela 2**).

Tabela 2. Indicadores de validade do SRQ-20 segundo características sociodemográficas para a população estudada – Município de Feira de Santana (BA) – 2007

Variável	Ponto de Corte	Sensibilidade (%)	Especificidade (%)	Área abaixo da Curva
Sexo				
Masculino	5	80,3	83,4	0,919
Feminino	7	64,5	64,5	0,708
Escolaridade				
Educação (< 5anos)	8	58,0	66,0	0,702
Educação (≥ 5 anos)	5	76,8	74,0	0,816

DISCUSSÃO

Os resultados apontaram para um desempenho satisfatório do SRQ-20, identificando particularidades entre características sociodemográficas, como sugerido anteriormente na literatura.^{22,23,27}

O ponto de corte estimado pela presente investigação (6/7) apresenta-se ligeiramente mais baixo que os encontrados em outros estudos.^{1,11,28-31}

A habilidade discriminativa determinada pela área sob a curva ROC para este estudo conferiu ao SRQ-20 uma razoável capacidade de detecção dos TMC, concordando com achados anteriores.¹¹

As taxas de falso negativo (29,3%) e falso positivo (32%) encontradas neste estudo representam dificuldade no processo de mensuração do SRQ-20, ao se considerar a natureza do objeto de investigação.

Deve-se considerar que os valores dos coeficientes de validade utilizados na avaliação do desempenho de instrumentos que mensuram distúrbios psiquiátricos devem ser ampliados, uma vez que a natureza do fenômeno investigado abrange múltiplas dimensões.³²

Entre os aspectos sociodemográficos investigados (sexo/escolaridade) foi identificado um melhor desempenho do SRQ-20 entre os indivíduos do sexo masculino e entre aqueles que possuíam maior escolaridade.

A literatura tem revelado que há diferença consistente entre os sexos na ocorrência dos transtornos mentais em todas as sociedades, e aponta alguns fatores que intensificam este quadro entre as mulheres: acesso à escola, violência intradomiciliar, oportunidades de emprego e sobrecarga doméstica.³³

Observou-se que o desempenho do SRQ-20 entre os indivíduos com baixa escolaridade, considerando o melhor ponto de corte, foi menos satisfatório que o estimado entre os indivíduos com maior nível de instrução. Este fato foi verificado em pesquisa anterior,³⁴ ao se observar que, entre populações de países periféricos que apresentam um baixo nível educacional, é comum a dificuldade de expressar desconfortos emocionais.

A relação entre o nível educacional baixo e os transtornos mentais pode ser confundida ou explicada por alguns fatores como a desnutrição na infância, fator que retarda o desenvolvimento intelectual e conduz a um baixo desempenho educacional e psicossocial no futuro. As consequências do baixo nível de instrução revelam a falta de oportunidade no contexto social e tem sido apontado como um fator de risco para a demência.³³

A variedade de pontos de corte encontrados na aplicação do SRQ-20 contradiz a hipótese de que exista um instrumento universal para avaliar os transtornos mentais, uma vez que, na investigação deste fenômeno, elementos culturais estão intimamente envolvidos, apresentando componentes verbais e não verbais para detecção de distúrbios emocionais;¹⁹ o presente estudo, porém, encontrou taxas aceitáveis de falsos positivos e falsos negativos na discriminação de casos e não casos.

Diante dos resultados apresentados neste estudo, com a observância da influência de fatores como idade e escolaridade no desempenho do SRQ-20, pode-se dizer que as questões relativas à validação de instrumentos de pesquisa conformam-se como requisitos importantes para a qualificação dos resultados produzidos, especialmente no que se refere à adequação dos instrumentos, de sua análise e interpretação na avaliação de populações específicas, com características sociais, econômicas e culturais próprias.

O baixo desempenho do SRQ-20 observado em populações de baixa escolaridade, por exemplo, alerta para o fato de que o seu uso pode apresentar dificuldades expressivas na avaliação de saúde mental em alguns grupos específicos, devendo-se considerar os resultados produzidos com cautela em populações com essa característica. Estudos futuros devem ser empreendidos para melhor compreensão dos fatores associados ao desempenho encontrado nesses grupos de modo a contribuir para a estruturação de instrumentos de medidas mais válidos e precisos.

Cabe lembrar, uma vez mais, que o SRQ-20 é um instrumento de triagem e destina-se à detecção de *sintomas*, sendo útil para a mensuração de nível de suspeição (presença/ausência) de transtorno mental. Assim, pode ter alcance relevante no campo da saúde pública, para as ações voltadas para os níveis primários de atenção, mas apresenta limitações evidentes para orientação da intervenção clínica.

Na análise geral do desempenho do SRQ-20, utilizado para avaliação de transtornos mentais em uma população residente em zona urbana, pode-se dizer que os aspectos psicométricos do instrumento, aqui avaliados, revelaram que, a despeito das limitações inerentes aos métodos de análise impostos às questões subjetivas como os transtornos mentais comuns, apresentou desempenho aceitável. Ou seja, não obstante os possíveis limites referentes ao pequeno número amostral (N=91) e da abrangência do construto avaliado (transtornos mentais), o desempenho do questionário foi satisfatório, identificando sintomas e classificando, em níveis aceitáveis, os indivíduos como suspeitos e não suspeitos de transtornos mentais comuns, sendo um instrumento útil para avaliação de morbidade psíquica em estudos epidemiológicos.

AGRADECIMENTOS

Agradecemos ao financiamento recebido da FAPESB (convênio nº 4183/2005) e CAPES (bolsa de mestrado) para a realização deste estudo.

REFERÊNCIAS

1. Mari JJ, Williams P. A validity study of a psychiatric screening questionnaire (SRQ-20) in primary care in the city of São Paulo. *Br. J. Psychiatry* 1986;148:23-6.
2. World Health Organization. Expert Committee on Mental Health: User's Guide to Self Reporting Questionnaire (SRQ). Geneva; 1994.
3. Borges LH, Medrado MA. Transtornos mentais menores entre trabalhadores de uma usina siderúrgica. *Rev. bras. Saúde Ocup.* 1993;21(77):7-18.
4. Goldberg D, Huxley P. Common mental disorders: a bio-social model. London: Tavistock; 1992.
5. Harding TW, Arango MV, Baltazar J, Climent CE, Ibrahim HHA, Ignacio LL, et al. Mental disorders in primary health care: a study of their frequency and diagnosis in four development countries. *Psychol. Medicine* 1980;10:231-41.
6. Aldana LL, Moreno LR, Carabantes, AD, Moscardó IB. Validacion del SRQ em los exámenes de salud mentale em la poblacion general. *Actas Luso-Esp. Neurol. Psiquiatr.* 1990;18(5):286-9.
7. Penayo U, Kullgren G, Caldera T. Mental disorders among primary health care patients in Nicaragua. *Acta Psychiatr. Scand.* 1990;82:82-5.
8. El-Rufaie OEF, Absood GH. Validity study of the Self-Reporting Questionnaire (SRQ-20) em Primary Health Care in the United Arab Emirades. *Int. J. Met. Psychiatr. Res.* 1994;4:45-53.

9. Husain N, Gater R, Tomenson B, Creed F. Social factors associated with chronic depression among a population-based sample of women in rural Pakistan. *Soc. Psychiatr. Psychiatr. Epidemiol.* 2004;39:618-24.
10. Giang KB, Allebeck P, Kullgren G, Tuan NV. The vietnamese version of the Self-Reporting Questionnaire 20 (SRQ-20) in detecting mental disorders in rural Vietnam: a validation study. *Int. J. Soc. Psychiatr.* 2006;52(2):175-84.
11. Araya R, Wynn R, Lewis G. Comparision of two self administered psychiatric questionnaires (GHQ-12 and SRQ-20) in primary care in Chile. *Social Psychiatr. Psychiatr. Epidemiol.* 1992;27(4):168-73.
12. Ludermir AB, Melo-Filho DA. Condições de vida e estrutura ocupacional associadas a transtornos mentais comuns. *Rev. Saúde Públ.* 2002;36:213-21.
13. Scazufca M, Menezes PR, Vallada H, Araya R. Validity of the self reporting questionnaire-20 in epidemiological studies with older adults: Results from the Sao Paulo Ageing & Health Study. *Social Psychiatr. Psychiatr. Epidemiol.* 2009;44:247-54.
14. Santos KOB, Araújo TM, Oliveira NE. Estrutura fatorial e consistência interna do Self-Reporting Questionnaire (SRQ-20) em população urbana. *Cad. Saúde Públ.* 2009;25:214-22.
15. Fletcher R, Fletcher S, Wagner E. *Epidemiologia clínica: elementos essenciais*. 3ª. ed. Porto Alegre: Artes Médicas; 1996.
16. Richardson RJ, Peres JAS, Wanderley JCV, Correia LM, Peres MHM. *Pesquisa social: métodos e técnicas*. 3ª. ed. São Paulo: Atlas; 1999.
17. Jekel JF, Elmore JG, Katz DL. Entendendo os erros em Medicina Clínica. In: Jekel JF, Elmore JG, Katz DL, editor. *Epidemiologia, Bioestatística e Medicina Preventiva*. 2ª ed. Porto Alegre: Artmed; 1999. p.114-28.
18. Kortmann F. Problems in communication in transcultural psychiatry. The self reporting questionnaire in Ethiopia. *Acta Psychiatr. Scand.* 1987;75(6):563-70.
19. Kortmann F, Ten Horn S. Comprehension and motivation in responses to a psychiatric screening instrument - Validity of SRQ in Ethiopia. *British J. Psychiatr.* 1988;153:95-101.
20. Dhadphale M, Ellison RH, Griffin L. Frequency of mental disorders among outpatients at a rural district hospital in Kenya. *Cent. Afr. J. Med.* 1982;28(4):85-9.
21. Gonçalves DM, Stein AT, Kapczinsk F. Performance of the Self-Reporting Questionnaire as a psychiatric screening questionnaire: a comparative study with Structured Clinical Interview for DSM-IV-TR. *Cad. Saúde Públ.* 2008;24(2):380-90.

22. Fernandes SRP, Almeida Filho N. Validação do SRQ-20 em amostra de trabalhadores de informática. *Rev. bras. Saúde Ocup.* 1998;89/90:105-12.
23. Palácios M, Jardim S, Ramos A, Silva Filho JF. Validação do Self-Report Questionnaire-20 (SRQ-20) numa população de trabalhadores de um banco estatal no Rio de Janeiro-Brasil. In: Silva Filho JF, Jardim S., editor. *A danação do trabalho – organização do trabalho e sofrimento psíquico*. Rio de Janeiro: Te Corá; 1998. p. 225-41.
24. Instituto Brasileiro de Geografia e Estatística. *Anuário estatístico de Feira de Santana*. Feira de Santana; 2000.
25. Organização Mundial de Saúde. *Relatório sobre a saúde no mundo: Saúde Mental: Nova Conceção, Nova esperança*. OPAS/OMS; 2001.
26. Lewis G, Pelosi AJ, Araya RI, Dunn G. Measuring psychiatric disorders in the community: the development of a standardized assessment for lay interviewers. *Psychol. Med.* 1992;22:465-86.
27. Mari JJ, Williams PA. A comparison of the validity of two psychiatric screening questionnaires (GHQ-12 and SRQ-20) in Brazil, using Relative Operating Characteristic (ROC) analysis. *Psychol. Med.* 1985;15:651-9.
28. Dhadphale R, Ellison RH, Griffin L. The frequency of psychiatric disorders among patients attending semi-urban and rural general out-patients clinics in Kenya. *British J. Psychiatr.* 1983;142:379-83.
29. Sen B, Wilkinson G, Mari JJ. Psychiatric morbidity in primary health care: a two-stage screening procedure in developing countries: choice of instruments and cost-effectiveness. *British J. Psychiatr.* 1987;151:33-8.
30. Deshpande SN, Sundaram KR, Wig NN. Psychiatric disorders among medical in patients in a Indian hospital. *British J. Psychiatr.* 1989;154:504-9.
31. Carta MG, Carpiniello B, Ciccone V, Sannais C, Paulis M, Rudas N. Standardization of a psychiatric screening test for use by general practitioners in Sardinia - preliminary results. *Acta Psychiatr. Scand.* 1993;87:342-4.
32. Iacoponi E, Mari JJ. Reliability and factor structure of the Portuguese version of Self-Reporting Questionnaire. *Intern. J. Soc. Psychiatr.* 1989;35(3):213-22.
33. Patel V, Kleinman A. Poverty and common mental disorders in developing countries. *Bull. World Health Organ.* 2003; 81(8):609-15.
34. Glubash R, Daradkeh T, El_Rufaie, OF, About-Saleh MT. A comparison of the validity of two psychiatric screening questionnaires: the Arabic General Health Questionnaire (AGHQ) and Self-Reporting Questionnaire (SRQ-20) in UAE, using Receiver Operating Characteristic (ROC) analysis. *Europ. Psychiatr.* 2001;16:122-6.

Recebido em 8.8.2009 e aprovado em 27.7.2010.