



## Versão brasileira do *Clinical COPD Questionnaire*, aplicado por meio de entrevista: propriedades de medida de confiabilidade e validade

Alexânia de Rê<sup>1,2</sup>, Fernanda Rodrigues Fonseca<sup>1,2</sup>, Ana Paula Queiroz<sup>1,2</sup>, Cardine Martins dos Reis<sup>1</sup>, Marina Mônica Bahl<sup>1</sup>, Janwillem Kocks<sup>3,4,5</sup>, Thys van der Molen<sup>6</sup>, Rosemeri Maurici<sup>1,2,7,8</sup>

1. Programa de Pós-Graduação em Ciências Médicas, Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.
2. Núcleo de Pesquisa em Asma e Inflamação das Vias Aéreas, Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.
3. General Practitioners Research Institute, Groningen, The Netherlands.
4. University of Groningen, University Medical Center Groningen, GRIAC Research Institute, Groningen, The Netherlands.
5. Observational and Pragmatic Research Institute, Singapore.
6. Department of General Practice, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.
7. Departamento de Clínica Médica, Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.
8. Hospital Universitário, Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.

Recebido: 25 julho 2020.  
Aprovado: 6 fevereiro 2021.

Trabalho realizado na Universidade Federal de Santa Catarina, Florianópolis (SC) Brasil.

### INTRODUÇÃO

As avaliações do impacto da doença e da estabilidade clínica em pacientes com DPOC devem ajudar os médicos a tomar decisões terapêuticas.<sup>(1)</sup> Medidas de desfechos relatados pelo paciente (PROMs, do inglês *patient-reported outcome measures*), como o *Clinical COPD Questionnaire* (CCQ) e o *COPD Assessment Test* (CAT), são úteis para a avaliação transversal do impacto da doença e para a avaliação longitudinal da estabilidade clínica.<sup>(1)</sup> De acordo com as recomendações da GOLD, o CCQ e o CAT são PROMs abrangentes e adequadas para a avaliação de sintomas em pacientes com DPOC.<sup>(2)</sup>

A maioria dos pacientes com DPOC prefere o CCQ ao CAT, pois o CCQ incorpora mais detalhes sobre problemas respiratórios diários do que o CAT e, portanto, reflete melhor seu estado de saúde. Alguns pacientes ressaltam

também que o CCQ apresenta um sistema de opções de resposta que é mais fácil de entender em comparação com o do CAT.<sup>(3)</sup> Além disso, o *International Primary Care Respiratory Group*<sup>(4)</sup> elegeu o CCQ como a melhor PROM para avaliar pacientes com DPOC na atenção primária. O CCQ foi a única PROM que recebeu nota máxima na pesquisa.

A seleção de uma PROM adequada para a avaliação do estado de saúde deve ser baseada na qualidade de suas propriedades de medida — uma PROM deve ser confiável e válida. Existem inúmeras PROMs que podem ser utilizadas para medir o estado de saúde. No entanto, uma seleção criteriosa é de extrema importância para evitar o risco de resultados imprecisos ou tendenciosos, os quais podem levar a conclusões errôneas.<sup>(5)</sup> As propriedades de medida podem diferir entre populações e, portanto,

### RESUMO

**Objetivo:** Testar a confiabilidade, validade e interpretabilidade da versão brasileira do *Clinical COPD Questionnaire* (CCQ) em pacientes com DPOC. **Métodos:** Cinquenta pacientes com DPOC preencheram o CCQ por meio de entrevista em duas ocasiões. Na primeira visita, o CCQ foi aplicado duas vezes, por dois avaliadores, com intervalo de aproximadamente 10 min; os pacientes também foram submetidos a espirometria e aplicação do *COPD Assessment Test*, da escala modificada do *Medical Research Council* e do *Saint George's Respiratory Questionnaire* (SGRQ). Na segunda visita (1-2 semanas depois), o CCQ foi reaplicado. Testamos a hipótese de que a pontuação total no CCQ se correlacionaria positivamente com a pontuação total e a pontuação nos domínios do SGRQ ( $r \geq 0,5$ ). **Resultados:** Dos 50 pacientes, 30 (60%) eram do sexo masculino. A média de idade foi de  $66 \pm 8$  anos, e a média do  $VEF_1$  foi de  $44,7 \pm 17,9\%$  do valor previsto. Para todos os itens do CCQ, o coeficiente alfa de Cronbach (IC95%) foi de 0,93 (0,91-0,96). Para analisar a confiabilidade interavaliadores e teste-reteste do CCQ, calculamos o coeficiente de correlação intraclasse de duas vias modelo de efeitos mistos para medidas únicas (0,97 [IC95%: 0,95-0,98] e 0,92 [IC95%: 0,86-0,95], respectivamente); erro-padrão de medida do tipo concordância (0,65 para ambas); a mínima mudança detectável individual (1,81 e 1,80, respectivamente) e no grupo (0,26 e 0,25, respectivamente); e os limites de concordância (-0,58 a 0,82 e -1,14 a 1,33, respectivamente). A pontuação total no CCQ correlacionou-se positivamente com todas as pontuações no SGRQ ( $r \geq 0,70$  para todas). **Conclusões:** A versão brasileira do CCQ apresentou erro de medida indeterminado, assim como confiabilidade interavaliadores/teste-reteste e validade de construto satisfatórias.

**Descritores:** Doença pulmonar obstrutiva crônica; Nível de saúde; Medidas de resultados relatados pelo paciente; Estudo de validação.

### Endereço para correspondência:

Alexânia de Rê. Núcleo de Pesquisa em Asma e Inflamação das Vias Aéreas, Universidade Federal de Santa Catarina, Hospital Universitário Polydoro Ernani de São Thiago, Campus Universitário, Rua Professora Maria Flora Pausewang, s/n, Trindade, CEP 88036-800, Florianópolis, SC, Brasil.

Tel./Fax: 55 48 3234-7711. E-mail: alexania\_dere@yahoo.com.br

Apoio financeiro: Alexânia de Rê foi bolsista de mestrado e Fernanda Rodrigues Fonseca foi bolsista de doutorado da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Rosemeri Maurici é bolsista de produtividade do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

devem ser testadas e consideradas adequadas para a população específica a ser avaliada.<sup>(6)</sup>

O CCQ foi desenvolvido por van der Molen et al.<sup>(7)</sup> em 2003 com o objetivo de promover a avaliação do controle clínico em pacientes com DPOC. Os domínios selecionados como os mais importantes para o controle clínico foram estado funcional, sintomas e estado mental.<sup>(7)</sup> Uma versão em português do CCQ para uso no Brasil está disponível no site do CCQ, mas suas propriedades de medida ainda não foram investigadas. O presente estudo objetivou analisar a consistência interna, confiabilidade, erro de medida e validade de construto, assim como os efeitos de piso e teto, da versão brasileira do CCQ, quando aplicada por meio de entrevista, em pacientes com DPOC.

## MÉTODOS

### Seleção de pacientes

Pacientes encaminhados a um ambulatório público especializado em DPOC foram convidados a participar do estudo. Os critérios de inclusão foram os seguintes: ter diagnóstico confirmado de DPOC, ter idade  $\geq 40$  anos, ser fumante ou ex-fumante, não apresentar outra comorbidade (como doenças cardiovasculares, neurológicas, ortopédicas, reumáticas ou respiratórias além da DPOC) que tivesse impacto negativo nas atividades da vida diária, e apresentar pontuação  $\geq 25$  (indivíduos alfabetizados) ou  $\geq 19$  (analfabetos) no Minixame do Estado Mental.<sup>(8)</sup> Os critérios de exclusão foram apresentar alterações na estabilidade clínica ou no impacto da doença no mês anterior ao estudo ou durante a coleta de dados, avaliadas por meio de perguntas fechadas, e não participar das avaliações do estudo. Todos os pacientes que aceitaram participar assinaram um termo de consentimento informado. Este estudo foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Santa Catarina (CAAE n. 33299214.8.0000.0121).

### Desenho do estudo

O estudo foi realizado em um ambulatório público especializado em DPOC em duas visitas no período da manhã entre 2017 e 2019. As PROMs selecionadas — CAT, CCQ, escala modificada *Medical Research Council* (mMRC) e *Saint George's Respiratory Questionnaire* (SGRQ) — foram aplicadas por meio de entrevistas em razão da baixa escolaridade de parte da amostra. As PROMs foram aplicadas em uma sala separada, com a presença apenas do paciente e do avaliador. Os avaliadores apenas leram as instruções e itens da PROM e anotaram as escolhas dos pacientes. Os avaliadores são fisioterapeutas com experiência na avaliação do estado de saúde em pacientes com DPOC. Na primeira visita, foi realizada espirometria, e foram aplicados o CAT, a escala do mMRC e o SGRQ. Além disso, o CCQ foi aplicado primeiro pelo avaliador 1 e então, após aproximadamente 10 min, pelo avaliador 2, para análise da confiabilidade interavaliadores. Na segunda visita, após uma a duas semanas, o CCQ foi

reaplicado pelo avaliador 1 para análise da confiabilidade teste-reteste.<sup>(9)</sup>

### Avaliações

A função pulmonar foi avaliada segundo os padrões recomendados pela *American Thoracic Society/European Respiratory Society*<sup>(10)</sup> utilizando um espirômetro (KoKo Sx 1000; nSpire Health Inc., Longmont, CO, EUA). Os valores de referência para variáveis espirométricas pós-broncodilatador foram os estabelecidos por Pereira et al.<sup>(11)</sup> A gravidade da limitação do fluxo aéreo foi baseada no  $VEF_1$  e classificada como GOLD I, II, III ou IV.<sup>(12)</sup>

As pontuações no CAT<sup>(13)</sup> e na escala mMRC,<sup>(14)</sup> assim como o número de exacerbações nos últimos 12 meses, independentemente das internações hospitalares, foram utilizadas para classificar o impacto da DPOC no estado de saúde e o risco de futuros eventos como GOLD A, B, C ou D.<sup>(12)</sup>

O SGRQ<sup>(15)</sup> foi utilizado para avaliar a qualidade de vida relacionada à saúde. O questionário é composto por 76 itens distribuídos em três domínios (sintomas, atividade e impacto psicossocial). A pontuação total varia de 0 a 100, sendo que pontuações mais altas significam pior qualidade de vida.

O CCQ<sup>(7)</sup> foi utilizado para avaliar o controle clínico. É composto por 10 itens distribuídos em três domínios (sintomas, estado mental e estado funcional). A pontuação total e a pontuação nos domínios variam de 0 a 6, sendo que pontuações mais altas significam pior controle.

### Análise estatística

A normalidade dos dados foi analisada por meio do teste de Shapiro-Wilk. O nível de significância estatística adotado foi de  $p < 0,05$ . A consistência interna dos itens do CCQ foi analisada por meio do coeficiente alfa de Cronbach ( $\alpha$ ) e do respectivo IC95%.<sup>(16)</sup> Para comparar as pontuações entre os avaliadores e entre o teste e reteste, utilizou-se o teste t de Student ou o teste de Wilcoxon, de acordo com a normalidade dos dados. A confiabilidade interavaliadores e teste-reteste das pontuações no CCQ foram analisadas por meio do coeficiente de correlação intraclass de duas vias/modelo de efeitos mistos para medidas únicas ( $CCI_{3,1}$ ) e o do respectivo IC95%.<sup>(17)</sup> Para as análises de confiabilidade interavaliadores, confiabilidade teste-reteste e erro-padrão de medida do tipo concordância ( $EPM_{concordância}$ ), calculamos a mínima mudança detectável individual ( $MMD_{individual}$ ) e no grupo ( $MMD_{grupo}$ ) e os limites de concordância (LC).<sup>(9)</sup> Para visualizar a pontuação total e a concordância entre as medidas do CCQ, foram utilizadas disposições gráficas de Bland-Altman.<sup>(18)</sup> Para analisar a validade de construto, as seguintes hipóteses foram utilizadas: a pontuação total no CCQ se correlacionaria com a pontuação total e a pontuação nos domínios do SGRQ, e o coeficiente de correlação ( $r$ ) seria  $\geq 0,5$ . A proporção de ocorrência da pontuação mínima e máxima no CCQ foi utilizada para analisar os efeitos de piso e teto, respectivamente, que foram classificados como ausentes ou presentes.<sup>(9)</sup>

## RESULTADOS

Cinquenta pacientes com DPOC foram incluídos no estudo. As características gerais da amostra são descritas na Tabela 1.

A mediana do tempo de aplicação do CCQ por meio de entrevista foi de 2,76 min (2,38-3,38 min). Na análise

**Tabela 1.** Características gerais da amostra.<sup>a</sup>

Variáveis	(N = 50)
Sexo masculino	30 (60)
Idade, anos	66 ± 8
IMC, kg/m <sup>2</sup>	24,7 ± 4,7
Carga tabágica, anos-maço	50 [23-73]
Função pulmonar	
VEF <sub>1</sub> /CVF	0,53 [0,44-0,61]
VEF <sub>1</sub> , L	1,15 [0,80-1,69]
VEF <sub>1</sub> , % do previsto	44,7 ± 17,9
CVF, L	2,14 [1,69-2,80]
CVF, % do previsto	66,2 [54,0-75,5]
GOLD, gravidade	
I	1 (2)
II	17 (34)
III	21 (42)
IV	11 (22)
GOLD, classificação	
A	12 (24)
B	19 (38)
C	0 (0)
D	19 (38)
Pontuação no CAT	18 ± 10
Pontuação na escala mMRC	1 [1-4]
Pontuação no SGRQ	
Total	38,3 [18,0-67,8]
Sintomas	44,6 ± 22,9
Atividade	51,5 [25,1-85,1]
Impacto psicossocial	28,9 [14,4-55,8]

CAT: *COPD Assessment Test*; mMRC: modificada do *Medical Research Council*; e SGRQ: *Saint George's Respiratory Questionnaire*. <sup>a</sup>Valores expressos em forma de n (%), média ± dp ou mediana [IIQ].

de consistência interna, os valores de  $\alpha$  (IC95%) para todos os itens do CCQ (pontuação total no CCQ) e para os domínios sintomas (itens 1, 2, 5 e 6), estado mental (itens 3 e 4) e estado funcional (itens 7, 8, 9 e 10) foram, respectivamente, de 0,93 (0,91-0,96), 0,77 (0,66-0,85), 0,79 (0,64-0,87) e 0,94 (0,91-0,96).

A Tabela 2 apresenta a pontuação total e a pontuação nos domínios do CCQ em cada aplicação. Foram observadas diferenças entre os avaliadores para a pontuação total e nos domínios estado mental e estado funcional, assim como para a pontuação no domínio estado mental entre o teste e reteste ( $p \leq 0,05$  para todas). A Tabela 2 também apresenta a análise da confiabilidade interavaliadores, da confiabilidade teste-reteste e do erro de medida. Todos os CCI<sub>3,1</sub> foram  $\geq 0,80$ , sendo que a faixa de IC95% foi mais ampla entre o teste e reteste do que entre os avaliadores. Os resultados de EPM<sub>c</sub> concordância, MMD<sub>i</sub> individual e MMD<sub>g</sub> grupo foram semelhantes entre os avaliadores e entre o teste e reteste para a pontuação total no CCQ, mas foram menores entre os avaliadores para os domínios sintomas e estado mental, assim como entre o teste e reteste para o domínio estado funcional. A Figura 1 também apresenta o erro de medida segundo os LC da pontuação total no CCQ, que foram mais amplos entre o teste e reteste do que entre os avaliadores.

A Tabela 3 e a Figura 2 apresentam as correlações entre as pontuações no CCQ e as pontuações no SGRQ. Todas as correlações foram fortes ( $r > 0,70$ ), exceto a correlação entre o domínio estado mental do CCQ e os domínios sintomas e atividade do SGRQ, que foram boas ( $0,50 < r < 0,70$ ).<sup>(19)</sup>

Apenas 4 pacientes apresentaram pontuação mínima (8%), e 1 apresentou pontuação máxima (2%), indicando a ausência de efeitos de piso e teto.<sup>(9)</sup>

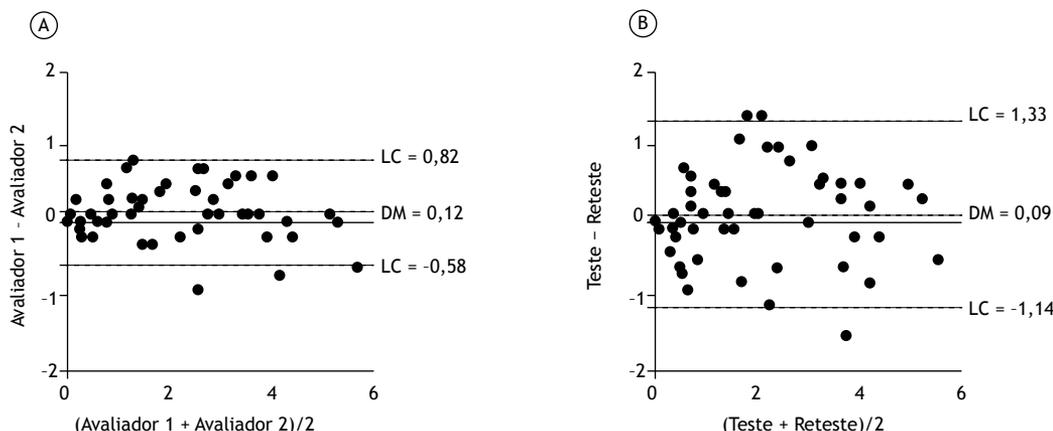
## DISCUSSÃO

No presente estudo, as propriedades de medida da versão brasileira do CCQ foram testadas em uma amostra de pacientes com DPOC no Brasil. Os resultados

**Tabela 2.** Pontuações, confiabilidade e erro de medida do *Clinical COPD Questionnaire* entre as medidas.<sup>a</sup>

CCQ	Avaliador 1 (Teste)		Avaliador 2				Avaliador 1 (Reteste)				
	Pontuação	Pontuação	CCI <sub>3,1</sub> (IC95%)	EPM <sub>c</sub>	MMD <sub>i</sub>	MMD <sub>g</sub>	Pontuação	CCI <sub>3,1</sub> (IC95%)	EPM <sub>c</sub>	MMD <sub>i</sub>	MMD <sub>g</sub>
Total	1,85 [0,77-3,52]	1,65 [0,60-3,07]	0,97 (0,95-0,98)	0,65	1,81	0,26	1,50 [0,80-3,10]	0,92 (0,86-0,95)	0,65	1,80	0,25
Sintomas	2,21 ± 1,44	2,25 ± 1,60	0,92 (0,86-0,95)	0,47	1,30	0,18	2,30 ± 1,42	0,81 (0,69-0,89)	0,77	2,14	0,30
Estado mental	1,5 [0,0-4,0]	1,00 [0,00-3,50]	0,96 (0,92-0,97)	0,70	1,91	0,27	0,75 [0,00-2,50]	0,80 (0,67-0,88)	2,04	5,66	0,80
Estado funcional	1,75 [0,68-3,5]	1,50 [0,44-3,00]	0,93 (0,85-0,96)	1,46	4,06	0,57	1,37 [0,25-3,06]	0,90 (0,83-0,94)	0,90	2,49	0,35

CCQ: *Clinical COPD Questionnaire*; CCI<sub>3,1</sub>: coeficiente de correlação intraclasse de duas vias modelo de efeitos mistos para medidas únicas; EPM<sub>c</sub>: erro-padrão de medida do tipo concordância; MMD<sub>i</sub>: mínima mudança detectável individual; e MMD<sub>g</sub>: mínima mudança detectável no grupo. <sup>a</sup>Valores expressos em mediana. [IIQ] ou média ± dp. \* $p > 0,05$  vs. avaliador 1 (teste).



**Figura 1.** Disposições gráficas de Bland-Altman da pontuação total no *Clinical COPD Questionnaire*, mostrando a confiabilidade interavaliadores (em A) e a confiabilidade teste-reteste (em B). LC: limite de concordância; e DM: diferença média.

sugerem que essa PROM é confiável e válida quando aplicada por meio de entrevista. Pelo que sabemos, este foi o primeiro estudo a analisar as propriedades de medida da versão brasileira do CCQ.

A análise de consistência interna da versão brasileira do CCQ revelou valores entre 0,77 e 0,94. No estudo original de desenvolvimento do CCQ,<sup>(7)</sup> esses valores variaram de 0,78 a 0,91. Entre os domínios, o maior valor foi no domínio estado funcional,<sup>(7)</sup> o que foi semelhante no presente estudo. Em outros estudos de validação, também foram encontrados valores acima de 0,70.<sup>(20-25)</sup> Por definição, a consistência interna determina o grau de inter-relação entre os itens.<sup>(26)</sup> Valores abaixo de 0,70 indicam falta de correlação entre os itens da PROM.<sup>(27)</sup> No entanto, valores acima de 0,95 podem indicar que a PROM contém muitos itens que estão avaliando o mesmo construto, sugerindo redundância.<sup>(28)</sup> Portanto, a consistência interna da versão brasileira do CCQ e seus domínios foi positiva ( $0,70 \leq \alpha \leq 0,95$ )<sup>(9)</sup> e suficiente ( $\alpha \geq 0,70$ ).<sup>(29)</sup>

Este estudo apresentou a análise de confiabilidade da versão brasileira do CCQ entre dois avaliadores e ao longo de um intervalo de tempo. O CCI mínimo foi de 0,80. No estudo que apresentou as propriedades de medida da versão original do CCQ,<sup>(7)</sup> o CCI foi de 0,94 para a pontuação total entre o teste e reteste. Nosso achado é semelhante aos relatados nos estudos de validação das versões italiana<sup>(20)</sup> (CCI = 0,99) e persa<sup>(30)</sup> (CCI = 0,98) do CCQ. A confiabilidade é definida como a proporção da variância total das medidas que é devida a diferenças verdadeiras entre os pacientes. A análise estatística deve preferencialmente ser realizada por meio do cálculo do CCI, pois ele considera os erros sistemáticos entre medidas repetidas.<sup>(26)</sup> No presente estudo, optamos pelo uso do CCI<sub>3,17</sub>, no qual cada indivíduo é avaliado por cada avaliador, sendo que esses são os únicos avaliadores de interesse, e a confiabilidade é calculada a partir de uma única medida.<sup>(17)</sup> Os CCI variam de 0 a 1; valores próximos de 1 indicam uma variação de erro pequena em comparação com a variação do paciente.

Isso significa que esses valores também dependem da heterogeneidade da população, ou seja, quando a população é mais homogênea, é mais fácil encontrar um CCI mais próximo de 0.<sup>(6)</sup> Considerando um CCI de pelo menos 0,70 como critério de qualidade, pode-se dizer que a confiabilidade da versão brasileira do CCQ entre os avaliadores e entre o teste e reteste foi positiva<sup>(9)</sup> e suficiente<sup>(29)</sup> em nossa amostra.

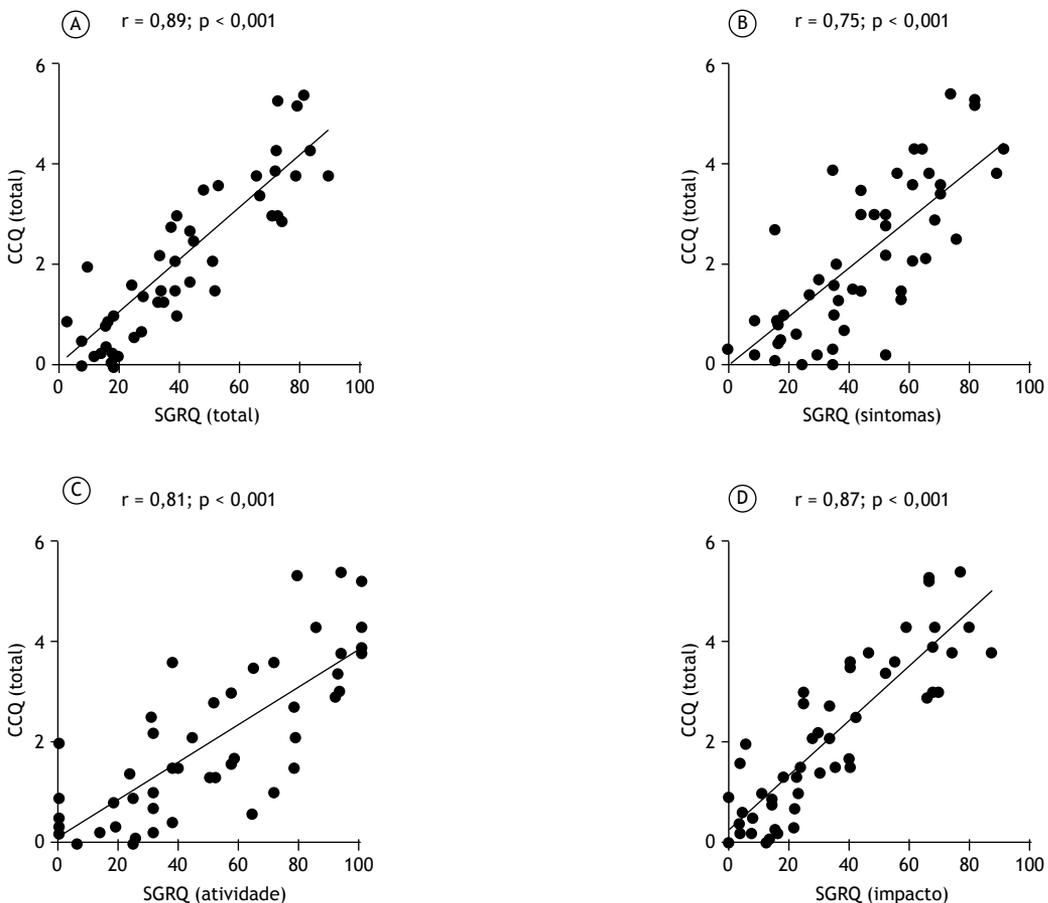
Houve diferenças nas pontuações entre as aplicações do CCQ. No entanto, testes de comparação não são recomendados pelas Normas Consensuais para a Seleção de Instrumentos de Medida de Saúde<sup>(29)</sup> para a análise de confiabilidade. Isso porque esses testes mostram apenas a concordância entre as aplicações para os valores centrais, mas não fornecem informações sobre a concordância entre as aplicações para valores individuais.<sup>(31)</sup> No presente estudo, as pontuações no CCQ nas três aplicações foram maiores nos domínios sintomas e estado funcional, corroborando os resultados da validação original do CCQ.<sup>(7)</sup>

O erro de medida foi analisado por meio de  $EPM_{concordância}$ ,  $MMD_{individual}$ ,  $MMD_{grupo}$  e LC para determinar a confiabilidade interavaliadores e teste-reteste da versão brasileira do CCQ. O  $EPM_{concordância}$  da pontuação total no CCQ foi o mesmo entre os avaliadores e entre o teste e reteste (0,65). Tsiligianni et al.,<sup>(3)</sup> estudando uma amostra de pacientes com DPOC clinicamente estáveis, relataram um valor menor do EPM para o CCQ quanto à confiabilidade teste-reteste. No entanto, esses autores calcularam o EPM utilizando uma equação diferente, que não considera a variação causada por diferenças sistemáticas entre os avaliadores.<sup>(6)</sup> Essa propriedade de medida representa o erro sistemático e aleatório da pontuação de um paciente que não é atribuído a mudanças reais no construto a ser medido.<sup>(26)</sup> No presente estudo, calculamos o EPM, que representa o desvio-padrão de medidas repetidas de um indivíduo; e então a MMD, que consiste na mudança mínima que deve ser superada para garantir uma mudança real do indivíduo.<sup>(26)</sup> Isso significa que uma mudança observada deve ser maior do que o

**Tabela 3.** Correlações da pontuação nos domínios do *Clinical COPD Questionnaire* (CCQ) com a pontuação total e a pontuação nos domínios do *Saint George's Respiratory Questionnaire* (SGRQ).

SGRQ	Domínios do CCQ		
	Sintomas	Estado mental	Estado funcional
Total	0,83 <sup>*</sup>	0,78 <sup>*</sup>	0,88 <sup>*</sup>
Sintomas	0,75 <sup>**</sup>	0,64 <sup>*</sup>	0,73 <sup>*</sup>
Atividade	0,75 <sup>*</sup>	0,64 <sup>*</sup>	0,82 <sup>*</sup>
Impacto psicossocial	0,78 <sup>*</sup>	0,81 <sup>*</sup>	0,85 <sup>*</sup>

\* $p < 0,01$ ; correlação ranqueada de Spearman. \*\* $p < 0,01$ ; correlação ranqueada de Pearson.



**Figura 2.** Correlações da pontuação total no *Clinical COPD Questionnaire* (CCQ) com a pontuação total (em A) e a pontuação nos domínios (em B, C e D) do *Saint George's Respiratory Questionnaire* (SGRQ).

limite da MMD para ser considerada verdadeira.<sup>(6)</sup> Os LC foram demonstrados em disposições gráficas de Bland-Altman para apoiar a interpretação do tamanho do erro de medida. É possível visualizar a magnitude do erro de medida quando relacionamos os LC com a faixa de pontuação. Por definição, 95% das diferenças entre medidas repetidas devem estar dentro dos LC. Um valor fora dos LC pode indicar uma mudança real.<sup>(6)</sup> Para saber se o erro de medida é aceitável ou não, deve-se também analisar a mínima mudança importante (MMI). A MMD e a MMI podem ser utilizadas para decidir se uma mudança real e clinicamente importante ocorreu com um paciente.<sup>(28)</sup> Outros estudos sobre o CCQ descreveram um valor de MMI próximo de 0,4.<sup>(32-35)</sup> No presente estudo, no entanto,

a MMI não foi calculada. Portanto, os erros de medida interavaliadores e teste-reteste da versão brasileira do CCQ foram classificados como indeterminados.<sup>(9,29)</sup>

De acordo com os resultados, a hipótese relacionada à análise do construto foi confirmada. Houve correlação positiva de valores de  $r$  de pelo menos 0,5 entre a pontuação total na versão brasileira do CCQ e a pontuação total e dos domínios do SGRQ. Para a elaboração da hipótese, foi considerado o valor mínimo de  $r$  (0,53) encontrado entre a pontuação total no CCQ e as pontuações no SGRQ no estudo de validade da versão original do CCQ.<sup>(7)</sup> Outros estudos também relataram valores de  $r$  semelhantes entre o CCQ e o SGRQ.<sup>(22,33)</sup> Resultados semelhantes também foram encontrados entre o CCQ e outras PROMs que

avaliam a qualidade de vida relacionada à saúde geral e específica de pacientes com DPOC.<sup>(20,21)</sup> A pontuação nos domínios sintomas e estado funcional do CCQ apresentou forte correlação ( $r > 0,70$ )<sup>(19)</sup> com todos os domínios do SGRQ. A pontuação no domínio estado mental do CCQ apresentou forte correlação apenas com a pontuação total e a pontuação no domínio impacto do SGRQ, provavelmente porque esse é o único domínio do SGRQ que apresenta perguntas sobre alterações psicossociais.<sup>(15)</sup> A validade de construto estima o grau de consistência entre as pontuações na PROM e as suposições baseadas na hipótese de que a PROM mede validamente o construto que se pretende medir.<sup>(9,28)</sup> Na amostra estudada, a validade de construto do CCQ atingiu os critérios de qualidade, sendo avaliada como positiva<sup>(9)</sup> e suficiente,<sup>(29)</sup> pois a hipótese da validade de construto foi atendida.

Na análise de interpretabilidade, não foram observados efeitos de piso e teto. Em estudos semelhantes, também não foram detectados efeitos de piso e teto na pontuação total no CCQ.<sup>(21,22)</sup> A presença de um efeito de piso ou teto pode indicar que itens extremos estão ausentes na extremidade inferior ou superior da PROM e pode limitar sua capacidade de discriminar pacientes e medir mudanças.<sup>(9)</sup>

Em razão da baixa escolaridade de parte da amostra estudada, o CCQ foi aplicado por meio de entrevista. Como se esperava, o preenchimento do CCQ por meio de entrevista no presente estudo demorou um pouco mais do que o preenchimento do CCQ original autoaplicável (aproximadamente 2 min).<sup>(7)</sup> Já foi relatada concordância entre as pontuações no CCQ autoaplicável e as pontuações no CCQ aplicado pelo clínico durante a consulta, assim como entre as pontuações no CCQ autoaplicável e as pontuações no CCQ aplicado pelo clínico por meio de entrevista semi-estruturada aprofundada.<sup>(36)</sup> No entanto, este foi o primeiro estudo a analisar as propriedades de medida do CCQ preenchido por meio de entrevista, sem interferência dos avaliadores. Uma meta-análise<sup>(37)</sup> relatou que, em geral, o autopreenchimento e o preenchimento assistido de uma PROM produzem pontuações equivalentes, sustentando que o formato de entrevista é um modo válido de aplicação. Além disso, no presente estudo, o preenchimento assistido do CCQ permitiu a análise sem precedentes da confiabilidade interavaliadores e do erro de medida.

O intervalo de tempo entre as aplicações do CCQ pelos avaliadores pode ter sido muito curto para evitar vies de memória e, portanto, pode ter comprometido a confiabilidade interavaliadores. No entanto, até onde sabemos, não há na literatura nenhuma recomendação quanto a um intervalo de tempo adequado para a

aplicação de uma PROM por avaliadores na literatura. Além disso, embora a faixa de IC95% entre os avaliadores tenha sido menor do que entre o teste e reteste, os CCI<sub>3,1</sub> interavaliadores e teste-reteste foram semelhantes e maiores que 0,70.<sup>(9,29)</sup> Outra possível limitação do estudo foi o desenho transversal, que pode ter afetado o erro de medida. No entanto, este foi o primeiro estudo que relatou o EPM<sup>concordância</sup> a MMD<sup>individual</sup> a MMD<sup>grupo</sup> e os LC da versão brasileira do CCQ. Nossos resultados revelam que as mudanças são reais, e não devidas a erro de medida, ao mostrar que quanto a pontuação precisa mudar antes de garantir que uma mudança real ocorreu, fornecendo condições para a interpretação de medidas longitudinais. Além disso, os valores de EPM<sup>concordância</sup>, MMD<sup>individual</sup>, MMD<sup>grupo</sup> e LC são apresentados na mesma unidade de medida da PROM que está sendo estudada, o que facilita a interpretação das pontuações pelos profissionais de saúde na prática clínica.<sup>(6)</sup>

Em conclusão, a versão brasileira do CCQ apresenta consistência interna e confiabilidade suficientes, ou seja, os itens da PROM estão inter-relacionados e suas pontuações são estáveis e capazes de reproduzir resultados consistentes em medidas repetidas entre avaliadores diferentes e ao longo do tempo. Além disso, a versão em português do CCQ para uso no Brasil demonstra validade de construto suficiente, e as correlações entre a pontuação total no CCQ e as pontuações no SGRQ são consistentes. No presente estudo, foram apresentados os parâmetros EPM<sup>concordância</sup>, MMD<sup>individual</sup>, MMD<sup>grupo</sup> e LC do erro de medida. No entanto, recomenda-se que mais estudos sejam realizados para testar a suficiência do erro de medida por meio do cálculo da MMI. Não foram encontrados efeitos de piso ou teto na pontuação total na versão brasileira do CCQ. Pelo que sabemos, este foi o primeiro estudo a avaliar as propriedades de medida do CCQ em uma amostra de pacientes com DPOC no Brasil, contribuindo para a divulgação e promoção do uso dessa PROM por profissionais de saúde e pesquisadores para avaliação do estado de saúde de seus pacientes.

## CONTRIBUIÇÕES DOS AUTORES

AR, FRF e RM: concepção e planejamento do estudo, interpretação das evidências, redação e revisão das versões preliminares e final e aprovação da versão final. APQ: redação e revisão das versões preliminares e final e aprovação da versão final. CMR: concepção e planejamento do estudo, interpretação das evidências e aprovação da versão final. MMB: redação e revisão das versões preliminares e final e aprovação da versão final. JK e TvdM: redação e revisão das versões preliminares e final e aprovação da versão final.

## REFERÊNCIAS

1. Soler-Cataluña JJ, Alcázar-Navarrete B, Miravittles M. The concept of control of COPD in clinical practice. *Int J Chron Obstruct Pulmon Dis*. 2014;9:1397-1405. <https://doi.org/10.2147/COPD.S71370>
2. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive

- summary. *Am J Respir Crit Care Med.* 2013;187(4):347-365. <https://doi.org/10.1164/rccm.201204-0596PP>
3. Tsiligianni IG, van der Molen T, Moraitaki D, Lopez I, Kocks JW, Karagiannis K, et al. Assessing health status in COPD. A head-to-head comparison between the COPD assessment test (CAT) and the clinical COPD questionnaire (CCQ). *BMC Pulm Med.* 2012;12:20. <https://doi.org/10.1186/1471-2466-12-20>
  4. Cave AJ, Atkinson L, Tsiligianni IG, Kaplan AG. Assessment of COPD wellness tools for use in primary care: an IPCRG initiative. *Int J Chron Obstruct Pulmon Dis.* 2012;7:447-456. <https://doi.org/10.2147/COPD.S29868>
  5. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19(4):539-549. <https://doi.org/10.1007/s11136-010-9606-8>
  6. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide (Practical Guides to Biostatistics and Epidemiology)*. Cambridge: Cambridge University Press; 2011. <https://doi.org/10.1017/CBO9780511996214>
  7. van der Molen T, Willemsse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes.* 2003;1:13. <https://doi.org/10.1186/1477-7525-1-13>
  8. Lourenço RA, Veras RP. Mini-Mental State Examination: psychometric characteristics in elderly outpatients. *Rev Saude Publica.* 2006;40(4):712-719. <https://doi.org/10.1590/S0034-89102006000500023>
  9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
  10. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26(2):319-338. <https://doi.org/10.1183/09031936.05.00034805>
  11. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol.* 2007;33(4):397-406. <https://doi.org/10.1590/S1806-37132007000400008>
  12. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary [published correction appears in *Eur Respir J.* 2017 Jun 22;49(6):]. *Eur Respir J.* 2017;49(3):1700214. <https://doi.org/10.1183/13993003.00214-2017>
  13. Silva GP, Morano MT, Viana CM, Magalhães CB, Pereira ED. Portuguese-language version of the COPD Assessment Test: validation for use in Brazil. *J Bras Pneumol.* 2013;39(4):402-408. <https://doi.org/10.1590/S1806-37132013000400002>
  14. Kovelis D, Segretti NO, Probst VS, Lareau SC, Brunetto AF, Pitta F. Validation of the Modified Pulmonary Functional Status and Dyspnea Questionnaire and the Medical Research Council scale for use in Brazilian patients with chronic obstructive pulmonary disease. *J Bras Pneumol.* 2008;34(12):1008-1018. <https://doi.org/10.1590/S1806-37132008001200005>
  15. Camelier A, Rosa FW, Salim C, Nascimento OA, Cardoso F, Jardim JR. Using the Saint George's Respiratory Questionnaire to evaluate quality of life in patients with chronic obstructive pulmonary disease: validating a new version for use in Brazil. *J Bras Pneumol.* 2006;32(2):114-122. <https://doi.org/10.1590/S1806-37132006000200006>
  16. Fan X, Thompson B. Confidence Intervals for Effect Sizes: Confidence Intervals about Score Reliability Coefficients, Please: An EPM Guidelines Editorial. *Educ Psychol Meas.* 2001;61(4):517-531. <https://doi.org/10.1177/00131644011614001>
  17. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods.* 1996;1(1):30-46. <https://doi.org/10.1037/1082-989X.1.1.30>
  18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
  19. Hazra A, Gogtay N. *Biostatistics Series Module 6: Correlation and Linear Regression.* Indian J Dermatol. 2016;61(6):593-601. <https://doi.org/10.4103/0019-5154.193662>
  20. Damato S, Bonatti C, Frigo V, Pappagallo S, Racanelli R, Rampoldi C, et al. Validation of the Clinical COPD questionnaire in Italian language. *Health Qual Life Outcomes.* 2005;3:9. <https://doi.org/10.1186/1477-7525-3-9>
  21. Reda AA, Kotz D, Kocks JW, Wesseling G, van Schayck CP. Reliability and validity of the clinical COPD questionnaire and chronic respiratory questionnaire. *Respir Med.* 2010;104(11):1675-1682. <https://doi.org/10.1016/j.rmed.2010.04.023>
  22. Ställberg B, Nokela M, Ehms PO, Hjemdal P, Jonsson EW. Validation of the clinical COPD Questionnaire (CCQ) in primary care. *Health Qual Life Outcomes.* 2009;7:26. <https://doi.org/10.1186/1477-7525-7-26>
  23. Papadopoulos G, Vardavas CI, Limperi M, Linardis A, Georgoudis G, Behrakis P. Smoking cessation can improve quality of life among COPD patients: validation of the clinical COPD questionnaire into Greek. *BMC Pulm Med.* 2011;11:13. <https://doi.org/10.1186/1471-2466-11-13>
  24. Antoniu SA, Puiu A, Zaharia B, Azoicai D. Health status during hospitalisations for chronic obstructive pulmonary disease exacerbations: the validity of the Clinical COPD Questionnaire. *Expert Rev Pharmacoecon Outcomes Res.* 2014;14(2):283-287. <https://doi.org/10.1586/14737167.2014.887446>
  25. Berkhof FF, Metzemaekers L, Uil SM, Kerstjens HA, van den Berg JW. Health status in patients with coexistent COPD and heart failure: a validation and comparison between the Clinical COPD Questionnaire and the Minnesota Living with Heart Failure Questionnaire. *Int J Chron Obstruct Pulmon Dis.* 2014;9:999-1008. <https://doi.org/10.2147/COPD.S66028>
  26. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
  27. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol.* 1993;78(1):98-104. <https://doi.org/10.1037/0021-9010.78.1.98>
  28. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable?. *Injury.* 2011;42(3):236-240. <https://doi.org/10.1016/j.injury.2010.11.042>
  29. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147-1157. <https://doi.org/10.1007/s11136-018-1798-3>
  30. Hasanpour N, Attarbashi Moghadam B, Sami R, Tavakol K. Cross-Cultural Adaptation, Reliability and Validity Study of the Persian Version of the Clinical COPD Questionnaire. *Acta Med Iran.* 2016;54(8):518-524.
  31. Bland JM, Altman DG. Comparing two methods of clinical measurement: a personal history. *Int J Epidemiol.* 1995;24 Suppl 1:S7-S14. [https://doi.org/10.1093/ije/24.Supplement\\_1.S7](https://doi.org/10.1093/ije/24.Supplement_1.S7)
  32. Kocks JW, Tuinenga MG, Uil SM, van den Berg JW, Ståhl E, van der Molen T. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res.* 2006;7(1):62. <https://doi.org/10.1186/1465-9921-7-62>
  33. Canavan JL, Dilaver D, Clark AL, Jones SE, Nolan CM, Kon SS, et al. Clinical COPD Questionnaire in patients with chronic respiratory disease. *Respirology.* 2014;19(7):1006-1012. <https://doi.org/10.1111/resp.12350>
  34. Smid DE, Franssen FM, Houben-Wilke S, Vanfleteren LE, Janssen DJ, Wouters EF, et al. Responsiveness and MCID Estimates for CAT, CCO, and HADS in Patients With COPD Undergoing Pulmonary Rehabilitation: A Prospective Analysis. *J Am Med Dir Assoc.* 2017;18(1):53-58. <https://doi.org/10.1016/j.jamda.2016.08.002>
  35. Alma HJ, de Jong C, Jelusic D, Wittmann M, Schuler M, Kollen BJ, et al. Assessing health status over time: impact of recall period and anchor question on the minimal clinically important difference of copd health status tools. *Health Qual Life Outcomes.* 2018;16(1):130. <https://doi.org/10.1186/s12955-018-0950-7>
  36. Kocks JW, Kerstjens HA, Sniijders SL, de Vos B, Biermann JJ, van Hengel P, et al. Health status in routine clinical practice: validity of the clinical COPD questionnaire at the individual patient level. *Health Qual Life Outcomes.* 2010;8:135. <https://doi.org/10.1186/1477-7525-8-135>
  37. Rutherford C, Costa D, Mercieca-Bebber R, Rice H, Gabb L, King M. Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Qual Life Res.* 2016;25(3):559-574. <https://doi.org/10.1007/s11136-015-1110-8>