



## ARTIGO DE REVISÃO

### Revisão sistemática e metanálise com enfoque na acurácia de testes diagnósticos

#### *Systematic reviews and meta-analysis of studies evaluating the accuracy of diagnostic tests*

Lídia Rosi Medeiros<sup>1</sup>, Priscyla Waleska Azevedo Simões<sup>2,3</sup>, Kristian Madeira<sup>2,3</sup>, Napoleão Chiaramonte Silva<sup>3</sup>, Suéli Lumertz<sup>3</sup>, Maria Inês da Rosa<sup>2,3</sup>

#### Resumo

Uma revisão sistemática (RS) é um resumo da literatura médica usando métodos definidos para pesquisar os estudos na literatura, fazendo análise crítica dos artigos individualmente e, após, usando-se técnica estatística adequada, combinar todos esses artigos selecionados pela estratégia de busca. O entendimento de uma RS e como implementar os achados na prática tem grande importância para todos profissionais envolvidos no cuidado da saúde. O objetivo desse estudo foi auxiliar aos pesquisadores a um melhor entendimento de uma RS de enfoque diagnóstico e entender como planejar e conduzi-la. RS para avaliar testes diagnósticos são realizadas para estimar a performance dos testes e impacto baseado na evidência dos estudos. Revisões de enfoque diagnóstico envolvem etapas como as de enfoque de tratamento, como: busca da literatura, avaliação da elegibilidade e qualidade dos estudos e extração e síntese dos dados.

**Descritores:** Revisão sistemática. Testes diagnósticos. Metanálise.

#### Abstract

A systematic review is a summary of the medical literature that uses explicit methods to perform a thorough literature search and critical appraisal of individual studies and that uses appropriate statistical techniques to combine these valid studies. An understanding and systematic reviews and how to implement them in practice is of extraordinary importance for all professionals involved in the delivery of healthcare. The objective this study was to help people reach a good understanding systematic reviews of diagnostic tests, and to learn how to plan and conduct a systematic review. Systematic reviews of evaluations of tests are undertaken to produce estimates of test performance and impact based on all available evidence. Reviews of studies of diagnostic accuracy involve the same key stages as systematic reviews of treatment interventions: defining questions, searching the literature, evaluating studies for eligibility and quality, and extracting and synthesizing data.

**Key words:** Systematic review. Diagnostic tests. Meta-analysis.

1. Programa de Pós-Graduação em Ciência Médicas – UFRGS, RS.

2. Programa de Pós-Graduação em Ciências da Saúde Universidade do Extremo Sul Catarinense (UNESC) - Criciúma- SC.

3. Laboratório de Epidemiologia –UNESC, SC.

## Introdução

A revisão sistemática é um tipo de estudo secundário que facilita a elaboração de diretrizes clínicas, sendo extremamente útil para os tomadores de decisão na área de saúde, entre os quais estão os médicos e administradores de saúde, tanto do setor público como do privado. Além disso, as revisões sistemáticas (RS) também constituem método moderno para avaliação de um conjunto de dados simultâneos e permite ao pesquisador distinguir um tratamento eficiente daquele que não o é, resolver controvérsias em condutas e determinar terapêuticas que devam ser implementadas<sup>1,2,3</sup>. Os protocolos de condutas (*guidelines*) estão preferencialmente baseados em estudos de revisões sistemáticas<sup>4</sup>, haja vista integrarem, de forma eficiente e racional, informações e dados que auxiliam em tomadas de decisões na área da saúde<sup>3,4,5</sup>.

A RS envolve aplicação de estratégias científicas para sua elaboração, utilizando métodos específicos sistemáticos para identificar, selecionar e avaliar pesquisas relevantes, bem como coletar e analisar dados de estudos incluídos na revisão, tendo como objetivo principal limitar os vieses, devendo responder a uma pergunta específica que suscite controvérsia<sup>4,6,7</sup>.

Para tanto, reúne, de forma organizada e crítica, grande quantidade de resultados de pesquisas clínicas<sup>4</sup>. Segundo esses mesmos autores, quando os resultados de estudos primários são sumarizados, mas não estatisticamente combinados, a revisão é chamada de "revisão sistemática qualitativa"<sup>7</sup>. Já a que utiliza métodos estatísticos para sumarizar seus resultados é denominada de "revisão sistemática quantitativa" ou metanálise<sup>4,8,9</sup>.

Nos trabalhos de RS com enfoque terapêutico, devem-se escolher os estudos randomizados, pois esse tipo de delineamento tipicamente mensura e compara diferentes eventos que estão presentes ou ausentes após os participantes receberem determinado tipo de intervenção<sup>4,10</sup>. As revisões sistemáticas de estudos observacionais, que avaliam a acurácia de testes diagnósticos, genericamente denominadas RS diagnósticas estão aparecendo mais frequentemente na literatura médica<sup>11</sup>, pois se entende que a escolha do melhor método disponível é fundamental para um diagnóstico correto. Vale lembrar que sem um diagnóstico não é possível propor tratamento ou estabelecer prognóstico.

O objetivo dessa revisão foi estudar elementos fundamentais envolvidos numa revisão sistemática e metanálise de testes diagnósticos.

## Revisão sistemática com enfoque diagnóstico

### Histórico e definição

Testes diagnósticos são usados rotineiramente na medicina para rastreamento, diagnóstico e monitorização da progressão de uma doença. Informações diagnósticas são obtidas de diversas maneiras, incluindo história clínica, exame físico, análises bioquímicas, imagens e avaliações histológicas<sup>12,13,14</sup>.

Estudos de RS com enfoque diagnóstico podem contribuir para avaliar o impacto de determinados testes diagnósticos e são essenciais para ajudar a escolher testes com ótima acurácia e que proporcionem apropriada interpretação dos resultados<sup>15</sup>.

RSs com enfoque diagnóstico são realizadas pela mesma razão com que são feitas as revisões com enfoque de intervenção: para produzir estimativas baseadas nas melhores evidências no assunto e para avaliar a variação dos achados entre os artigos<sup>16,17,18</sup>.

### Roteiro para Leitura de uma Revisão Sistemática de enfoque Diagnóstico (Quadro 1)

#### 1. Questão de pesquisa e objetivo do estudo:

Deve-se definir claramente a pergunta que se quer responder e traçar os objetivos geral e específicos, definindo o teste diagnóstico que será estudado e o padrão-ouro<sup>16</sup>.

#### 2. Definição dos critérios de inclusão e exclusão.

Definir o(s) desfecho(s) a serem estudados, qual o teste diagnóstico, definindo qual a metodologia laboratorial ou outra a ser empregada para determinar a positividade do exame e qual o padrão-ouro usado para definir doentes ou não doentes para o desfecho objeto do estudo<sup>3</sup>.

O delineamento dos estudos e qualidade metodológica também devem ser pré-definidos. Informações detalhadas sobre os participantes é fundamental para a precisão. O ideal é que entrem estudos em todas as línguas, para evitar viés de seleção.

#### 3. Estratégias de busca

No preparo de uma RS não existe uma única fonte de busca de trabalhos, havendo necessidade de pesquisarem-se estudos relevantes em bancos de dados eletrônicos (PUBMED, EMBASE, LILACS, CANCERLIT, CINAHL e Cochrane Library). verificar as referências bibliográficas das publicações relevantes, solicitar estudos a colegas e pesquisar manualmente revistas e anais de congressos, sendo que para cada fonte utilizada

deve-se detalhar o método adotado<sup>1,2,4,6</sup>. Normalmente são utilizados os seguintes termos de busca para RS com enfoque diagnóstico<sup>18</sup>: Medical Subject Headings (MeSH terms): explode "sensitivity and specificity"/; explode "mass screening"; "predictive-value-of teste"; "ROC-curve"; specificity\*; false negative\*; accuracy; screening; sensitivity; predictive value\*; likelihood ratio\*; diagnosis\* As palavras-chave acima são associadas às palavras-chave do que se deseja pesquisar, como, por exemplo, tumoração ovariana (sub-headings): Ovarian cysts ovar\*[tw] AND tumor\*[tw] ovar\*[tw] AND cancer [tw]; Adnexal diseases [mh]; Ovarian neoplasms [mh]; Pelvic\*[tw] AND tumor\* [tw]; Pelvic\*[tw] AND masses\*[tw]; "CA 125"; "Serum CA125"; frozen section; "color"[mh] AND Doppler [all fields]; colour [tw] AND Doppler [all fields]; "ultrasonography"[Tw] AND Doppler [all fields]; "color" [mh] AND transvaginal[All Fields]; flow[All Fields] AND Doppler[All Fields].

#### 4. Seleção dos estudos

Deve ser feito por mais de um observador e desenvolver-se estratégias para resolver as discordâncias. A princípio faz-se um rastreamento pelos títulos e resumos. Após consenso devem ser lidos todos os artigos selecionados para certificar-se que preenchem os critérios de inclusão, sendo excluídos os outros. Deve-se informar os motivos pelos quais os artigos foram excluídos<sup>19</sup>.

#### 5. Avaliação da qualidade metodológica

Deve-se avaliar os artigos em sua qualidade metodológica para elegibilidade, como também fazer uma lista de dados importantes extraídos de cada artigo (checklist), após escolher a melhor maneira para efetuar à síntese dos dados a serem analisados<sup>18</sup>.

Os estudos de acurácia diagnóstica diferem, em seu desenho clínico, dos estudos de intervenção, o que significa que se deve ter outra metodologia para avaliação da qualidade e do potencial viés de cada estudo. Com este intuito foi criado o STARD (Standards for Reporting of Diagnostic Accuracy) um checklist com 25 itens que auxiliam na avaliação metodológica dos artigos com enfoque diagnóstico<sup>20</sup>. Com a mesma intenção, existem mais duas publicações, extremamente importantes para quem trabalha com RS com enfoque diagnóstico, que proporcionam ao pesquisador respostas mais precisas na avaliação dos artigos com enfoque diagnóstico tendo como finalidade avaliar a inclusão do artigo ou não na RS<sup>20,21</sup>.

Uma RS com enfoque diagnóstico deve excluir estudos que não possuem os critérios de qualidade metodológica e que são mais suscetíveis a vieses, ou então incluir esses estudos e explorar as características e as diferenças entre os mesmos<sup>18</sup>. O poder estatístico é raramente discutido em estudos com enfoque diagnósti-

co, por não compararem dois grupos e não possuírem formalmente uma hipótese diagnóstica. Entretanto, o aumento do tamanho da amostra no somatório geral dos resultados aumenta a precisão da estimativa do teste diagnóstico<sup>18</sup>.

Estudos de teste de acurácia comparam os resultados entre dois grupos, os pacientes com a doença e os sem a doença, sendo que estes podem apresentar teste diagnóstico positivo ou negativo, sugerindo ter a doença ou não. O resultado do teste diagnóstico é comparado com exame considerado de referência para determinada doença. Tem-se, então, a mensuração da sensibilidade, da especificidade e dos resultados falso-positivos e falso-negativos. A partir destes dados pode-se calcular a razão de verossimilhança positiva e negativa, o valor preditivo positivo (probabilidade pós-teste positiva) e o valor preditivo negativo, bem como diagnostic odds ratio (DOR)<sup>22</sup>.

Estudos ideais para inclusão de uma RS com enfoque diagnóstico são aqueles em que os pacientes, recrutados de uma população ampla, sejam selecionados de maneira consecutiva e cegados tanto o pesquisador quanto os pacientes. Entretanto, geralmente encontram-se estudos não cegados, retrospectivos, com pacientes oriundos de uma população restrita<sup>21</sup>. É de extrema importância que as amostras dos estudos sejam selecionadas de grupos de saúde com aspectos semelhantes<sup>15</sup>.

Todos os estudos devem ter o mesmo teste diagnóstico como teste de referência (padrão-ouro), sendo a referência padrão em todos os estudos selecionados para a RS com enfoque diagnóstico. É imprescindível que os dois testes, o da pesquisa e o de referência, sejam mensurados de maneira independentes<sup>15</sup>.

Outro aspecto importante quando se inclui um estudo é que ele deva ter descrição clara do teste de referência e do teste experimental, que é o enfoque do estudo de acurácia diagnóstica. Devem estar descritos os desfechos positivos e negativos, as características demográficas da população em estudo, as co-morbidades e o histórico de cada paciente<sup>15</sup>.

A qualidade metodológica dos artigos por recomendação Cochrane Library é a quality assessment tool for diagnostic accuracy studies, (QUADAS). A versão modificada é constituído por 11 itens sobre as características do estudo com o potencial de introduzir viés. Os itens são classificados como positivos (sem viés), negativo (com viés), ou informação insuficiente (não claro).<sup>23</sup>

#### 6. Extração dos dados

Dois revisores independentes devem extrair os da-

dos dos estudos primários. Detalhes completos dos participantes deve ser extraídos. A coleta dos dados é a etapa da RS na qual deve ser relacionado um conjunto de informações sobre cada estudo que foi selecionado. A sistematização é importante para evitar viés e para conferir reprodutibilidade à revisão. A coleta dos dados deve ser feita por meio de formulários padronizados, e dela devem constar alguns componentes fundamentais: informações sobre o estudo, revisores e características dos estudos primários. Extraí-se os dados que possibilitem montar a tabela 2x2, ou seja: Verdadeiros positivos, Falsos Positivos, Verdadeiros Negativos e Falsos Negativos, para cada desfecho determinado nos objetivos do estudo<sup>15</sup>.

### 7. Análise de dados - Metanálise nos estudos de acurácia diagnóstica

A metanálise nos estudos de acurácia diagnóstica deve ser considerada somente quando os estudos tenham sido recrutados de população clinicamente semelhante, usando o mesmo teste diagnóstico comparando-o com o teste considerado padrão-ouro e preferencialmente que os estudos sejam prospectivos e cegados. Embora, muitas vezes, estes critérios sejam obedecidos na seleção dos artigos, ocorre heterogeneidade dentro da metanálise, porque nem todos os artigos apresentam alto nível de evidência científica<sup>15,21,24</sup>. Conforme o grau de heterogeneidade encontrado na metanálise e conforme os testes estatísticos realizados na investigação da heterogeneidade, muitas vezes torna-se impossível sumarizar o resultado da RS em dados únicos, por exemplo, para: sensibilidade, especificidade e DOR<sup>15</sup>.

A metanálise passa por dois tipos de processo para realizar os somatórios dos resultados dos estudos de teste de acurácia. Inicialmente deve-se elaborar em cada estudo uma tabela 2 x 2 e depois realizar, através de sistema de computação, com o software específico Meta-DiSc® (version Beta 1.1.1), o sumário estatístico de todos os estudos incluídos na revisão sistemática<sup>15,25</sup>.

A escolha do método estatístico para combinar os estudos depende da heterogeneidade observada entre os estudos incluídos na metanálise. O grau de heterogeneidade pode ser observado pelos gráficos que sumarizam a sensibilidade e a especificidade a partir dos resultados encontrados em cada estudo e através do sumário da curva ROC (Receiver operator characteristic). Divergências entre os estudos são esperadas, mas, conforme o tipo de seleção dos pacientes e o desenho clínico de cada estudo, a variabilidade ou heterogeneidade pode assumir valores importantes e deve ser investigada através de cálculos estatísticos específicos<sup>15</sup>. Essa variabilidade leva à introdução da variação no li-

miar do teste diagnóstico. Os estudos então incluídos na RS podem usar diferentes limiares para definir se o resultado do teste é positivo ou negativo.

#### *Somatório da sensibilidade e especificidade na metanálise diagnóstica - Verificando heterogeneidade*

Um método simples é combinar estudos de acurácia diagnóstica e computar o peso das médias da sensibilidade e da especificidade de cada estudo. Entretanto, esse método só deve ser utilizado se não houver variabilidade no limiar do teste diagnóstico. Quando existe heterogeneidade, ela pode ser investigada através da curva ROC. A homogeneidade da sensibilidade e especificidade pode ser avaliada pelo teste QT (Cochran) distribuição do  $\chi^2$  com 1 grau de liberdade. E o cálculo da correlação entre sensibilidade e especificidade é realizado através do teste de correlação de Spearman. Quando não existe correlação, o somatório das sensibilidades e das especificidades é calculado da seguinte forma, tendo como resultado dois tipos de categoria (positiva e negativa)<sup>18</sup>.

$$\text{sensibilidade } p_i = \frac{Y_i}{n_i} ; \text{ especificidade } p_i = \frac{Y_i}{n_i}$$

Para todos os estudos se calcula a sensibilidade e a especificidade, para, após, sumarizar o resultado usando aproximação com inverso da variância. A estimativa total para sensibilidade e especificidade é calculada de acordo com a seguinte fórmula<sup>15</sup>.

$$p = \frac{\sum Y_i}{\sum n_i}$$

Onde  $\sum Y_i$  é o somatório dos verdadeiros positivos (sensibilidade) ou somatório dos verdadeiros negativos (especificidade), e  $\sum n_i$  é o somatório de doentes (sensibilidade) ou não doentes (especificidade)<sup>15</sup>. O erro padrão é computado através da fórmula<sup>15</sup>:

$$SE(p) = \sqrt{\frac{p(1-p)}{\sum n_i}}$$

#### *Decidindo sobre o modelo estatístico*

Se os cálculos acima mostram homogeneidade e ausência de correlação entre sensibilidade e especifi-

cidade, usam-se modelos de efeitos fixos para encontrar o sumário da sensibilidade e da especificidade bem como da DOR<sup>15</sup>. Entretanto, quando existir heterogeneidade pelo teste QT (Cochran) distribuição do  $\chi^2$  com 1 grau de liberdade, mas a correlação de Spearman for negativa entre especificidade e sensibilidade<sup>15</sup>. Os cálculos sumarizados da sensibilidade, especificidade e da DOR podem ser apresentados realizando-se os cálculos através do modelo de afeitos aleatórios<sup>15</sup>. Neste caso, o cálculo para sumarizar os dados da sensibilidade e da especificidade é feito, por exemplo, utilizando-se o valor da sensibilidade de cada estudo e multiplicando-se o valor pelo número de sujeitos de cada estudo; após realiza-se o somatório dos valores acima descritos e divide-se pelo número total dos sujeitos de todos os  $Y_i$   $Y_i$ ; sensibilidade  $p_i =$ ; especificidade  $p_i$   $n_i$   $\sum Y_i$   $p = \sum n_i$   $p(1-p)$   $SE(p) = \sum n_i \sqrt{2}$  estudos. A mesma fórmula utiliza-se para cálculo da especificidade<sup>26</sup>.

$$\Sigma[(\text{sensibilidade}_{ij} \times n_j) / (n_1 + n_2 + n_3 + \dots + n_j)]$$

*DOR (diagnostic Odds Ratio)*

A DOR é uma conveniente mensuração quando se combinam estudos de uma RS. Resume acurácia diagnóstica descrevendo muito bem como o teste sumariza os dados da doença e de quem não tem a doença (sensibilidade e especificidade). Entretanto, é difícil aplicar-se diretamente na prática clínica. Uma maneira simples de sumarizar-se a DOR é através da curva ROC que corresponde à sensibilidade e à especificidade encontradas na metanálise com enfoque diagnóstico<sup>15,18</sup>. O valor da DOR igual a 1 indica que o teste não tem poder discriminatório, sendo que quanto maior for o valor da DOR maior será o grau de relevância do teste diagnóstico.

$$DOR = \frac{\text{sensibilidade}}{1 - \text{Sensibilidade}} / \frac{1 - \text{especificidade}}{\text{especificidade}}$$

Outra maneira simples de ser calculada a DOR é com o uso de tabela 2 x 2. No caso de alguma casela conter valor 0, adiciona-se nesta nela 0,5 e em todas as demais também realiza-se essa adição, para evitar problemas no cálculo final<sup>15,18,27</sup>.

$$DOR = \frac{\text{verdadeiros positivos} \times \text{verdadeiros negativos}}{\text{falso-positivos} \times \text{falso-negativos}} = \frac{A \times D}{B \times C}$$

Quando se observa heterogeneidade entre os es-

tudos, observa-se concomitante variação no limiar do teste diagnóstico, e com isso as estimativas de sumário de especificidade e sensibilidade são superestimadas. Nesta situação, o apropriado sumário da metanálise, não é somente um simples ponto na curva ROC, mas o sumário da curva ROC (SROC) que deve ser analisada como um todo, sendo um balanço final entre a sensibilidade e a especificidade. Com a curva ROC pode-se também estimar a área sob a curva, conhecida como AUC (area under the curve) que pode sumarizar a capacidade do teste discriminatório para doença. Testes perfeitos usualmente possuem AUCs próximos ao valor 1, e testes considerados não satisfatórios possuem valor igual ou inferior a 0,5. A área sob a curva pode sumarizar a capacidade inerente de um teste diagnóstico ou de um biomarcador para discriminar doença de não doença<sup>27, 28, 29</sup>.

A curva ROC assimétrica ocorre quando a DOR muda com o limiar diagnóstico da sensibilidade e da especificidade. Para estimar-se o sumário da curva ROC usa-se um modelo linear de Littenberg e Moses com a seguinte fórmula<sup>27, 28, 29</sup>:

$$D = a + \beta S$$

$$D = \ln \text{DOR}$$

$$S = \text{logiti verdadeiros positivos} + \text{logiti de falso-positivos}$$

$$A = \ln \text{OR (intercepto)}$$

$$\beta = \text{examina quanto a OR depende do valor da variação do limiar}$$

Caso o coeficiente de regressão fique próximo do zero, a acurácia de cada estudo primário pode ser sumarizada por OR, dada pelo intercepto. Pode-se avaliar também de outra forma, caso o valor de p do intercepto  $\beta$  seja menor que 0,05, ou com valor muito próximo a este, o que sugere que a DOR muda com o limiar, ou seja, não se pode ter um valor único que represente aquele teste diagnóstico. Entretanto, se o valor de p do intercepto  $\beta$  for maior que 0,1 isso sugere que a DOR não modifica seu limiar, mesmo havendo heterogeneidade entre os estudos, e pode-se ter um valor único para DOR (Tabela 1)<sup>27,29,30</sup>. A Figura 1, demonstra a curva ROC correspondente a esse cálculo.

**8. Interpretação dos resultados**

Considerar as limitações, relatando vieses, fazer inferências sobre evidências de aplicabilidade do teste. Verificar se existe evidências ou não da acurácia do teste em estudo, se existe evidência se é suficiente com poder estatístico e se deve ser usado ou não.

Para redigir uma metanálise de estudos observacionais criteriosa, deve-se obedecer ao checklist de recomendações, MOOSE (Meta-analysis of Observational

studies in Epidemiology), consenso de um Workshop de expertises realizado em Atlanta em 1997<sup>32</sup>.

## Discussão

As RSs para avaliação de testes diagnósticos são realizadas pelas mesmas razões das revisões com enfoque de tratamento, ou seja, para estimar a performance e o impacto dos testes baseado em todas as evidências disponíveis na literatura, de estudos com a qualidade metodológica avaliada com uma meta avaliação estatística de todos os resultados encontrados pelos autores selecionados para compor a revisão sistemática, controlando-se a heterogeneidade dos estudos.

A discussão dos achados, é fundamental, pois os pesquisadores que realizaram a RS, assim como os autores de um artigo original, são as pessoas mais aptas a discutir os achados. O(s) desfecho(s) que foram considerados na RS devem ser claramente identificados. A seguir deve ser feito um balanço dos efeitos benéficos e os potenciais riscos do teste estudado<sup>33</sup>.

As pessoas, em diferentes situações, utilizam as RSs publicadas. Por isso é importante que sejam feitas com critério, seriedade e responsabilidade, mencionando todos os fatores que poderiam implicar na aplicabilidade e as limitações do estudo na discussão.

Existem muitos fatores que podem influenciar nos estudos incluídos na revisão, como a heterogeneidade entre os estudos, que podem ser minimizados com a escolha adequada do método estatístico para análise dos dados.

A discussão também pode contribuir, incluindo alguma informação sobre a frequência do problema de saúde ao qual se está direcionando a RS.

Deve-se ter o máximo critério na interpretação dos resultados e nas inferências a serem feitas, sempre fazendo-se a seguinte pergunta: Esses achados podem contribuir para agregar algum conhecimento relevante? Se a resposta for provavelmente não, deve-se ter muito cuidado nas inferências finais<sup>33</sup>.

O primeiro componente da conclusão será colocar todas as evidências encontradas na revisão, colocando-se a conclusão e a confiança da interpretação realizada com todos os desfechos estudados. É necessário refletir a estimativa de custo, benefícios e riscos dos achados, pois a partir desses achados resulta a tomada de decisão clínica.

Alternativamente pode-se concluir dizendo que não há evidências suficientes para concluir, sugerindo que futuros estudos devem ser realizados para responder a pergunta norteadora da RS realizada<sup>33</sup>.

A medicina tem evoluído com uma velocidade espantosa e práticas atuais são substituídas por novas muito mais rapidamente nos últimos anos. A grande quantidade de informações sobre cuidados com a saúde, diariamente publicada e o tempo limitado dos profissionais da área da saúde trazem a necessidade do desenvolvimento de processos que proporcionem caminhos concisos para obtenção de uma resposta mais rápida aos seus questionamentos, por isso observa-se um número crescente de publicações onde usa-se como desenho do estudo a revisão sistemática com ou sem metanálise, que organiza e resume criteriosamente o que há de importante na literatura mundial.

## Referências

1. Egger M, Smith G D, Phillips A N. Meta-analysis: principles and procedures. *BMJ* 1997; (315): 1533-1537.
2. Mulrow C D, Cook D J, Davidoff F. Systematic reviews: critical links in the great chain of evidence. *Ann Intern Med* 1997; (126): 389-391.
3. Walter S D, Jadad A R. Meta-analysis of screening data: a survey of the literature. *Stat Med* 1999; (18): 3409-3424.
4. Cook D J, Mulrow C D, Haynes R B. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997; (126): 376-380.
5. Mulrow C D. Rationale for systematic reviews. *BMJ* 1994; (309): 597-599.
6. Oxman A D, Cook D J, Guyatt G H. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA* 1994; (272): 1367-1371.
7. Egger M, Smith G D, Schneider M. Systematic reviews of observational studies. In: *Systematic reviews in Health Care - Meta-analysis in context*. (Eds. Egger M, Smith G D, Schneider M). London: BMJ Publishing, 2001; 2nd: 211-227.
8. Egger M, Smith G D. Misleading meta-analysis. *BMJ* 1995; (310): 752-754.
9. Egger M, Ebrahim S, Smith G D. Where now for meta-analysis? *Int J Epidemiol* 2002; (31): 1-
10. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; (282): 1054-1060
11. Irwig L, Tosteson A N, Gatsonis C, Lau J, Colditz G, Chalmers T C, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994; (120): 667-676.

12. Sackett D L. Clinical epidemiology. *Am J Epidemiol* 1969; (89): 125-128. 3 Referências Bibliográficas 68
13. Sackett D L. Clinical diagnosis and the clinical laboratory. *Clin Invest Med* 1978; (1): 37-43.
14. Sackett D L. Clinical epidemiology. what, who, and whither. *J Clin Epidemiol* 2002; (55): 1161-1166.
15. Deeks J J. Systematic reviews of evaluations of diagnostic and screening tests. In: *Systematic reviews in health care - Meta-analysis in context*. (Eds. Egger M, Smit DG, Altman DG). London: BMJ Publishing Group, 2001; 248-282.
16. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995; (48): 119-130.
17. Vamvakas E C. Meta-analyses of studies of the diagnostic accuracy of laboratory tests: a review of the concepts and methods. *Arch Pathol Lab Med* 1998; (122): 675-686.
18. Deeks J J. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001b; (323): 157-162.
19. Egger M, Smith G D, Schneider M. Systematic reviews of observational studies. In: *Systematic reviews in Health Care - Meta-analysis in context*. (Eds. Egger M, Smith G D, Schneider M). London: BMJ Publishing, 2001; 2nd: 211-227.
20. Bossuyt P M, Reitsma J B, Bruns D E, Gatsonis C A, Glasziou P P, Irwig L M, Lijmer J G, Moher D, Rennie D, de Vet H C. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; (138): 40-44.
21. Lijmer J G, Mol B W, Heisterkamp S, Bossel G J, Prins M H, van der Meulen J H, Bossuyt P M. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; (282): 1061-1066.
22. Go A S. Refinando a probabilidade: introdução à solicitação de exames complementares. In: *Medicina baseada em evidências*. (Eds. Friedland DJ, Go A S, Davoren J B, Shlipak M G, Bent S W, Subak L L, Mendelson T). Rio de Janeiro: Editora Guanabara Koogan, 2001; 1 ed: 12-31.
23. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006; 6:9.
24. Richardson W S, Wilson M C, Guyatt G H, Cook D J, Nishikawa J. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. *JAMA* 1999; (281): 1214-1219.
25. Zamora J, Abraira V, Muriel A, Khan K S, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. In: *BMC Medical Research Methodology*, 2006; 31.
26. Smith-Bindman R, Kerlikowske K, Feldstein V A, Subak L, Scheidler J, Segal M, Brand R, Grady D. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. *JAMA* 1998; (280): 1510-1517.
27. Deeks J J, Altman D G, Bradburn M J. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: *Systematic Reviews in Health Care - Meta-analysis in context*. (Eds. Egger M, Smit D G, Altman D G). London: BMJ Publishing Group, 2001; 2nd: 285-312.
28. Tosteson A N, Begg C B. A general regression methodology for ROC curve estimation. *Med Decis Making* 1988; (8): 204-215.
29. Littenberg B, Moses L E. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993; (13): 313-321.
30. Sutton A J, Abrams K R, Jones D R, Sheldon T A, Song F. Meta-analysis of different types of data. In: *Methods for Meta-Analysis in Medical Research*. (Eds. Sutton A J, Abrams K R, Jones D R, Sheldon T A, Song F). Chichester: John Wiley & Sons, 2000; 1: 205-228.
31. Medeiros LR, Rosa DD, da Rosa MI, Bozzetti MC. Accuracy of CA 125 in the diagnosis of ovarian tumors: A quantitative systematic review. *Eur J Obstet Gynecol Reprod Biol*. 2008 Nov 6. [Epub ahead of print].
32. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000 Apr 19;283(15):2008-12.
33. *Cochrane Handbook for Systematic reviews of Interventions*. Disponível em: *Cochrane Handbook for Systematic Reviews of Interventions*. Acessado em 20/novembro de 2008.

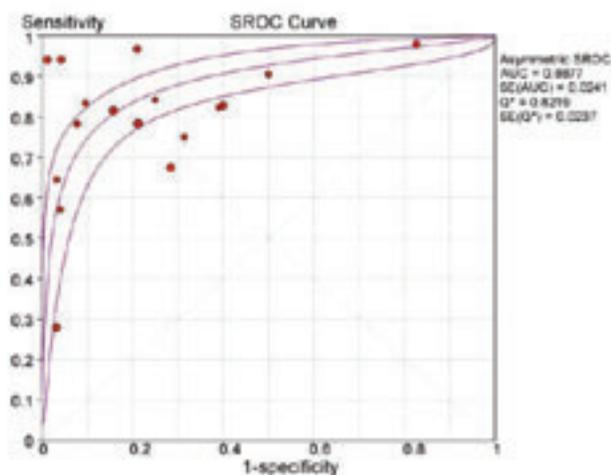
**Quadro 1:** Etapas da elaboração de uma revisão sistemática

|   |
|---|
| 1 - Formular a questão da revisão (assunto que gera incerteza na conduta)                                   |
| 2 - Definição dos critérios de inclusão e exclusão:   |
| ☒ Participantes   |
| ☒ intervenção e comparação  |
| ☒ desfechos   |
| ☒ delineamento do estudo e qualidade metodológica   |
| 3 - Estratégias de busca:   |
| ☒ The Cochrane Controlled Trials Register (CCTR)  |
| ☒ pesquisa em bancos de dados (Medline, Embase, Lilacs) informatizados e de estudos não registrados no CCTR |
| ☒ verificação da lista de referências dos artigos   |
| ☒ pesquisa manual de revistas na área do estudo   |
| ☒ contato pessoal com pesquisadores na área do estudo   |
| 4 - Seleção de estudos:   |
| ☒ critérios de elegibilidade conferidos por mais de um observador   |
| ☒ desenvolver estratégias para resolver as discordâncias  |
| ☒ informar os artigos excluídos e as razões   |
| 5 - Avaliação da qualidade metodológica dos estudos:  |
| ☒ avaliação deve ser realizada por mais de um pesquisador   |
| ☒ usar uma lista simples (checklist), não adotar escala de qualidade  |
| ☒ sempre avaliar a maneira de alocação e cegamento dos dois grupos  |
| ☒ considerar o cegamento dos observadores para os autores, instituição e revistas                           |
| 6 - Extração dos dados:   |
| ☒ desenhar um piloto para extração dos dados  |
| ☒ considerar extração dos dados por mais de um observador   |
| ☒ considerar cegamento dos observadores para os autores, instituição e revistas                             |
| 7 - Apresentação e análise dos resultados:  |
| ☒ deve-se tabular os resultados dos estudos de forma individual   |
| ☒ examinar os gráficos (forest plot)  |
| ☒ explorar as causas possíveis de heterogeneidade   |
| ☒ considerar na metanálise todos os estudos ou subgrupos de cada estudo                                     |
| ☒ realizar a análise de sensibilidade examinando o gráfico do funil (funnel plots)                          |
| ☒ disponibilizar uma lista com os estudos excluídos, caso o leitor tenha interesse                          |
| 8 - Interpretação dos resultados:   |
| ☒ considerar as limitações, incluindo as publicações, e relatar os vieses                                   |
| ☒ considerar a força das evidências   |
| ☒ considerar a aplicabilidade   |
| ☒ considerar o número necessário para tratar (benefício e malefício)  |
| ☒ considerar as implicações econômicas  |
| ☒ considerar as implicações para estudos futuros  |

Fonte: Egger e Smith, 2001

**Tabela 1:** Coeficiente de regressão pelo método de Moses e Littenberger.

| Parâmetro | Intercepto | Erro padrão | T     | P    |
|-----------|------------|-------------|-------|------|
| a         | 3,059      | 0,324       | 9,4   | 0,00 |
| $\beta$   | -0,203     | 0,165       | 1,233 | 0,23 |

**Figura 1:** Sumário da curva ROC calculado a partir do método de Moses e Littenberger, referente à regressão linear mostrada acima, com intervalo de confiança de 95%.**Endereço para correspondência**

Maria Inês da Rosa  
Rua Cruz e Souza, 510.  
Bairro Pio Correa  
Criciúma-SC- Brazil  
CEP 88811-550  
E-mail: mir@unesc.net