

# Revisão sistemática da acurácia dos testes diagnósticos: uma revisão narrativa

## *Systematic review of diagnostic tests accuracy: a narrative review*

GLÓRIA MARIA DE OLIVEIRA<sup>1</sup>; FÁBIO TRINCA CAMARGO<sup>2</sup>; EDUARDO COSTA GONÇALVES<sup>3</sup>; CARLOS VINÍCIUS NASCIMENTO DUARTE<sup>4</sup>; CARLOS ALBERTO GUIMARÃES, TCBC-RJ<sup>5</sup>

### R E S U M O

Este artigo tem o objetivo de realizar uma revisão narrativa sobre revisão sistemática da acurácia dos testes diagnósticos. Foi realizada busca na *Cochrane Methodology Reviews* (*Cochrane Reviews of Diagnostic Test Accuracy*), *Medline* e LILACS, bem como busca manual das listas de referências dos artigos incluídos na revisão. As estratégias de busca empregadas foram as seguintes, empregando-se títulos de assuntos e termos livres: 1. na *Cochrane Methodology Reviews*: accuracy study "Methodology" 2. Na *Pubmed* "Meta-Analysis" [Publication Type] AND "Evidence-Based Medicine" [Mesh] AND "Sensitivity and Specificity" [Mesh]; 3. Na LILACS: (revisao sistematica) or "literatura de REVISAO como assunto" [Descritor de assunto] and (sistematica) or "SISTEMATICA" [Descritor de assunto] and (acuracia) or "SENSIBILIDADE e especificidade" [Descritor de assunto]. Em suma, a preparação e o planejamento metodológicos das revisões sistemáticas de testes diagnósticos é ulterior àqueles empregados nas revisões sistemáticas das intervenções terapêuticas. Há muitas fontes de heterogeneidade nos desenhos dos estudos de teste diagnóstico, o que dificulta muito a síntese – metanálise – dos seus resultados. Para contornar esse problema, existem atualmente normas, exigidas pelas principais revistas biomédicas, para a submissão de um manuscrito sobre testes diagnósticos.

**Descritores:** Diagnóstico, metanálise, revisão, literatura de revisão como assunto.

### INTRODUÇÃO

A revisão sistemática da literatura é um estudo secundário que tem por objetivo reunir estudos primários, publicados ou não, avaliando-os criticamente em sua metodologia e reunindo-os numa análise estatística – a metanálise – quando isso é possível<sup>1</sup>.

A acurácia de um teste diagnóstico leva em conta como esse teste pode identificar corretamente as pessoas com uma determinada doença e excluir as pessoas que não têm a doença. Essa acurácia pode ser estimada por meio da comparação dos resultados de um teste, chamado de teste índice, com os resultados de um teste padrão de referência. Os participantes de um estudo de teste diagnóstico devem submeter-se tanto ao teste índice quanto ao teste padrão, para que, em seguida, os resultados dos dois testes sejam comparados.

Quando há uma única condição alvo para testes dicotômicos, a acurácia de um teste diagnóstico é geralmente calculada pela proporção de indivíduos verdadeiramente positivos com aqueles verdadeiramente negativos em relação aos falso positivos e falso negativos.

Recentemente, surgiu uma demanda para revisões sistemáticas – importante ferramenta da Medicina Baseada em Evidências – sobre estudos da acurácia dos testes

diagnósticos. As revisões sistemáticas são desenhadas para focalizar uma questão, utilizando fontes compreensivas, definindo estratégias explícitas de pesquisa, uniformizando a seleção dos artigos, avaliando criticamente e, quando metodologicamente possível, sintetizando de forma quantitativa os resultados. Em outubro de 2008 foi publicada pela primeira vez uma revisão sistemática, sobre a acurácia de um teste diagnóstico, pela Biblioteca Cochrane<sup>2</sup>.

A *Cochrane Reviews of Diagnostic Test Accuracy*, grupo criado para desenvolver e implementar as revisões sistemáticas da acurácia dos testes diagnósticos, dentro da Colaboração Cochrane, procura ser abrangente na identificação da literatura, avaliar a qualidade da investigação e usar métodos estatísticos validados para sintetizar os resultados. Os detalhes técnicos de sua realização variam devido a diferenças na indexação, concepção, execução e análise de estudos da acurácia (3). A busca dos estudos sobre acurácia diagnóstica é mais difícil e trabalhosa do que a de ensaios clínicos aleatorizados e controlados, devido tanto aos relatos inadequados pelos autores desses estudos, e à indexação inconsistente na *Medline* e em outras bases de dados, tornando difícil desenvolver estratégias de busca que sejam ao mesmo tempo específicas e sensíveis na recuperação de todos os artigos<sup>4</sup>.

Trabalho realizado pelo Grupo de Estudos de Revisão Sistemática do Rio de Janeiro (GERS-Rio)<sup>6</sup>

1. Mestranda do Programa de Pós-Graduação em Ciências Cirúrgicas da Faculdade de Medicina da UFRJ; 2. Residente de Clínica Médica do Hospital do Andaraí, Ministério da Saúde, Rio de Janeiro; 3. Residente de Clínica Médica do Hospital do Andaraí, Ministério da Saúde, Rio de Janeiro; 4. Residente de Clínica Médica do Hospital do Andaraí, Ministério da Saúde, Rio de Janeiro; 5. Professor Adjunto do Departamento de Cirurgia – UFRJ; 6. Grupo de Estudos de Revisão Sistemática do Rio de Janeiro (GERS-Rio).

O estabelecimento da acurácia diagnóstica é de fundamental importância para a tomada de decisões na área da saúde, sendo relevante para a escolha dos testes de rastreamento na população, assim como, para prever o curso de diversas terapias ou fornecer alguma indicação do prognóstico do paciente.

O objetivo deste estudo é realizar uma revisão narrativa sobre revisão sistemática da acurácia dos testes diagnósticos.

## MÉTODOS

Foi realizada busca na *Cochrane Methodology Reviews* (*Cochrane Reviews of Diagnostic Test Accuracy*), *Medline* e *LILACS*. Foi realizada a busca manual das listas de referências dos artigos incluídos na revisão. As estratégias de busca empregadas foram: 1. na *Cochrane Methodology Reviews*: accuracy study "Methodology" 2. Na *Pubmed* "Meta-Analysis" [Publication Type] AND "Evidence-Based Medicine" [Mesh] AND "Sensitivity and Specificity" [Mesh]; 2. Na *LILACS*: (revisao sistemática) or "literatura de REVISAO como assunto" [Descritor de assunto] and (sistemática) or "SISTEMATICA" [Descritor de assunto] and (acurácia) or "SENSIBILIDADE e especificidade" [Descritor de assunto].

## HISTÓRICO

No início da década de 70, o prof. Archibald Leman Cochrane (*University of York* do Reino Unido) chamou a atenção para o fato de que muitos problemas de ineficiência do Serviço Nacional de Saúde Britânico dependiam de decisões e práticas pouco fundamentadas, sem suporte dos dados provenientes de bons ensaios clínicos<sup>5</sup>.

No final daquela década, um grupo de epidemiologistas clínicos, formado por David Sackett, Brian Haynes, Peter Tugwell e Victor Neufeld, da *Mac Master University*, Canadá, estabeleceram os princípios do que viria a ser denominado Medicina Baseada em Evidências, isto é, a utilização da Epidemiologia, da Estatística, da Metodologia Científica, e da Informática para a tomada de decisão clínica.

No final de 1992, foi criado o Centro Cochrane do Reino Unido pela *National Health Service Research and Development Programme*, para facilitar e coordenar a preparação e a manutenção das revisões sistemáticas de estudos aleatorizados e controlados em cuidados de saúde.

Em 2003, a Colaboração Cochrane decidiu incluir as revisões sistemáticas de acurácia dos testes diagnósticos na *Cochrane Database of Systematic Reviews* da *Cochrane Library* (*Steering Group Meeting*, Melbourne, abril de 2003). Para permitir isso, um grupo de trabalho foi formado para desenvolver uma metodologia, incluindo um *software* e um manual<sup>6</sup>.

A primeira revisão sistemática sobre acurácia de um teste diagnóstico da Biblioteca Cochrane foi publicada em outubro de 2008 (2), porém, desde 1994 já havia uma

diretriz para avaliação das revisões sistemáticas sobre acurácia dos testes diagnósticos<sup>7</sup>.

## DISCUSSÃO

As revisões sistemáticas da literatura são realizadas com critérios metodológicos rigorosos e são reconhecidas menos propensas a vieses, o que as tornam importantes instrumentos para as decisões relacionadas a condutas de saúde<sup>8</sup>. As revisões sistemáticas para avaliação da acurácia dos testes diagnósticos são realizadas pelas mesmas razões que as de intervenções terapêuticas e envolvem as mesmas etapas das outras revisões sistemáticas: elaboração da pergunta clínica, busca na literatura (localização e seleção dos estudos), sistematização do método de avaliação da qualidade da informação, extração e síntese dos dados, interpretação dos resultados e resolução do cenário clínico.

O teste perfeito seria aquele sem resultado falso positivo ou falso negativo. A avaliação pode se referir a uma doença, ou um estágio da doença. O teste em si é uma intervenção e faz parte dos cuidados de saúde do paciente. Novos testes são adotados na prática clínica por uma série de razões: substituição de um teste existente (em que o novo teste é capaz de reduzir o impacto negativo dos testes já existentes), melhor informação ou informação equivalente por um custo menor, triagem (para decidir se há necessidade de um teste mais caro ou invasivo) ou como um complemento aos testes existentes<sup>9</sup>. Os testes perfeitos normalmente não existem, mas é necessário saber como um teste diferencia os indivíduos que têm a condição alvo daqueles que não a têm. Isso é chamado de acurácia do teste diagnóstico<sup>10</sup>.

A acurácia de um teste diagnóstico deve ser avaliada, comparando seu resultado com um teste "padrão ouro", ou padrão de referência (o melhor disponível). O teste de referência pode ser um único teste, uma combinação de diferentes testes ou o desfecho clínico dos pacientes<sup>11</sup>. Todos os pacientes devem ser submetidos aos dois testes. Os testes podem ser de laboratório, procedimentos cirúrgicos, exame clínico, de imagem ou histopatológico.

A acurácia do teste depende do seu contexto clínico, não sendo uma propriedade fixa do teste em si. É importante compreender não só as características do teste, tais como a sensibilidade e a especificidade, mas também os benefícios e os danos dos resultados dos testes para os subgrupos dos pacientes. A escolha de um teste diagnóstico depende da história médica pessoal, familiar e do contexto em que está inserido para ser determinada sua realização. Nesse cenário clínico, é que devemos solicitar um teste diagnóstico<sup>12</sup>; comparando o seu desempenho dentro de um contexto, teremos uma melhor idéia de como utilizá-lo na prática clínica.

Ao considerar uma revisão sistemática de estudos de acurácia diagnóstica, é importante avaliar se os resultados da revisão serão capazes de fornecer as informações necessárias para a prática clínica<sup>9</sup>.

O principal objetivo de uma revisão sistemática da acurácia dos testes de diagnóstico é encontrar as evi-

dências sobre a acurácia de um teste (sensibilidade, especificidade, razão de verossimilhança, curva ROC-Receiver Operating Characteristic). A sensibilidade é a probabilidade de um teste dar positivo na presença da doença, isto é, avalia a capacidade do teste detectar a doença quando ela está presente. A especificidade é a probabilidade de um teste dar negativo na ausência da doença, isto é, avalia a capacidade de o teste afastar a doença quando ela está ausente. A razão entre a probabilidade de um determinado resultado de um teste diagnóstico em indivíduos portadores da doença e a probabilidade do mesmo resultado em indivíduos sem a doença é chamada razão de verossimilhança. A curva ROC é uma representação gráfica da sensibilidade (eixo Y) em relação a 1 – especificidade (eixo X). Para isso é necessário avaliar criticamente a qualidade dos estudos primários, para verificar a homogeneidade nos resultados entre os estudos e determinar as fontes da eventual heterogeneidade<sup>13</sup>.

A avaliação crítica da qualidade dos estudos selecionados deve ser realizada por pelo menos dois revisores de forma independente. A avaliação metodológica dos estudos primários é frequentemente dificultada pela falta das informações contidas nos artigos. Os métodos da pesquisa, as características da população do estudo e os procedimentos são muitas vezes mal relatados nos estudos primários. Nesses casos, os revisores podem entrar em contato com os autores dos estudos. Em 1999, o *Cochrane Diagnostic and Screening Test Methods Working Group* discutiu esse problema, a baixa qualidade metodológica dos estudos primários, e deu o primeiro passo para sua correção, seguindo a iniciativa do CONSORT (*Consolidated Standards of Reporting Trials*)<sup>14</sup>, criando o grupo STARD (*STAndards for the Reporting of Diagnostic accuracy studies*)<sup>15</sup>.

Esse grupo criou uma lista de 25 itens que devem ser avaliados para elaboração de estudos sobre a acurácia diagnóstica. Avalia título, resumo, palavras-chave, método (participantes, testes, estatística), resultados e discussão. Atualmente, mais de 200 periódicos utilizam o STARD nas normas para instruções aos autores (*International Seminars in Surgical Oncology, Journal of Cardiothoracic Surgery, Journal of Orthopaedic Surgery and Research, Lancet, World Journal of Emergency Surgery*, entre outros). Apesar dessa iniciativa, as revisões sistemáticas

sobre acurácia dos testes diagnósticos têm muitas vezes uma baixa qualidade metodológica. A variação nos desenhos dos estudos, nos padrões de referência, nas fontes de vieses, na multiplicidade de medidas faz da revisão sistemática de estudos de acurácia diagnóstica mais difícil do que as de intervenção que usam ensaios clínicos aleatorizados e controlados<sup>16</sup>. Em 2003, foi desenvolvida uma ferramenta para avaliar os estudos primários de acurácia diagnóstica, o QUADAS (*Quality Assessment of Diagnostic Accuracy Studies*), que consiste em um questionário de 14 itens, formulados como perguntas, cada uma das quais deve ser marcada como sim, não ou pouco clara. O questionário é simples e rápido para ser concluído: analisa o espectro dos pacientes, o critério de seleção, o padrão de referência, a utilização dos testes índice e de referência em todos os pacientes, a descrição detalhada dos dois testes, os vieses da progressão da doença, a interpretação dos resultados por quem não sabia do outro resultado, a aplicabilidade do teste, os estudos retirados e os resultados indeterminados<sup>17</sup>. É um ponto de partida útil para avaliação dos estudos que relatam a sensibilidade e especificidade da acurácia de um teste diagnóstico<sup>18</sup>.

A utilização das revisões sistemáticas sobre acurácia diagnóstica pelo médico tem de considerar a integração da acurácia do teste com o ganho esperado quando é utilizado na prática clínica, comparando o desempenho dos testes com sua utilização dentro de um cenário clínico. Precisamos saber como um novo teste diagnóstico classifica os doentes, particularmente aqueles que têm um risco intermediário de doença. A melhor maneira de avaliarmos é compreender os benefícios clínicos esperados e os danos imputáveis à sua utilização<sup>16</sup>. A análise da decisão tem como objetivo avaliar o valor de um teste clínico, atribuindo pesos para cada consequência possível, para justificar seu uso no cuidado do paciente<sup>19</sup>.

A metodologia das revisões sistemáticas de testes diagnósticos é ulterior àquela empregada nas revisões sistemáticas das intervenções terapêuticas. Observa-se grande heterogeneidade nos desenhos dos estudos de teste diagnóstico, o que dificulta muito a síntese – metanálise – dos seus resultados. Para contornar esse problema, existem atualmente normas, exigidas pelos principais periódicos, para a submissão de um manuscrito sobre teste diagnóstico.

## A B S T R A C T

*The aim of this study is to perform a narrative review of systematic reviews of diagnostic tests accuracy. We undertook a search using The Cochrane Methodology Reviews (Cochrane Reviews of Diagnostic Test Accuracy), Medline and LILACS up to October 2009. Reference lists of included studies were also hand searched. The following search strategy was constructed by using a combination of subject headings and text words: 1. Cochrane Methodology Reviews: accuracy study "Methodology" 2. In Pubmed "Meta-Analysis" [Publication Type] AND "Evidence-Based Medicine" [Mesh] AND "Sensitivity and Specificity" [Mesh] 3. LILACS (revisao sistematica) or "literatura de REVISAO como assunto" [Descritor de assunto] and (sistematica) or "SISTEMATICA" [Descritor de assunto] and (acuracia) or "SENSIBILIDADE e especificidade" [Descritor de assunto]. In summary, the methodological planning and preparation of systematic reviews of therapeutic interventions are prior to that used in systematic reviews of diagnostic tests accuracy. There are more sources of heterogeneity in design of diagnostic test studies, which impair the synthesis – meta-analysis – of the results. To work around this problem, there are currently uniform requirements for diagnostic test manuscripts submitted to leading biomedical journals.*

**Key words:** *Diagnosis, meta-analysis, review, review Literature as Topic.*

## REFERÊNCIAS

1. Centro Cochrane do Brasil [Internet]. Brasil: [acessado em 15 de out 2009]. Disponível em: <http://www.centrocochranedobrasil.org.br/revisoes.html#>
2. Leeflang MM, Debets-Ossenkopp YJ, Visser CE, Scholten RJ, Hooft L, Bijlmer HA, Reitsma JB, Bossuyt PM, Vandenbroucke-Grauls CM. Galactomannan detection for invasive aspergillosis in immunocompromized patients. *Cochrane Database of Systematic Reviews* 2008 Oct 8;(4):CD007394.
3. Cochrane Collaboration; the Diagnostic Test Accuracy Working Group [database on the Internet]. Reino Unido. [acessado em 15 de out 2009]. Disponível em: <http://srdta.cochrane.org/en/index.html>.
4. Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately? *Health Info Libr J.* 2007;24(3):188-92.
5. Cochrane AL. Effectiveness and Efficiency. Random Reflections on Health Services. London: Nuffield Provincial Hospitals Trust;1972.
6. Deeks J, Gatsonis C, Bossuyt P, Antes G. Cochrane reviews of diagnostic test accuracy. A progress report. [Internet]. *Cochrane News*.2004 Aug Jan [acessado em 22 out 2009]. Disponível em: <http://www.cochrane.org/newslett/ccnews31-lowres.pdf>
7. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med.* 1994 Apr 15;120(8):667-76.
8. Clarke M, Horton R. Bringing it all together: Lancet-Cochrane collaborate on systematic reviews. *Lancet.* 2001 Jun 2;357(9270):1728.
9. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care* [Internet]. York: CRD; 2009 Jan [acessado em 22 out 2009]. 294 p. Disponível em: [http://www.york.ac.uk/inst/crd/pdf/Systematic\\_Reviews.pdf](http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)
10. Cochrane Collaboration; the Diagnostic Test Accuracy Working Group [database on the Internet]. Reino Unido: [acessado em 15 de out 2009]. Disponível em: [http://www3.interscience.wiley.com/homepages/106568753/what\\_diagnostic.pdf](http://www3.interscience.wiley.com/homepages/106568753/what_diagnostic.pdf)
11. Devillé WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, Bezemer DP. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol.* 2002 Jul 3;2:9.
12. Cornell J, Mulrow CD, Localio AR. Diagnostic Test Accuracy and Clinical Decision Making. *Ann Intern Med.* 2008;149:904-6.
13. Pai M, McCulloch M, Enanoria W, Colford Jr JM. Systematic reviews of diagnostic test evaluations: What's behind the scenes? *ACP J Club.* [Internet]. 2004 Jul-Aug;141(1): A11-3. [acessado em 15 de out de 2009]. Disponível em: [http://src.tums.ac.ir/SystematicReview/Assets/Pai\\_ACPJC\\_2004\\_Editorial.pdf](http://src.tums.ac.ir/SystematicReview/Assets/Pai_ACPJC_2004_Editorial.pdf)
14. CONSORT Statement [Internet]. [acessado em 15 de out de 2009]. Disponível em: <http://www.consort-statement.org/>
15. STARD Statement [Internet]. [acessado em 15 de out de 2009]. Disponível em: <http://www.stard-statement.org/>
16. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med.* 2008 Dec 16;149:889-97.
17. Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* [Internet]. 2003 Nov 10;3:25. Reino Unido:[acessado em 15 de out de 2009]. Disponível em: <http://www.biomedcentral.com/1471-2288/3/25>
18. Meads CA, Davenport CF. Quality assessment of diagnostic before-after studies: development of methodology in the context of a systematic review. *BMC Med Res Methodol.* [Internet]. 2009 Jan 19;9:3. Reino Unido:[acessado em 15 de out de 2009]. Disponível em: <http://www.biomedcentral.com/1471-2288/9/3>
19. Vickers AJ. Decision Analysis for the Evaluation of Diagnostic Tests, Prediction Models, and Molecular Markers. *The American Statistician.* 2008 November 1, 62(4): 314-20.

Recebido em 15/11/2009

Aceito para publicação em 15/01/2010

Conflito de interesse: nenhum

Fontes de financiamento: nenhuma

### Como citar esse artigo:

Oliveira GM, Camargo FT, Gonçalves EC, Duarte CVN, Guimarães, CA. Revisão sistemática da acurácia dos testes diagnósticos: uma revisão narrativa. *Rev Col Bras Cir.* [periódico na Internet] 2010; 37(2). Disponível em URL: <http://www.scielo.br/rcbc>

### Endereço para correspondência:

Glória Maria de Oliveira

E-mail: [gloriamo1@yahoo.com.br](mailto:gloriamo1@yahoo.com.br)