

ARTIGO DE PESQUISA

Utilidade do Teste de Avaliação da DPOC (CAT) em consultas do serviço primário de saúde: estudo controlado randomizado

Kevin Gruffydd-Jones,¹ *Helen C Marsden,² Steve Holmes,³ Peter Kardos,⁴
Roger Escamilla,⁵ Roberto Dal Negro,⁶ June Roberts,⁷ Gilbert Nadeau,² Mathieu Vasselle,⁸
David A Leather,² Paul Jones⁹

¹ Box Surgery, Box, Corsham, Wiltshire, Reino Unido

² GlaxoSmithKline, Londres, Reino Unido

³ Clínica Médica The Park, Shepton Mallet, Somerset, Reino Unido

⁴ Gemeinschaftspraxis und Zentrum für Pneumologie, Allergologie, Schlafmedizin, Frankfurt, Alemanha

⁵ Hospital Larrey-CHU Toulouse, Toulouse, França

⁶ Ospedale Civile Orlandi di Bussolengo, Itália

⁷ Salford Royal Hospital Foundation Trust, Salford, Reino Unido

⁸ GlaxoSmithKline, Marly-le-Roi, França

⁹ St George's University of London, Reino Unido

Originalmente recebido em 20/06/2012; revisado em 14/09/2012; aceito em 7/10/2012; publicado on-line em 2/01/2013

Resumo

Fundamentos: Um dos objetivos do Teste de Avaliação da DPOC (CAT) é auxiliar na comunicação entre o médico e o paciente quanto às consequências da doença pulmonar obstrutiva crônica (DPOC) para a vida do paciente.

Objetivos: Investigar o impacto do CAT na qualidade das consultas nos serviços primários de saúde envolvendo pacientes com DPOC.

Métodos: Médicos de serviços primários de saúde da Europa realizaram seis consultas com pacientes padronizados com DPOC (representados por atores treinados). Os médicos foram randomizados para atender ao paciente com (braço CAT+) ou sem (CAT-) preenchimento do CAT durante a consulta. Essas consultas foram gravadas em vídeo e avaliadores independentes atribuíram pontuações aos médicos quanto à sua capacidade de identificar e abordar problemas específicos dos pacientes, como depressão (subescore A), revisar problemas padronizados da DPOC, como dispneia (subescore B), sua compreensão a respeito do caso (escore de compreensão) e seu desempenho geral. O desfecho primário foi o escore global (subescores A + B, variando de 0 a 40).

Resultados: Ao todo, 165 médicos foram incluídos no estudo e realizaram seis consultas cada um; 882 consultas foram consideradas adequadas para análise. Não se observou diferenças entre os braços quanto ao escore global (20,3 no braço CAT- e 20,7 no braço CAT+; IC-95% = -1,0 a 1,8; p = 0,606) ou ao subescore A (p = 0,255). Observou-se uma diferença estatisticamente significativa, embora de limitada relevância clínica, na média do subescore B (8,8 no braço CAT- e 9,6 no braço CAT+; IC-95% = 0,0 a 1,6; p = 0,045). Não houve diferença no escore de compreensão (p = 0,824) ou no desempenho geral (p = 0,655).

Conclusões: O CAT é um instrumento específico que auxilia na avaliação médica da DPOC. Ele parece melhorar a detecção de sintomas não ligados à DPOC e de comorbidades.

© 2013 Primary Care Respiratory Society UK. Todos os direitos reservados.

K Gruffydd-Jones et al. *Prim Care Respir J* 2013; 22(1): 37-43

<http://dx.doi.org/10.4104/pcrj.2013.00001>

Palavras-chave: DPOC, Teste de Avaliação da DPOC, serviço primário de saúde, estudo controlado randomizado, casos de pacientes padronizados.

Keywords: COPD, COPD Assessment Test, primary care, randomised controlled trial, standardised patient cases

Veja o editorial associado, de Langhammer e Jones, na *Prim Care Respir J* 2013; 22(1) página 8.

A versão integral deste artigo, assim como o apêndice on-line, está disponível no website www.thepcrj.org

Introdução

A qualidade da consulta proporcionada por um médico pode ter profundo impacto na qualidade do atendimento e do engajamento do paciente nas decisões terapêuticas.¹ As consultas mais efetivas são aquelas nas quais os médicos reconhecem diretamente e respondem

* **Correspondência:** Dr Helen C Marsden, GlaxoSmithKline, Stockley Park West, 1-3 Ironbridge Road, Uxbridge UB11 1BT, UK.
Tel: +44 (0)20 8990 2776. Fax: +44 (0)20 8990 3504. E-mail: Helen.c.marsden@gsk.com

aos problemas e preocupações dos pacientes.² Os pacientes geralmente não expressam todos os seus problemas numa consulta, o que pode levar a maus resultados.³ Portanto, instrumentos que melhorem a comunicação entre o paciente e o médico podem permitir que os problemas do paciente sejam explicitados e abordados.

O Teste de Avaliação da DPOC (CAT) é um questionário preenchido pelo paciente, elaborado para indicar uma medida simples e confiável do estado de saúde do paciente com doença pulmonar obstrutiva crônica (DPOC).⁴ O CAT é formado por oito questões que cobrem os sintomas da DPOC que mais incomodam o paciente, como a dispnéia e as limitações às atividades cotidianas. Já foi demonstrado que o instrumento tem propriedades similares às de questionários mais complexos sobre o estado de saúde do paciente, como o Questionário Respiratório de St. George⁵ e o Questionário de Doenças Respiratórias Crônicas.⁶ Entretanto, é mais breve, o que o torna adequado ao uso rotineiro na clínica. Já foi demonstrado também que esse teste é sensível às exacerbações e às melhoras da população geral com DPOC após uma reabilitação pulmonar.⁷

Quando o CAT foi desenvolvido, um dos seus objetivos era auxiliar na comunicação entre o médico e o paciente quanto ao impacto da DPOC. Até o momento, esse aspecto do CAT ainda não foi testado. Assim, procuramos avaliar o impacto do CAT na comunicação médico-paciente. Para testar a robustez dos resultados do CAT, delineamos um novo estudo, que por padronizar os pacientes e os critérios de avaliação, permitiu a sua condução em vários países.⁸

Métodos

Este estudo aberto randomizado (1:1) de grupos paralelos com uma única visita comparou a qualidade da consulta médica com ou sem o CAT em um ambiente simulado padronizado. Médicos de serviços primários de saúde foram selecionados por meio de entrevista telefônica em cinco países europeus. Aqueles que referiram experiência no manejo da DPOC (ao menos três pacientes), mas não no uso do CAT, foram convidados a participar do estudo de comunicação médico-paciente. As questões da seleção inicial cobriram também outras doenças e medidas de resultados referidas pelo paciente (PROMs), de forma que os médicos não sabiam que o estudo envolvia o CAT até que comparece-

ram à sessão agendada. Utilizaram-se dois ou quatro locais espalhados geograficamente em cada país, dependendo do número de médicos necessários. Os médicos compareceram às sessões em grupos de cinco ou seis, com dois grupos por local. Cada grupo foi randomizado para atender ao paciente com (braço CAT+) ou sem (braço CAT-) o preenchimento do CAT. Usou-se um desenho com dois níveis hierárquicos, tamanho do bloco de randomização de dois, sem estratificação, de forma que um grupo de cada local foi alocado em um dos dois braços. Os recrutadores não tiveram acesso aos dados da randomização.

O Serviço Nacional de Ética em Pesquisas do Reino Unido confirmou que não era necessário obter aprovação ética para a realização do estudo. Os médicos deram seu consentimento para participarem do estudo e foram questionados quanto aos seus dados demográficos, histórico de tabagismo e experiência na prática médica.

Os médicos participantes receberam um breve treinamento em DPOC. Aqueles que foram alocados para o braço CAT+ também receberam um breve treinamento em CAT (cerca de 15 minutos de informações de fundo, como interpretar as pontuações gerais no teste e como identificar áreas específicas de preocupação para o paciente). Não foram fornecidas orientações específicas quanto às ações a serem tomadas com base no escore do teste CAT.

Os médicos realizaram, então, as consultas com seis pacientes padronizados com DPOC (representados por atores treinados). Vinte atores fluentes no idioma relevante fizeram o papel dos pacientes com DPOC. Eles foram treinados em aspectos clínicos da DPOC, nos detalhes do seu caso e em como representar um paciente com DPOC acuradamente. Seis casos de DPOC foram delineados para cobrir variados níveis de gravidade da doença e cenários relevantes para a prática clínica, incluindo-se quatro ou cinco “problemas do paciente” específicos (Tabela 1). Os atores foram treinados para não tomarem a iniciativa de relatar seus problemas ao médico, que precisou descobri-los por meio do questionamento direto.

Cada médico realizou seis consultas. As consultas foram gravadas em vídeo e submetidas a avaliadores independentes com experiência na avaliação de desempenho médico. Os avaliadores foram treinados para atribuir uma pontuação aos médicos, caso estes identificassem e abordassem problemas relevantes para os pacientes e revisassem 10

Tabela 1. Sumário dos casos padronizados

Caso	Idade	VEF ₁	Escore no CAT	Histórico médico	Problemas do paciente
1	68	40	34	DPOC grave, alta carga de manifestações da doença	Levemente deprimido, com restrição de atividades, necessidade de reabilitação pulmonar
2	60	30	21	Estilo de vida sedentário, após exacerbação grave	Perdeu a confiança; necessita de revisão do estilo de vida e do tratamento
3	50	70	9	Diagnóstico recente de DPOC leve, baixa carga de manifestações da doença	Ansiedade em relação ao diagnóstico, necessidade de orientação quanto ao estilo de vida e manejo geral da DPOC
4	65	45	16	Doença cardiovascular como comorbidade, bem tratada. DPOC mal controlada.	Continua fumando, pouca atividade física, impacto da doença nas atividades diárias, baixa adesão ao tratamento
5	70	68	23	Gravemente limitado pela doença; em franca depressão	Manifestações de depressão, baixa adesão ao tratamento, necessita de reabilitação pulmonar e suporte social
6	63	65	19	Imigrante do Oriente Médio / Norte da África. Infecção torácica grave, quer antibióticos. Tosse intensa	Não acredita que tem DPOC; precisa observar o impacto das manifestações da doença

VEF₁ = volume expiratório forçado em um segundo; CAT = Teste de Avaliação da DPOC; DPOC = doença pulmonar obstrutiva crônica.

manifestações padronizadas da DPOC (em três níveis: 0 = nenhum, 1 = pouco, 2 = muito). Os escores foram captados em uma planilha on-line e os escores de problemas dos pacientes (subscore A, até 20 pontos) e de revisão da DPOC (subscore B, até 20 pontos) foram calculados. Além disso, os médicos receberam uma pontuação em relação à sua compreensão do caso (“escore de compreensão”: baixa, aceitável ou acurada) e ao seu desempenho geral (muito baixo, baixo, bom ou muito bom).

O tempo máximo permitido para cada caso foi de dez minutos, e o tempo dedicado à consulta foi registrado.

Os atores apresentaram seus comentários depois de cada consulta, dizendo se sentiram que o médico havia identificado e abordado seus problemas, se a duração da consulta havia sido suficiente e se a consulta havia sido satisfatória.

Conduziu-se um estudo-piloto no Reino Unido, para confirmar a viabilidade da metodologia e fornecer informações sobre o tamanho da amostra do estudo. O estudo-piloto foi conduzido conforme a descrição acima, com 10 médicos e 7 avaliadores. Os resultados do estudo-piloto indicaram que seria possível detectar diferenças entre os braços com essa metodologia.⁸

Conduziu-se ainda um teste comparativo envolvendo todos os avaliadores do estudo principal, para avaliar a variabilidade e a confiabilidade dos escores. Os avaliadores revisaram duas consultas de alta pontuação e duas consultas de baixa pontuação do estudo-piloto, com e sem CAT. Embora se identificasse uma significativa variabilidade na pontuação real e uma diferença nas médias dos escores totais entre os países, a classificação e a diferença entre as consultas de alta e baixa pontuação foram geralmente consistentes.⁸

O desfecho primário do estudo foi a média do escore global (combinação dos subscores A e B), com uma escala de 0 (pior) a 40 (melhor). Como essa medida não havia sido submetida a nenhum estudo prévio, não se conhecia a sensibilidade ou a possível magnitude da diferença entre os braços. Os pesquisadores concordaram com uma diferença de 10% (ou seja, de 4 pontos) entre os braços como convincente quanto a uma diferença real entre os braços. O tamanho da amostra foi calculado com base nos resultados conservadores do estudo-piloto. Visando um poder estatístico de 90% para se identificar uma diferença de pelo menos 3 pontos, com desvio padrão de 12 pontos, seriam necessárias 752 consultas, admitindo-se 10% de dados perdidos, sem fatores de confusão.

Os subscores A e B também foram definidos como desfechos secundários. Todos os escores foram submetidos a uma análise de medidas repetidas da variância, com um modelo misto linear. As diferenças no escore global e nos subscores A e B foram testados com um teste do qui-quadrado ou um teste exato de Fisher. O “escore de compreensão” e a graduação da avaliação global foram analisados por meio de um modelo de equações de estimativa generalizada. Para compensar o efeito do avaliador, identificado no teste comparativo, o avaliador e o caso foram incluídos como variáveis de ajuste nos modelos. A ordem em que os médicos viram os casos também foi incluída, para compensar qualquer efeito de treinamento. Conduziu-se ainda uma análise secundária na qual os dados demográficos do médico foram incluídos como covariáveis. A análise estatística foi realizada com o uso do programa SAS, versão 9.1.

Resultados Participantes

Ao todo, 165 médicos foram recrutados entre outubro de 2010 e fevereiro de 2011. Por questões técnicas, os vídeos de dois médicos não

Figura 1. Diagrama Consort dos médicos recrutados para o estudo

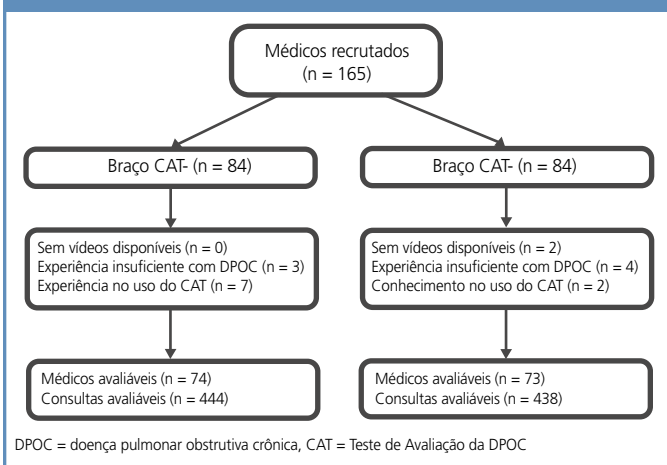


Tabela 2. Características dos médicos avaliáveis

	CAT+ (N = 74)	CAT- (N = 73)	Total (N = 147)
Média de idade (DP)	50,9 (8,5)	51,7 (8,2)	51,3 (8,3)
Proporção entre os gêneros (F/M)	30/70%	36/64%	33/67%
Tabagismo (atual/ prévio/nunca)	4/31/65%	8/32/60%	6/31/63%
Local da clínica atual (país/cidade)	7/93%	7/93%	7/93%
Média de anos (DP) de experiência	18,1 (8,0)	18,8 (8,6)	18,5 (8,2)
≤ 5	4 (5%)	5 (7%)	9 (6%)
5 a 10	12 (16%)	10 (14%)	22 (15%)
10 a 20	33 (45%)	25 (35%)	58 (40%)
20 a 30	21 (28%)	26 (36%)	47 (32%)
> 30	4 (5%)	6 (8%)	10 (7%)
Média (DP) de pacientes com DPOC já tratados	648 (940)	722 (1242)	685 (1099)
Média (DP) de pacientes com DPOC tratados atualmente	96 (150)	82 (159)	89 (154)
Já ouviram falar do CAT	28%	27%	28%
Áustria	9 (12%)	10 (14%)	19 (13%)
França	19 (26%)	17 (23%)	36 (24%)
Alemanha	20 (27%)	16 (22%)	36 (24%)
Irlanda	10 (14%)	11 (15%)	21 (14%)
Reino Unido	16 (22%)	19 (26%)	35 (24%)

DPOC = doença pulmonar obstrutiva crônica, CAT = Teste de Avaliação da DPOC, DP = desvio padrão.

puderam ser avaliados. Outros 16 médicos não preencheram os critérios de inclusão (referiram ter experiência com o uso do CAT ou experiência insuficiente no manejo de pacientes com DPOC na época das gravações), resultando em 147 médicos avaliáveis no estudo (Figura 1). Os médicos avaliáveis representam uma ampla variedade de experiências, sem grandes diferenças entre os braços (Tabela 2). Todos os médicos avaliáveis tiveram seis consultas avaliadas, resultando em 882 vídeos avaliados. Vinte e quatro avaliadores de três países avaliaram de 28 a 51 vídeos cada um, incluindo ao menos uma ocorrência de

Tabela 3. Perfil e carga de trabalho dos avaliadores	
Total (N = 24)	
Número de avaliadores que avaliaram todos os 6 casos	24 (100%)
País do avaliador	
França	5 (21%)
Alemanha	9 (38%)
Reino Unido	10 (42%)
Mín; Máx	
Número de consultas avaliáveis avaliadas por avaliador	
CAT+	11; 28
CAT-	14; 24
Número de médicos avaliáveis avaliados por avaliador	
	7; 21
CAT = Teste de Avaliação da DPOC.	

cada caso padronizado, divididos equitativamente entre vídeos CAT- e CAT+ (Tabela 3).

Análise de eficácia

O escore global médio ajustado de médicos avaliáveis foi de 20,3 no braço CAT- e de 20,7 no braço CAT+ (Tabela 4). A diferença entre os

braços não foi estatisticamente significativa (diferença de 0,4, IC-95% = -1,0 a 1,8, p = 0,606). Uma análise de sensibilidade que incluiu todas as consultas avaliáveis também não mostrou diferença entre os braços (p = 0,548). De maneira similar, não houve diferença entre os braços após o controle de fatores como o país e a idade (p = 0,635). Quando o escore global foi avaliado de acordo com o caso clínico, nítidas diferenças puderam ser observadas entre os casos (Figura 2). O CAT melhorou significativamente o escore global em pacientes com carga elevada de manifestações da DPOC (caso 1), enquanto o escore global foi significativamente reduzido no caso com depressão maior (caso 5).

O impacto do CAT variou entre os subescores A e B e entre os vários tipos de casos (Tabela 4 e Figura 3). As médias ajustadas do subescore A nos braços CAT- e CAT+ foram de 11,5 e 11,1, respectivamente (p = 0,255), e somente o caso 5 mostrou diferença significativa entre os braços (em favor do braço CAT-). Quando os itens individuais do subescore A foram investigados, o item "inapetência e dieta" no caso 5 foi significativamente mais identificado e abordado no braço CAT- que no braço CAT+ (p < 0,002, dados não apresentados).

As médias ajustadas do subescore B nos braços CAT- e CAT+ foram de 8,8 e 9,6, respectivamente (p = 0,045). Considerando-se os casos separadamente, as médias ajustadas do subescore B foram significativamente diferentes entre os braços nos casos 2, 4 e 6 (todos em favor do braço CAT+). Os médicos do braço CAT+ obtiveram significativa-

Tabela 4. Análise média ajustada do escore global, análises de sensibilidade, subescore B e subescore B					
	CAT+ (N = 444)	CAT- (N = 438)	Diferença média	IC-95%	Valor de p
	Média ajustada (EP)	Média ajustada (EP)	(EP)		
Análise primária: escore global de médicos avaliáveis	20,7 (0,5)	20,3 (0,5)	0,4 (0,7)	-1 a 1,8	0,606
Análise secundária: com ajuste em relação à idade do médico, país etc.	20,7 (0,5)	20,3 (0,5)	0,3 (0,7)	-1,1 a 1,7	0,635
Capacidade de identificar e abordar problemas dos pacientes (subescore A)	11,1 (0,3)	11,5 (0,3)	-0,4 (0,4)	-1,2 a 0,3	0,255
Revisão de aspectos padronizados da DPOC (subescore B)	9,6 (0,3)	8,8 (0,2)	0,8 (0,4)	0,0 a 1,6	0,045
DPOC = doença pulmonar obstrutiva crônica, CAT = Teste de Avaliação da DPOC, EP = erro padrão.					

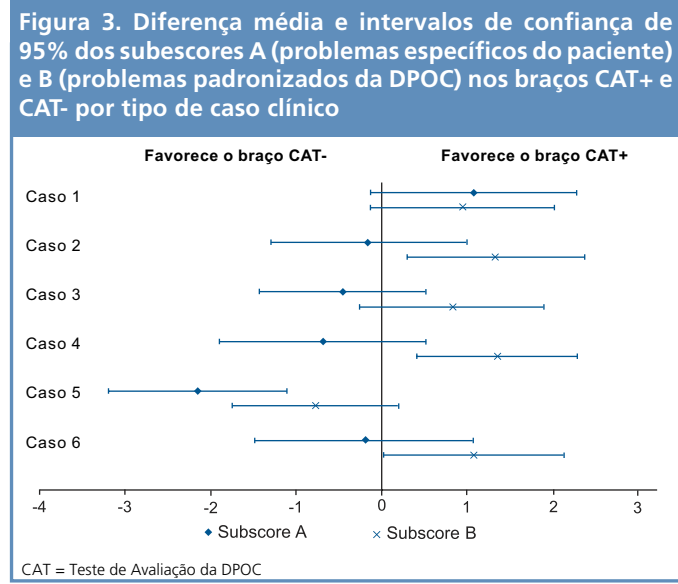
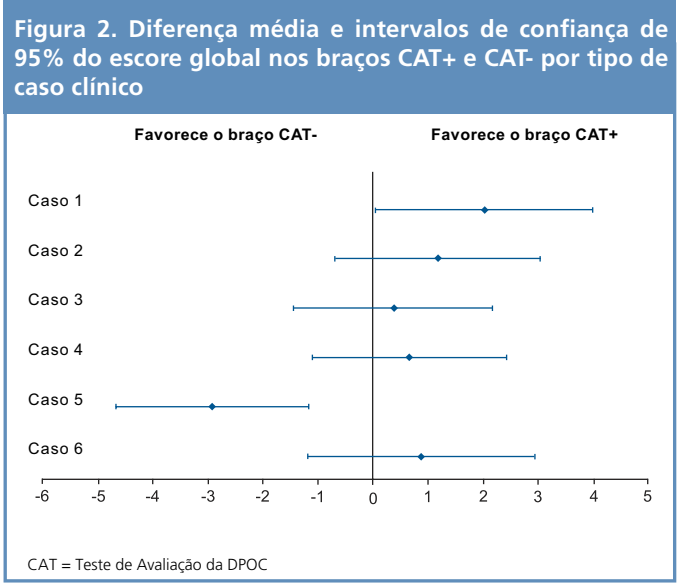


Tabela 5. Análise do subescore B por item

	CAT+ (N = 444)	CAT- (N = 438)	Valor de p
	Não feita/ Pouco/Muito	Não feita/ Pouco/Muito	
Revisão da tosse	20%/47%/33%	32%/44%/24%	<0,001
Revisão do muco	35%/37%/28%	47%/31%/22%	0,002
Revisão da opressão torácica	56%/32%/13%	70%/23%/7%	<0,001
Revisão da dispneia	9%/47%/43%	14%/48%/38%	0,069
Revisão das atividades diárias	16%/39%/45%	18%/44%/38%	0,098
Revisão da confiança	50%/35%/15%	58%/31%/12%	0,051
Revisão do sono	58%/21%/21%	70%/16%/13%	<0,001
Revisão do nível de energia	40%/41%/19%	51%/34%/15%	0,003
Revisão das exacerbações	33%/42%/25%	28%/46%/25%	0,265
Revisão do tabagismo	16%/14%/70%	9%/14%/77%	0,018

CAT = Teste de Avaliação da DPOC.

mente mais revisões de “alta qualidade” nos itens que compuseram o subescore B (exceto na revisão do tabagismo e das exacerbações) que os médicos do braço CAT- (Tabela 5).

Para testar se os resultados atípicos do caso 5 (paciente com depressão maior) estariam enviesando os resultados globais, conduziu-se uma análise *post hoc* removendo-se os dados desse caso. O escore global médio aumentou para 20,4 no braço CAT- e 21,3 no braço CAT+ (diferença de 0,93, IC-95% = -0,5 a 2,4, $p = 0,207$). A diferença entre os braços foi reduzida no subescore A para -0,1 ($p = 0,748$) e aumentou no subescore B para 1,1 ($p = 0,010$).

Com exceção do caso 5, o CAT teve pouco impacto no “escore de compreensão” dos médicos ou no seu escore global de avaliação (Figura 4). Globalmente, proporções similares de médicos receberam “escore de compreensão” baixa (19% nos dois braços), aceitável (39% no braço CAT- e 42% no braço CAT+) ou acurada (42% no braço CAT- e 40% no braço CAT+; OR = 1,0, IC-95% = 0,7 a 1,4; $p = 0,824$). De maneira similar, as proporções de médicos com escore de avaliação global ruim/muito ruim (39% nos dois braços) ou boa/muito boa (61% nos dois braços) foram as mesmas (OR = 0,9; IC-95% = 0,6 a 1,4; $p = 0,655$).

Não se observou correlação entre a experiência do médico (em anos de atividade como médico do serviço primário de saúde) e o efeito do CAT em uma análise *post hoc*.

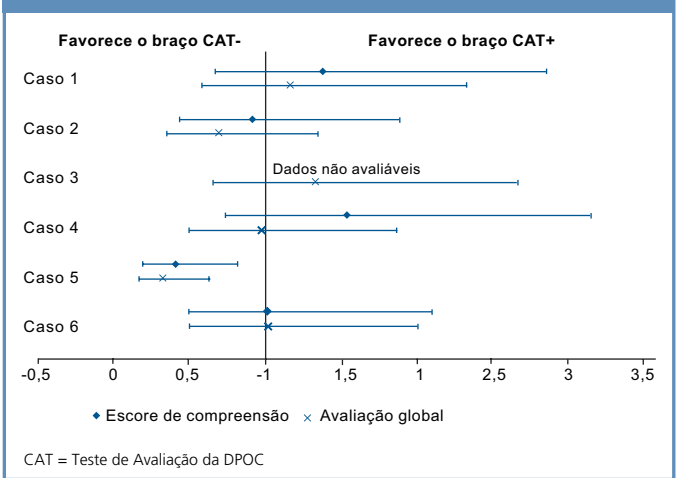
O tempo médio de consulta foi de 9,7 minutos no braço CAT- e de 9,8 minutos no braço CAT+. O comentário devolutivo dos atores foi muito similar entre os médicos dos braços CAT- e CAT+ (dados não apresentados).

Discussão

Principais achados

Nós delineamos este estudo considerando que o CAT poderia estimular uma discussão mais holística com o paciente e permitir a rápida identificação de aspectos fundamentais da DPOC, deixando mais tempo para a revisão de aspectos da condição do paciente não relacionados à DPOC. Não observamos efeito do CAT no desfecho primário do estudo, que incluiu a capacidade do médico de revelar problemas do paciente que podiam não ser específicos da DPOC. Por outro lado, o CAT melhorou de maneira estatisticamente significativa a avaliação

Figura 4. Razão de chances e intervalos de confiança de 95% do “escore de compreensão” e escore de avaliação global nos braços CAT+ e CAT- por tipo de caso clínico



de aspectos padronizados da DPOC (subescore B), embora a relevância clínica dessa melhora possa ser limitada. As revisões dos médicos do braço CAT+ não foram melhores nem piores que as dos médicos do braço CAT- em relação a dois aspectos da DPOC (tabagismo e exacerbações), que não fazem parte do CAT, mas são aspectos fundamentais de uma revisão da DPOC. Esse achado destaca o papel do CAT como uma avaliação não específica para a DPOC que complementa uma avaliação clínica abrangente.

Este estudo demonstra que a introdução do CAT pode ter um efeito específico. Ele não aumenta a conscientização do médico a respeito de aspectos da DPOC que o teste não cobre, nem assegura que os médicos possam detectar melhor fatores subjacentes relacionados à saúde, como a depressão, na condição de um paciente com DPOC. Talvez tenha sido muito otimista incluir no desfecho primário do estudo uma avaliação do desempenho do médico, que não melhorou com o CAT.

Embora o CAT tenha melhorado a avaliação da DPOC no subescore B, não melhorou o “escore de compreensão”. Um motivo para esse achado pode estar nas avaliações de 81% dos médicos que foram consideradas aceitáveis ou acuradas, enquanto somente 19% foram consideradas ruins, de maneira que havia pouco espaço para melhora. Claramente, os médicos são capazes de obter uma boa compreensão do diagnóstico de um paciente sem o CAT, portanto, é improvável que este produza uma grande melhora nesse sentido.

Os comentários objetivos feitos pelos atores foram muito similares nos dois braços, sugerindo uma diferença na percepção do valor do CAT entre os atores e os avaliadores, ou que as questões respondidas pelos atores não foram sensíveis a quaisquer diferenças em sua experiência.

Uma área não abordada por este estudo é se o CAT afetaria as decisões de manejo feitas pelo médico. Esta foi uma limitação necessária, pois não havia diretrizes de manejo com base no CAT, na época em que o estudo foi delineado. Essa situação se modificou, e o CAT atualmente faz parte do algoritmo de avaliação da DPOC de 2011 da Iniciativa Global contra a Doença Pulmonar Obstrutiva Crônica (GOLD), assim como a espirometria e o histórico de exacerbações.⁹

Em todo o estudo, os escores de um dos casos foram melhores no braço CAT- que no braço CAT+. Esse caso (número 5) refere-se a uma senhora idosa que sofria de depressão maior, com sintomas típicos como rebaixamento do humor e anorexia, sua DPOC estava

mal controlada e a paciente não aderiu ao tratamento. Seria possível postular que os médicos do braço CAT+ estariam mais concentrados na DPOC do que nos aspectos mais amplos do caso, mas eles também obtiveram escores bem piores no subescore B. A presença de um questionário específico para a doença pode, portanto, limitar a comunicação natural entre o paciente e o médico. Como os pacientes com comorbidades constituem uma grande proporção da população com DPOC atendida nos serviços primários, os médicos precisam compreender plenamente a DPOC subjacente e também as comorbidades, o que pode requerer consultas mais longas com o seu médico. Nós interrompemos as consultas que continuavam depois de 10 minutos, para manter as pressões das condições reais nos serviços primários,¹⁰ mas essa prática pode ter cerceado as discussões de maneira injusta.

Vantagens e limitações do estudo

A principal vantagem deste estudo é a padronização de casos e as avaliações para assegurar uma comparação simples entre os braços. O estudo foi estritamente controlado, atentando-se para a seleção e o treinamento dos médicos, a inclusão de um estudo-piloto como teste do seu desenho e atividades de controle de qualidade para compreensão e abordagem da variação nos escores dos avaliadores.

A situação artificial das gravações poderia ter afetado o comportamento dos médicos no estudo, o que poderia ter alterado e/ou limitado a interpretação dos resultados na prática clínica diária. O fenômeno da melhora do desempenho de uma pessoa que está sendo observada é reconhecido (efeito de Hawthorne). Embora não seja possível eliminá-lo, um eventual impacto seria similar nos dois braços. O aspecto mais importante foi que o desenho do estudo exigia interação em uma visita única, sem relação ou história entre os dois indivíduos (médico e paciente). Também não houve oportunidade para avaliação do impacto do CAT no manejo e seguimento em longo prazo. Foi interessante observar que os médicos modificaram as suas práticas com a experiência de usar o Questionário Clínico da DPOC.¹¹ Se esse fenômeno for representativo, pode haver um impacto maior do CAT na comunicação médico-paciente ao longo do tempo.

O treinamento sobre o uso do CAT foi oferecido por meio de um material de leitura e da discussão entre os participantes do braço CAT+, como ocorre frequentemente na vida real, presumindo-se que assim os médicos usariam os escores do CAT em suas consultas e decisões. Admite-se agora que essa abordagem foi provavelmente insuficiente, e uma abordagem educacional mais prática poderia encorajar mais as alterações comportamentais dos médicos no braço CAT+.¹² Na verdade, os atores que representaram os pacientes referiram que os médicos não mencionaram o CAT durante a consulta.

Os critérios de avaliação, inclusive os comentários dos pacientes, foram criados especificamente para este estudo, o que constitui uma vantagem e uma desvantagem ao mesmo tempo. Eles permitiram testar a utilidade do CAT nas áreas que acreditamos que o teste poderia ter algum impacto (sem garantia de sucesso), mas não foram testados quanto à confiabilidade ou à sensibilidade. Não podemos ter certeza da relevância clínica de uma diferença estatisticamente significativa.

Os médicos não foram selecionados randomicamente, e a nossa população certamente tem um viés de pertencer mais a clínicas urbanas, em comparação com relatos prévios.¹⁰ Essa condição deve-se a aspectos práticos da condução do estudo: 12 médicos dispostos a participar do estudo precisavam estar a uma distância viável dos locais das gravações. Provavelmente, médicos que atuam em áreas rurais têm habilidades de comunicação similares às dos seus pares que atuam em áreas urbanas, de forma que esperamos que um eventual viés seja razoavelmente pequeno. Os médicos com interesse ou expertise

específico em doenças respiratórias também poderiam gerar um viés. Para limitá-lo, nem os recrutadores nem os médicos conheciam plenamente a natureza do estudo na época do recrutamento.

O nosso estudo foi multinacional, cobrindo várias culturas e sistemas de saúde diferentes. O uso de atores bilíngues, que passaram algum tempo com pacientes locais portadores de DPOC, permitiu que os casos fossem retratados com acurácia cultural. Os casos foram muitas vezes desempenhados pelo mesmo ator em vários países, gerando uma consistência adicional. De maneira similar, os avaliadores da Alemanha e do Reino Unido foram usados para avaliar os médicos da Áustria e da Irlanda, respectivamente. O avaliador foi considerado uma covariável na nossa análise, para remoção de um possível viés decorrente de diferenças entre os avaliadores. Os resultados do estudo podem, portanto, ser considerados razoavelmente passíveis de generalização.

Interpretação dos achados em relação a trabalhos previamente publicados

Até o momento, nenhum outro estudo avaliou o impacto do CAT na interação entre pacientes com DPOC e seu médico do serviço primário de saúde. Na verdade, até onde temos conhecimento, nenhum outro estudo avaliou o impacto de uma PROM em uma única consulta, seja no atendimento primário ou secundário. Contudo, vários projetos de pesquisa sobre intervenções com impacto na consulta paciente-médico já foram publicados.¹³⁻²⁰ Intervenções como encorajar o paciente a desempenhar um papel mais ativo na consulta, o treinamento dos profissionais de saúde em habilidades de comunicação e o fornecimento de mais informações sobre a doença, o tratamento ou a qualidade de vida, já foram avaliadas com sucesso variável.¹⁸ Um dos problemas com esse tipo de avaliação é que a metodologia nem sempre é bem definida e as intervenções frequentemente não estão claramente associadas às medidas do seu provável efeito.^{12,18}

Implicações para futuras pesquisas, políticas de saúde e prática clínica

Apresentamos um novo desenho de estudo baseado em metodologia validada, usada na avaliação do treinamento de especialidades na clínica geral do Reino Unido. Ela mostra-se promissora como método de avaliação do impacto de uma PROM ou de outro instrumento na consulta entre paciente e médico. Contudo, é preciso levar em conta lições aprendidas neste estudo para o uso futuro. Particularmente, é preciso considerar os critérios de avaliação aplicados. Recomendamos incluir atividades de controle de qualidade para uma compreensão e abordagem da variabilidade entre os avaliadores.

Este estudo considera somente a utilidade do CAT em uma única sessão de avaliação. O CAT permite o monitoramento do estado de saúde do paciente com DPOC ao longo do tempo, mas não seria possível realizar um estudo longitudinal desse tipo. O CAT continua a ser avaliado em diferentes cenários, a médio e longo prazo, para definição da sua utilidade e impacto no manejo e nos resultados a longo prazo.²⁰

Conclusões

O CAT funciona como um instrumento específico da doença que ajuda na avaliação da DPOC pelo médico, de acordo com o objetivo para o qual foi criado. Ele não parece melhorar a detecção de sintomas que não são da DPOC e de comorbidades.

Editor executivo: Arnulf Langhammer

Revisão estatística: Gopal Netuveli

Agradecimentos: Os autores gostariam de agradecer à equipe de filmagem,

aos atores, aos assessores e aos médicos que participaram do estudo, a James Dodd pelo apoio no treinamento dos atores e aos pacientes com DPOC que ajudaram a definir o desenho do estudo e a treinar os atores.

Conflitos de interesses: KG-J já atuou como consultor e palestrante em nome da GSK, AZ, Chiesi, Boehringer Ingelheim, MSD, Muni Pharma/Napp, Allmirall, Novartis e Sandoz. SH já recebeu honorários por palestras, fundos de viagem e honorários por participar de conselhos consultivos da AZ, Boehringer Ingelheim, Chiesi, GSK, MSD, Napp, Novartis e Nycomed. PK já recebeu honorários por participar de conselhos consultivos, fundos de viagem e honorários por palestras da AZ, Boehringer Ingelheim, Chiesi, GSK, MSD, Novartis e Nycomed. RE já recebeu honorários por participar de conselhos consultivos, fundos de viagem e honorários por palestras da AZ, Boehringer Ingelheim, Chiesi, GSK, MSD, Novartis e Nycomed. RDN não possui conflitos de interesses em relação a este artigo. JR já recebeu honorários por palestras, fundos de viagem e honorários por participar de conselhos consultivos da AZ, Boehringer Ingelheim, Chiesi, GSK, MSD, Novartis e Teva. PJ já recebeu honorários de empresas farmacêuticas, inclusive da GSK, para proferir palestras em eventos científicos e participar de reuniões de conselhos consultivos, e já recebeu fundos de apoio à pesquisa de empresas farmacêuticas, inclusive da GSK. HCM, GN, MV e DAL são funcionários da GSK.

Contribuições dos autores: Todos os autores deram sugestões para a elaboração do desenho do estudo, desde a definição dos casos dos pacientes e critérios de avaliação até os desfechos e análises estatísticas usados no estudo. Todos os autores revisaram criticamente o artigo. Além disso, KG-J, SH, PK, RE, JR, GN e DAL auxiliaram na avaliação das consultas, PJ auxiliou na descrição dos casos de pacientes com DPOC e MV conduziu a análise estatística.

Financiamento: Os fundos para realização deste estudo foram fornecidos pela GlaxoSmithKline.

Referências

- Coulter A, Ellins J. Patient-focused interventions: a review of the evidence Picker Institute Europe; Health Foundation, 2006.
- Freeman G, Horder JP, Howie JGR, et al. Evolving general practice consultation in Britain: issues of length and context. *BMJ* 2002;**324**:880-2. <http://dx.doi.org/10.1136/bmj.324.7342.880>
- Barry CA, Bradley CP, Britten N, Stevenson FA, Barber N. Patients' unvoiced agendas in general practice consultations: qualitative study. *BMJ* 2000;**320**:1246-50. <http://dx.doi.org/10.1136/bmj.320.7244.1246>
- Jones P, Harding G, Wiklund I, Berry P, Leidy N. Improving the process and outcome of care in COPD: development of a standardised assessment tool. *Prim Care Respir J* 2009;**18**:208-15. <http://dx.doi.org/10.4104/pcrj.2009.00053>
- Jones PW, Harding G, Berry P, Wiklund I, Chen W-H, Leidy NK. Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009;**34**:648-54. <http://dx.doi.org/10.1183/09031936.00102509>
- Jones P, Harding G, Wiklund I, et al. Tests of the responsiveness of the Chronic Obstructive Pulmonary Disease (COPD) Assessment Test™ (CAT) following acute exacerbation and pulmonary rehabilitation. *Chest* 2012. Published online: <http://dx.doi.org/10.1378/chest.11-0309>
- Dodd JW, Hogg L, Nolan J, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax* 2011;**66**:425-9. <http://dx.doi.org/10.1136/thx.2010.156372>
- Gruffydd-Jones K, Marsden H, Holmes S, et al. Novel study design to assess the utility of the COPD Assessment Test in a primary care setting. Presented at European Respiratory Society Annual Congress, Amsterdam, 24-28 September 2011. Abstract P3764.
- Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011. Available from: <http://www.goldcopd.org/>
- Deveugele M, Derese A, van den Brink-Muinen A, Bensing J, de Maeseneer J. Consultation length in general practice: cross sectional study in six European countries. *BMJ* 2002;**325**:472-7. <http://dx.doi.org/10.1136/bmj.325.7362.472>
- Kocks JWH, Kerstjens HAM, Snijders SL, 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. <http://dx.doi.org/10.1186/1477-7525-8-135>
- Greenhalgh J, Long A, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory. *Soc Sci Med* 2005;**60**:833-43. <http://dx.doi.org/10.1016/j.socscimed.2004.06.022>
- Rodin G, Mackay JA, Zimmermann C, et al. Clinician-patient communication: a systematic review. *Support Care Cancer* 2009;**17**:627-44. <http://dx.doi.org/10.1007/s00520-009-0601-y>
- Efrainsson EO, Klang B, Larsson K, Ehrenberg A, Fossum B. Communication and self-management education at nurse-led COPD clinics in primary health care. *Patient Educ Counsel* 2009;**77**:209-17. <http://dx.doi.org/10.1016/j.pec.2009.03.033>
- Mauksch LB, Dugdale DC, Dobson S, Epstein R. Relationship, communication and efficiency in the medical encounter. *Arch Intern Med* 2008;**168**:1387-95. <http://dx.doi.org/10.1001/archinte.168.13.1387>
- Zandbelt LC, Smets EMA, Oort FJ, Godfried MH, de Haes HCJM. Medical specialists' patient-centered communication and patient reported outcomes. *Med Care* 2007;**45**:330-9. <http://dx.doi.org/10.1097/01.mlr.0000250482.07970.5f>
- Francis N, Rollnick S, McCambridge J, Butler C, Lane C, Hood K. When smokers are resistant to change: experimental analysis of the effect of patient resistance on practitioner behaviour. *Addiction* 2005;**100**:1175-82. <http://dx.doi.org/10.1111/j.1360-0443.2005.01124.x>
- Griffin SJ, Kinmonth A-L, Veltman MWM, Gillard S, Grant J, Stewart M. Effect on health-related outcomes of interventions to alter the interaction between patients and practitioners: a systematic review of trials. *Ann Fam Med* 2004;**2**:595-608. <http://dx.doi.org/10.1370/afm.142>
- Detmar SB, Muller MJ, Schornagel JH, Wever LDV, Aaronson NK. Health-related Quality-of-Life assessments and patient-physician communication. A randomised controlled trial. *JAMA* 2002;**288**:3027-34. <http://dx.doi.org/10.1001/jama.288.23.3027>
- Jones PW, Brusselle G, Dal Negro RW, et al. Properties of the COPD assessment test in a cross-sectional European study. *Eur Respir J* 2011;**38**:29-35. <http://dx.doi.org/10.1183/09031936.00177210>

Apêndice 1

CONSORT 10: Lista de verificação de informações a incluir no relato de um estudo randomizado*

Seção/tópico	Nº	Item a verificar	Reportado na página
Título e resumo			
	1a	Identificação como estudo randomizado no título	37
	1b	Sumário estruturado do desenho do estudo, métodos, resultados e conclusões (para orientação específica, consulte o CONSORT para resumos)	37
Introdução			
Fundamentos e objetivos			
	2a	Fundamentos científicos e explicação da justificativa do estudo	37-38
	2b	Objetivos ou hipóteses específicos	38
Métodos			
Desenho do estudo			
	3a	Descrição do desenho do estudo (por exemplo, paralelo ou fatorial), inclusive com a proporção de randomização	38
	3b	Alterações importantes nos métodos depois do início do estudo (como critérios de elegibilidade), com os motivos	N/D
Participantes			
	4a	Critérios de elegibilidade dos participantes	38
	4b	Locais e condições em que os dados foram coletados	38 / Tabela 2
Intervenções			
	5	Intervenções em cada grupo, com detalhes suficientes para permitir uma replicação, inclusive como e quando foram de fato realizadas	38-39
Desfechos			
	6a	Definição completa dos desfechos primário e secundário pré-especificados, inclusive como e quando foram avaliados	39
	6b	Qualquer alteração nos desfechos do estudo após o seu início, com os motivos	N/D
Tamanho da amostra			
	7a	Como foi determinado o tamanho da amostra	N/D
	7b	Quando for o caso, explicar as análises preliminares e os critérios de interrupção	N/D
Randomização:			
Geração da sequência			
	8a	Método usado para gerar a sequência de alocação randômica	38
	8b	Tipo de randomização, detalhes de eventuais restrições (como o uso e o tamanho de blocos de randomização)	38
Mecanismo de ocultação da alocação			
	9	Mecanismo usado para implantar a sequência de alocação randômica (como embalagens numeradas sequencialmente), com descrição das etapas realizadas para ocultação da sequência até a definição das intervenções	38
Implantação			
	10	Quem gerou a sequência de alocação randômica, quem incluiu os participantes e quem atribuiu as intervenções aos participantes	38
Mascaramento			
	11a	Se foi o caso, quem não teve acesso às informações da randomização após a definição das intervenções (por exemplo, os participantes, os profissionais de saúde, os avaliadores dos resultados) e como	38
	11b	Se for relevante, descrever a similaridade das intervenções	N/D
Métodos estatísticos			
	12a	Métodos estatísticos usados para comparar os grupos quanto aos desfechos primário e secundário	39
	12b	Métodos de análises adicionais, como análises de subgrupos e análises ajustadas	39
Resultados			

Apêndice 1 continuação

Distribuição dos participantes (recomenda-se fortemente um fluxograma)	13a	Para cada grupo, números de participantes randomizados, que receberam o tratamento previsto e que foram analisados quanto ao desfecho primário	39, Figura 1
	13b	Para cada grupo, perdas e exclusões após a randomização, assim como os motivos	39, Figura 1
Recrutamento	14a	Datas que definem os períodos de recrutamento e seguimento	39
	14b	Por que o estudo foi encerrado ou interrompido	N/D
Dados iniciais	15	Tabela mostrando dados demográficos e características clínicas iniciais em cada grupo	Tabela 2
Números analisados	16	Para cada grupo, número de participantes (denominador) incluídos em cada análise e se a análise foi feita de acordo com os grupos originariamente definidos no estudo	Sim
Desfechos e estimativa	17a	Para cada desfecho, primário e secundário, os resultados de cada grupo, a magnitude estimada do efeito e sua precisão (por exemplo, com os intervalos de confiança de 95%)	40, Tabela 4, Figura 2
	17b	Em resultados binários, recomenda-se apresentar os valores absolutos e as magnitudes relativas do efeito	N/D
Análises auxiliares	18	Resultados de outras análises realizadas, inclusive análises de subgrupos e análises ajustadas, distinguindo-se as análises pré-especificadas das exploratórias	40-41, Figuras 3 e 4, Tabela 5
Efeitos adversos	19	Todos os efeitos adversos e não pretendidos relevantes ocorridos em cada grupo (para orientação específica, consulte o CONSORT sobre efeitos adversos)	N/D
Discussão			
Limitações	20	Limitações do estudo, discutindo-se as fontes de possíveis vies, imprecisão e, se for o caso, a multiplicidade de análises	42
Possibilidade de generalização	21	Possibilidade de generalização (validade externa, aplicabilidade) dos achados do estudo	42
Interpretação	22	Interpretação coerente com os resultados, ponderando-se benefícios e malefícios, e levando-se em conta outras evidências relevantes	42
Outras informações			
Registro	23	Número e nome do registro do estudo	N/D
Protocolo	24	Onde o protocolo completo do estudo pode ser consultado, se disponível	N/D
Financiamento	25	Fontes de financiamento e outros apoios (como o suprimento de fármacos), papel dos patrocinadores no estudo	43

*Recomendamos fortemente a leitura destas orientações, bem como a Explicação e Elaboração do CONSORT 2010, para importantes esclarecimentos sobre todos os itens. Se for relevante, recomendamos também as orientações estendidas sobre estudos randomizados agrupados, estudos de não inferioridade e de equivalência, tratamentos não farmacológicos, intervenções fitoterápicas e estudos pragmáticos. Outros complementos estão sendo elaborados. Para consultá-los e para obter referências atualizadas relevantes para esta lista de verificação, visite o website www.consort.statement.org

CONSORT 2010 lista de verificação