

MÉTODOS ESTATÍSTICOS EM METANÁLISE I: INTRODUÇÃO

Suleimy Cristina MAZIN¹
Edson Zangiacomi MARTINEZ²

- RESUMO: A metanálise é o método estatístico que integra os resultados de diferentes estudos em uma medida síntese. No presente artigo, descrevemos alguns aspectos da metanálise, como as medidas de efeito usuais, as medidas de heterogeneidade entre diferentes estudos, a apresentação dos resultados e os vieses que podem comprometer os resultados metanalíticos.
- PALAVRAS-CHAVE: Metanálise; ensaios clínicos; estimação de parâmetros; estatística médica.

1 Introdução

Em uma revisão sistemática da literatura (Mulrow, 1994; Atallah, 1996; Cook, 1997), os pesquisadores delimitam uma pergunta específica e fazem um exaustivo levantamento dos estudos que tiveram por objetivo responder a esta questão. É uma forma de sintetizar objetivamente as informações disponíveis sobre a questão de interesse, sendo que os estudos escolhidos para compor a revisão devem possuir boa qualidade metodológica, segundo critérios bem definidos (Silva Filho et al., 2005). As revisões sistemáticas são frequentes na área da saúde, onde os pesquisadores têm por interesse buscar bons níveis de evidência sobre os possíveis efeitos de um tratamento ou sobre os fatores de risco para uma determinada enfermidade. Por exemplo, Moreira et al. (2003) realizaram uma pesquisa bibliográfica utilizando bases eletrônicas de dados (ou seja, as bases *MedLine* e *Lilacs*) para identificar artigos relevantes, publicados entre 1990 e 2001, que avaliassem a associação entre o diabetes mellitus e a depressão.

¹Centro de Métodos Quantitativos – CEMEQ, Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo – USP, CEP: 14049-900, Ribeirão Preto, SP, Brasil. E-mail: sumazin@pop.com.br

²Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo – USP, CEP: 14049-900, Ribeirão Preto, SP, Brasil. E-mail: edson@fmrp.usp.br

Os dados de uma revisão sistemática podem ser analisados quantitativamente por uma metanálise, que é uma metodologia estatística que combina ou integra os resultados de diversos estudos independentes, voltados a uma única questão, sintetizando-os em uma só medida (Haines e Walters, 1995; Whitehead, 2002; Martinez, 2007). Suas bases surgiram no século XVII, na astronomia, onde se estabeleceu que a combinação dos dados de diferentes estudos poderia ser mais apropriada que a observação de alguns destes trabalhos (Egger e Smith, 1997). No início do século XX, o estatístico Karl Pearson foi provavelmente o primeiro pesquisador a usar técnicas formais para combinar dados de diferentes estudos médicos, quando examinou o efeito preventivo de inoculações contra febre entérica (Pearson, 1904). O termo metanálise apareceu pela primeira vez em 1976, em um artigo do psicólogo Gene Glass que o definiu como “*uma análise de uma grande coleção de resultados de análises provenientes de estudos individuais, tendo como propósito completar o que foi encontrado*”.

Em meados do século XX houve um grande crescimento da literatura científica, e com isso um aumento do uso da metanálise em várias áreas do conhecimento, como a pesquisa social (Glass et al., 1981), a educação (Kulik e Kulik, 1989), a medicina e a odontologia (Coutinho, 2005). Um grande salto para a utilização da metanálise nos estudos médicos resultou das palavras do médico e epidemiologista britânico Archie Cochrane, quando em 1979 criticou o fato de não haver resumos organizados e atualizados de todos os trabalhos relevantes em saúde, elogiando as revisões sistemáticas de estudos de intervenções durante a gravidez e o parto, publicadas na época, sugerindo que os métodos aplicados nestes estudos deveriam ser utilizados por pesquisadores de outras especialidades (Cochrane, 1989; Atallah, 1998). Para atender a este apelo e incentivar a execução de revisões sistemáticas na área da saúde, nasceu em 1992 o Centro Cochrane Britânico, uma organização que visa a colaboração entre instituições do Reino Unido e de outros países (ver www.cochrane.org). A Colaboração Cochrane no Reino Unido é uma organização sem fins lucrativos, com o objetivo de elaborar, manter e assegurar o acesso às revisões sistemáticas sobre efeitos de intervenções na área da saúde. Com a expansão da organização surgiu, em 1997, o Centro Cochrane do Brasil, seção brasileira da colaboração Cochrane (ver www.centrocochranedobrasil.org.br). Outro importante marco da metanálise foi o surgimento da chamada “*medicina baseada em evidências*”, onde as revisões sistemáticas com metanálise aparecem como o processo de investigação científica merecedor de maior nível de evidência (Medeiros e Stein, 2002).

O presente artigo tem como objetivo apresentar uma revisão sobre a metanálise, descrevendo as medidas de efeito mais usuais, as medidas de heterogeneidade entre diferentes estudos e sua importância, e um exemplo de aplicação em dados da área da saúde. Na Seção 2 apresentaremos as medidas de efeito em metanálise com seus respectivos intervalos de confiança; na Seção 3 discutiremos as medidas de heterogeneidade; na Seção 4 apresentaremos o método de DerSimonian e Laird, que tem como principal vantagem a incorporação da

medida de heterogeneidade em metanálise; na Seção 5, um exemplo de aplicação de metanálise; na Seção 6, a apresentação dos resultados da metanálise por *forest plots*; na Seção 7, uma discussão sobre os vieses que podem afetar os resultados das metanálises e, finalmente, na Seção 8, uma discussão encerra o artigo.

2 Medidas de efeito em metanálise

Nesta seção, apresentaremos algumas medidas de efeito utilizadas quando são comparadas em cada estudo respostas binárias entre dois grupos independentes de indivíduos. Para exemplificar estas medidas em um contexto de análise de dados da área médica, consideraremos resultados oriundos de ensaios clínicos aleatorizados. Um ensaio clínico é um experimento que envolve pacientes, desenhado para evidenciar a conduta mais adequada para o tratamento de indivíduos portadores de uma condição médica. Sua característica essencial é o uso de resultados baseados em uma amostra limitada de pacientes para fazer inferências sobre qual tratamento deve ser conduzido em uma população de indivíduos que necessitam ser tratados (Pocock, 1983). Quando é prevista a comparação da experiência dos indivíduos submetidos ao tratamento sob investigação com um grupo de pacientes com um tratamento padrão ou simulado, a um placebo, ou mesmo a nenhum tipo de tratamento, o ensaio clínico é chamado controlado. Um ensaio clínico aleatorizado é aquele que utiliza um processo aleatório para alocar os participantes nos diferentes grupos de procedimentos estudados. Vamos considerar que a resposta de cada indivíduo ao tratamento ao qual foi submetido é expressa por uma variável binária (são exemplos: óbito ou sobrevivência, redução completa ou parcial do tamanho de um tumor, alívio ou não da dor, cura ou não da doença).

Considere uma metanálise de J ensaios clínicos aleatorizados independentes, onde são comparados dois tratamentos, que denotaremos por A e B. Vamos considerar que no j -ésimo estudo ($j = 1, 2, \dots, J$), n_{1j} indivíduos foram submetidos ao tratamento A e n_{0j} indivíduos foram submetidos ao tratamento B. Podemos considerar, por exemplo, que o tratamento A é um tratamento sob investigação, e o tratamento B prevê a administração de um placebo (o que especifica um ensaio clínico controlado). A Tabela 1 apresenta a notação utilizada para os resultados do j -ésimo estudo, onde d_{1j} é o número de respostas ao tratamento A, $n_{1j} - d_{1j}$ é o número de indivíduos que não responderam ao tratamento A, d_{0j} é o número de respostas ao tratamento B e $n_{0j} - d_{0j}$ é o número de indivíduos que não responderam ao tratamento B. Considerando o j -ésimo estudo, a chance de resposta p_{1j} entre os indivíduos submetidos ao tratamento A é estimada por $\widehat{p}_{1j} = d_{1j}/n_{1j}$ e a chance de resposta p_{0j} entre os indivíduos submetidos ao tratamento B é estimada por $\widehat{p}_{0j} = d_{0j}/n_{0j}$. Assim $d_{1j} \sim \text{Binomial}(p_{1j}, n_{1j})$ e $d_{0j} \sim \text{Binomial}(p_{0j}, n_{0j})$.

Ao denotarmos genericamente uma medida de efeito para o j -ésimo estudo por θ_j ($j = 1, 2, \dots, J$), a metanálise objetiva estimar uma medida resumo θ_M , que sintetiza os resultados dos J estudos envolvidos na revisão sistemática. De

Tabela 1 - Notação para o j -ésimo estudo

	respostas	não respostas	total
Tratamento A	d_{1j}	$n_{1j} - d_{1j}$	n_{1j}
Tratamento B	d_{0j}	$n_{0j} - d_{0j}$	n_{0j}
Total	d_j	$n_j - d_j$	n_j

acordo com os objetivos da metanálise, a medida θ_j é expressa por um *odds ratio*, uma redução absoluta de risco, um risco relativo, ou outra quantidade de interesse. Sendo $\hat{\theta}_j$ uma estimativa amostral de θ_j , a medida resumo ou metanalítica θ_M é genericamente estimada pela média ponderada

$$\hat{\theta}_M = \frac{\sum_{j=1}^J w_j \hat{\theta}_j}{\sum_{j=1}^J w_j}, \quad (1)$$

onde w_j é o peso de cada estudo na metanálise. Este método para a estimação de uma medida metanalítica é chamado de método de Mantel-Haenszel (Mantel e Haenszel, 1959; Petitti, 1994, p.96), e considera que os estudos são, entre si, homogêneos (este pressuposto será discutido na Seção seguinte). Considerando que uma revisão sistemática envolve estudos de tamanhos amostrais diferentes, a “força” da evidência (a favor ou não do tratamento) de cada estudo não é igual. Desta forma, convencionou-se em uma metanálise que a contribuição de cada um dos estudos à determinação de uma medida resumo deve ser relacionada ao tamanho do estudo e ao número de eventos observados. Assim, w_j é geralmente definido pelo inverso da variância estimada para a medida do efeito associada ao j -ésimo estudo.

2.1 Odds ratio

Um *odds* é definido pela divisão entre a probabilidade de um evento ocorrer e a probabilidade deste mesmo evento não ocorrer. Considerando o tratamento A, a chance de uma resposta é estimada pela razão d_{1j}/n_{1j} e o seu complemento é estimado por $(n_{1j} - d_{1j})/n_{1j}$. Portanto, o *odds* para uma resposta considerando o tratamento A é estimado por $d_{1j}/(n_{1j} - d_{1j})$. Analogamente, o *odds* para uma resposta considerando o tratamento B é estimado por $d_{0j}/(n_{0j} - d_{0j})$. O *odds ratio*, ou razão de *odds*, expressa quantas vezes o *odds* da resposta no tratamento A equivale ao *odds* da resposta no tratamento B. Assim o *odds ratio* (*OR*) associado ao j -ésimo estudo é estimado por

$$\widehat{OR}_j = \frac{d_{1j} (n_{0j} - d_{0j})}{d_{0j} (n_{1j} - d_{1j})}, \quad (2)$$

para $j = 1, 2, \dots, J$ (Bland e Altman, 2000). Um *odds ratio* metanalítico, que resume as informações dos vários estudos que compõem a metanálise, é uma média dos *odds*

ratios de cada estudo, ponderada pelo peso w_j , o inverso da variância estimada para o j -ésimo estudo. De acordo com a expressão (1), o *odds ratio* metanalítico (ORM) é estimado por

$$\widehat{ORM} = \frac{\sum_{j=1}^J w_j \widehat{OR}_j}{\sum_{j=1}^J w_j}.$$

Uma aproximação para a variância do *odds ratio* considera $\widehat{var}(\widehat{OR}_j) = n_j [d_{0j} (n_{1j} - d_{1j})]^{-1}$ e $w_j^{-1} = \widehat{var}(\widehat{OR}_j)$, onde $n_j = n_{0j} + n_{1j}$. Assim, o *odds ratio* metanalítico de Mantel-Haenszel (ORM_{MH}) é estimado por

$$\widehat{ORM}_{MH} = \frac{\sum_{j=1}^J \frac{d_{1j} (n_{0j} - d_{0j})}{n_j}}{\sum_{j=1}^J \frac{d_{0j} (n_{1j} - d_{1j})}{n_j}}.$$

Outra maneira de estimar o *odds ratio* metanalítico é obtida utilizando a estimativa da variância de cada estudo pelo método de Woolf (1955), onde, do método delta,

$$w_j^{-1} = \widehat{var}(\ln \widehat{OR}_j) = \frac{1}{d_{1j}} + \frac{1}{n_{1j} - d_{1j}} + \frac{1}{d_{0j}} + \frac{1}{n_{0j} - d_{0j}}.$$

Um método para estimar a variância do *odds ratio* metanalítico de Mantel-Haenszel foi introduzido por Robins et al. (1986), considerando

$$\widehat{var}(\widehat{ORM}_{MH}) = \frac{\sum_{j=1}^J F_j}{2 \sum_{j=1}^J R_j^2} + \frac{\sum_{j=1}^J G_j}{2 \sum_{j=1}^J R_j \sum_{j=1}^J S_j} + \frac{\sum_{j=1}^J H_j}{2 \sum_{j=1}^J S_j^2},$$

em que

$$\begin{aligned} F_j &= d_{1j} (n_{0j} - d_{0j}) \frac{n_{1j}}{n_j^2}, \\ G_j &= \frac{d_{1j} (n_{0j} - d_{0j}) (d_{0j} + n_{1j} - d_{1j}) + d_{0j} (n_{1j} - d_{1j}) (d_{1j} + n_{0j} - d_{0j})}{n_j^2}, \\ H_j &= \frac{d_{0j} (n_{1j} - d_{1j}) (d_{0j} + n_{1j} - d_{1j})}{n_j^2}, \\ R_j &= \frac{d_{1j} (n_{0j} - d_{0j})}{n_j}, \text{ e} \\ S_j &= \frac{d_{0j} (n_{1j} - d_{1j})}{n_j}. \end{aligned}$$

Assim, um intervalo assintótico de confiança 95% para o ORM_{MH} é dado por

$$\exp \left[\ln \widehat{ORM}_{MH} \pm 1,96 \sqrt{\widehat{var}(\widehat{ORM}_{MH})} \right].$$

2.2 Odds ratio de Peto

Ao observarmos a expressão (2), notamos que OR_j não pode ser estimado quando d_{0j} é igual a zero ou d_{1j} é igual a n_{1j} , considerando algum $j = 1, 2, \dots, J$. Esta é uma situação comum em metanálises, em que frequentemente estão presentes estudos em que o tamanho da amostra não é grande o suficiente para a ocorrência de ao menos uma resposta (ou não resposta). Uma alternativa é o *odds ratio de Peto* (ORP_j) (Yusuf et al., 1985) que possui a mesma interpretação que o *odds ratio* convencional. O *odds ratio de Peto* (ORP_j) associado ao j -ésimo estudo é estimado por

$$\widehat{ORP}_j = \exp\left(\frac{O_j - E_j}{V_j}\right),$$

onde

$$\begin{aligned} O_j &= d_{1j}, \\ E_j &= \frac{n_{1j}d_j}{n_j}, \text{ e} \\ V_j &= \frac{n_{0j}n_{1j}d_j(n_{0j} + n_{1j} - d_{1j} - d_{0j})}{n_j^2(n_j - 1)}, \end{aligned}$$

$d_j = d_{1j} + d_{0j}$, sendo que V_j fornece uma estimativa para a variância da diferença entre O_j e E_j , e também uma quantificação para o peso do j -ésimo estudo na obtenção de uma medida metanalítica de *odds ratio*. Notar que O_j é o número de eventos observados no grupo de indivíduos submetidos ao tratamento, e E_j é o número esperado de eventos neste mesmo grupo, sob a hipótese de que as proporções de respostas entre os dois tratamentos são homogêneas. Um intervalo assintótico de confiança 95% para ORP_j é dado por

$$\exp\left(\frac{O_j - E_j \pm 1,96\sqrt{V_j}}{V_j}\right).$$

O *odds ratio* metanalítico de Peto ($ORMP$) é estimado por

$$\widehat{ORMP} = \exp\left[\frac{\sum_{j=1}^J (O_j - E_j)}{\sum_{j=1}^J V_j}\right],$$

e um intervalo assintótico de confiança 95% para $ORMP$ é obtido por

$$\exp\left[\frac{\sum_{j=1}^J (O_j - E_j) \pm 1,96\sqrt{\sum_{j=1}^J V_j}}{\sum_{j=1}^J V_j}\right].$$

2.3 Redução absoluta de risco

A redução absoluta de risco (RAR) associada ao j -ésimo estudo é estimada pela diferença entre a chance de resposta dos indivíduos submetidos ao tratamento A em relação ao tratamento B, ou seja,

$$\widehat{RAR}_j = \widehat{p}_{0j} - \widehat{p}_{1j} = \frac{d_{0j}}{n_{0j}} - \frac{d_{1j}}{n_{1j}},$$

para $j = 1, 2, \dots, J$. Petitti (1994) mostra que a variância estimada para \widehat{RAR}_j é aproximada por

$$var(\widehat{RAR}_j) = \frac{(d_{1j} + d_{0j})(n_{1j} - d_{1j} + n_{0j} - d_{0j})}{n_{1j}n_{0j}(n_{1j} + n_{0j})}.$$

Uma estimativa para a redução absoluta de risco metanalítica segundo o método de Mantel-Haenszel é dada por Robins et. al (1986), que considera

$$\widehat{RARM}_{MH} = \frac{\sum_{j=1}^J \left(\frac{d_{0j}n_{1j}}{n_j} - \frac{d_{1j}n_{0j}}{n_j} \right)}{\sum_{j=1}^J \frac{n_{0j}n_{1j}}{n_j}},$$

onde $n_j = n_{0j} + n_{1j}$. Sendo sua variância estimada através do método delta por

$$\widehat{var}(\widehat{RARM}_{MH}) = \frac{\sum_{j=1}^J \frac{d_{1j}(n_{1j} - d_{1j})n_{0j}^3 + d_{0j}(n_{0j} - d_{0j})n_{1j}^3}{n_{1j}n_{0j}n_j^2}}{\left(\sum_{j=1}^J \frac{n_{1j}n_{0j}}{n_j} \right)^2},$$

consequentemente, um intervalo assintótico de confiança 95% para a \widehat{RARM}_{MH} é dado por

$$\widehat{RARM}_{MH} \pm 1,96 \sqrt{\widehat{var}(\widehat{RARM}_{MH})}.$$

2.4 Risco relativo

O risco relativo (RR) expressa a razão entre a chance de resposta dos indivíduos submetidos ao tratamento A em relação ao tratamento B, ou seja, para o j -ésimo estudo

$$\widehat{RR}_j = \frac{d_{1j}/n_{1j}}{d_{0j}/n_{0j}} = \frac{d_{1j}n_{0j}}{d_{0j}n_{1j}}.$$

Um risco relativo próximo a 1 evidencia que a chance de resposta dos indivíduos submetidos ao tratamento A é igual a chance de resposta ao tratamento B. Quanto mais distante o risco relativo do valor 1, mais evidente é a associação entre o

tratamento e o evento de interesse. A relação entre o risco relativo e o *odds ratio* é explicada pela expressão

$$\widehat{RR}_j = \frac{\widehat{OR}_j}{1 - \frac{d_{0j}}{n_{0j}} + \widehat{OR}_j \frac{d_{0j}}{n_{0j}}}. \quad (3)$$

Quando a proporção d_{0j}/n_{0j} é próxima de zero, o denominador da expressão (3) aproxima-se de 1, tornando \widehat{RR}_j próximo de \widehat{OR}_j . Por este motivo, muitos pesquisadores utilizam o *odds ratio* como uma aproximação do risco relativo em estudos onde a ocorrência do evento de interesse é relativamente baixa. A principal vantagem do uso do *odds ratio* neste caso está na facilidade em estimá-lo através de modelos de regressão logística, que permitem ajustamentos por variáveis de controle. Entretanto, é sempre importante observar que o *odds ratio* pode trazer medidas que superestimam o risco relativo quando o evento de interesse é relativamente frequente. Os estudos de Simon (2001), Holcomb et al. (2001) e Schechtman (2002) comparam o uso do *odds ratio* e do risco relativo, discutindo as vantagens e as desvantagens de cada medida.

O risco relativo metanalítico (*RRM*) estimado pelo método de Mantel-Haenszel (Rothman e Greenland, 1998) considera $w_j = n_j^{-1}d_{0j}n_{1j}$. Portanto, uma estimativa de RRM_{MH} é dada por

$$\widehat{RRM}_{MH} = \frac{\sum_{j=1}^J \frac{d_{1j}n_{0j}}{n_j}}{\sum_{j=1}^J \frac{d_{0j}n_{1j}}{n_j}}, \quad (4)$$

em que $n_j = n_{0j} + n_{1j}$.

Um método para estimar a variância do logaritmo do risco relativo metanalítico de Mantel-Haenszel foi introduzido por Robins et al. (1986), considerando

$$\widehat{var} \left(\ln \widehat{RRM}_{MH} \right) = \frac{\sum_{j=1}^J \frac{n_{0j}n_{1j} (d_{1j} + d_{0j}) - d_{1j}d_{0j}n_j}{n_j^2}}{\left(\sum_{j=1}^J \frac{d_{1j}n_{0j}}{n_j} \right) \left(\sum_{j=1}^J \frac{d_{0j}n_{1j}}{n_j} \right)}$$

Assim, um intervalo assintótico de confiança 95% para o RRM_{MH} é dado por

$$\exp \left[\ln \widehat{RRM}_{MH} \pm 1,96 \sqrt{\widehat{var} \left(\ln \widehat{RRM}_{MH} \right)} \right]. \quad (5)$$

Outra maneira usual de obter intervalos de confiança assintóticos para o RR é através do uso da variância do logaritmo de \widehat{RR}_j , estimada pelo método delta por

$$w_j^{-1} = \widehat{var}(\ln \widehat{RR}_j) = \frac{1}{d_{1j}} - \frac{1}{n_{1j}} + \frac{1}{d_{0j}} - \frac{1}{n_{0j}}.$$

3 Medidas de heterogeneidade

As medidas metanalíticas apresentadas na seção anterior pressupõem que a variabilidade observada entre os resultados dos estudos que compõem a metanálise é devida apenas a flutuações amostrais aleatórias. Nesta situação, dizemos que os estudos são homogêneos (Engels et al., 2000; Xu et al., 2008). Quando a variabilidade entre os estudos tem origem não aleatória, dizemos que os estudos são heterogêneos. Song et al. (2001) citam que as possíveis causas da heterogeneidade podem ter origem nas diferenças entre protocolos empregados, definição da doença, planejamento do estudo, variações entre tratamentos ou variabilidades demográficas nas populações sob estudo.

A avaliação da heterogeneidade é de grande importância na metanálise, pois a ausência ou a presença da heterogeneidade indica o tratamento estatístico mais adequado a ser utilizado (Huedo-Medina et al., 2006; Song et al., 2001; Higgins et al., 2003; Sidik e Jonkman, 2007). Nos estudos em saúde, a heterogeneidade é geralmente avaliada pelo teste Q proposto por Cochran (1954) ou quantificada pela estatística I^2 definida por Higgins e Thompson (2002).

3.1 O teste Q de Cochran

Considere uma metanálise composta por J estudos, onde $\hat{\theta}_j$ é a estimativa da medida de efeito para o j -ésimo estudo. A estatística Q proposta por Cochran (1954) é dada por

$$Q = \sum_{j=1}^J w_j (\hat{\theta}_j - \hat{\theta}_M)^2, \quad (6)$$

onde $\hat{\theta}_M$ é a estimativa para a medida metanalítica θ_M e o peso w_j é dado pelo inverso da variância da medida para o j -ésimo estudo. Notar que, em metanálises onde a medida de efeito é o *odds ratio*, é comum estabelecer que $\hat{\theta}_j = \ln \widehat{OR}_j$ e $\hat{\theta}_M = \ln \widehat{OR}_M$. Sob a hipótese de homogeneidade entre os estudos, Q tem, assintoticamente, distribuição qui-quadrado com $J - 1$ graus de liberdade.

Quando a medida de efeito escolhida para a metanálise é o *odds ratio* de Peto, a estatística Q é dada por

$$Q = \sum_{j=1}^J \frac{(O_j - E_j)^2}{V_j} - \frac{\left[\sum_{j=1}^J (O_j - E_j) \right]^2}{\sum_{j=1}^J V_j}.$$

Nos casos em que o método de Mantel-Haenszel é utilizado, $\hat{\theta}_M$ é substituído por $\hat{\theta}_{MH}$ (a medida metanalítica estimada pelo método Mantel-Haenszel) em (6), onde w_j é o inverso da variância estimada pelo método Mantel-Haenszel.

Uma deficiência deste teste é que ele possui um baixo poder quando o número J de estudos que compõem a metanálise é pequeno. Por outro lado, quando o

número de estudos é muito grande ele pode detectar uma falsa heterogeneidade (Higgins e Thompson, 2002).

3.2 Estatísticas H^2 e I^2

Seja θ_j a medida de efeito para o j -ésimo estudo, $j = 1, \dots, J$, e seja $\hat{\theta}_j$ uma estimativa de θ_j . As estatísticas H^2 e I^2 são derivadas assumindo $\tilde{\sigma}^2 = \widehat{var}(\hat{\theta}_j) = w_j^{-1}$, ou seja, as variâncias das medidas de efeito de cada estudo são conhecidas (dado que w_j são conhecidos) e iguais. Sendo θ_M a medida metanalítica, vamos considerar τ^2 a variabilidade da medida de efeito θ_j entre os J estudos, sem que uma distribuição específica seja assumida. Seja

$$\hat{\tau}^2 = \tilde{\sigma}^2 \left(\frac{Q}{J-1} - 1 \right) \quad (7)$$

um estimador de momentos de τ^2 considerando w_j conhecidos (ver demonstração no apêndice), $j = 1, \dots, J$. A partir da razão $\rho = \hat{\tau}^2 / \tilde{\sigma}^2$ são geradas as estatísticas H^2 e I^2 .

Considerando

$$\rho + 1 = \frac{\hat{\tau}^2 + \tilde{\sigma}^2}{\tilde{\sigma}^2} = \frac{\tilde{\sigma}^2 \left(\frac{Q}{J-1} - 1 \right) + \tilde{\sigma}^2}{\tilde{\sigma}^2} = \frac{Q}{J-1} - 1 + 1 = \frac{Q}{J-1},$$

temos

$$H^2 = \frac{Q}{J-1}.$$

Lembrar que, quando os estudos são homogêneos, Q segue assintoticamente uma distribuição qui-quadrado com $J - 1$ graus de liberdade. Assim, quando H é próximo a 1, temos evidência de homogeneidade entre os estudos (Higgins e Thompson, 2002).

Considerando a relação

$$\frac{\rho}{1 + \rho} = \frac{\hat{\tau}^2}{\hat{\tau}^2 + \tilde{\sigma}^2} = \frac{H^2 - 1}{H^2} = \frac{\frac{Q}{J-1} - 1}{\frac{Q}{J-1}} = \frac{Q - (J - 1)}{Q},$$

Higgins e Thompson (2002) propuseram a estatística I^2 , dada por

$$I^2 = \frac{H^2 - 1}{H^2} 100\% = \frac{Q - (J - 1)}{Q} 100\%. \quad (8)$$

A estatística I^2 é a mais utilizada entre as duas estatísticas e descreve a proporção da variação total que é devida à heterogeneidade. Notar que este índice

não ultrapassa 100%, e quando negativo, é igualado a 0. Higgins e Thompson (2002) propuseram um intervalo de confiança 95% assintótico para H , dado por

$$\exp(\ln H \pm 1,96 \hat{e}p[\ln(H)])$$

onde $\hat{e}p[\ln(H)]$ é a estimativa do erro padrão de $\ln(H)$ estimado por

$$\hat{e}p[\ln(H)] = \frac{1}{2} \frac{\ln(Q) - \ln(J-1)}{\sqrt{2Q - (2J-3)}} \quad \text{se } Q > J$$

$$\hat{e}p[\ln(H)] = \sqrt{\frac{1}{2(J-2)} \left(1 - \frac{1}{3(J-2)^2}\right)} \quad \text{se } Q \leq J.$$

Assim, para encontrar o intervalo de confiança para I^2 basta substituir os limites de confiança em (8).

De modo subjetivo, Higgins et al. (2003) sugerem uma escala em que um valor de I^2 próximo a 0% indica não heterogeneidade entre os estudos, um valor de I^2 próximo 25% indica baixa heterogeneidade, um valor de I^2 próximo 50% indica heterogeneidade moderada, um valor de I^2 próximo 75% indica uma alta heterogeneidade entre os estudos.

Huedo-Medina et al. (2006) descrevem que a estatística I^2 tem vantagens importantes em relação ao teste Q de Cochran, como a sua facilidade de interpretação e a sua não dependência do número de graus de liberdade. Outra vantagem está na possibilidade de descrever a magnitude da heterogeneidade na metanálise, enquanto o teste Q de Cochran trata apenas a “significância estatística” da hipótese de homogeneidade.

4 Método de DerSimonian e Laird

A vantagem do método proposto por DerSimonian e Laird (1986) está na possibilidade de estimarmos medidas metanalíticas sem a necessidade de pressupormos que os estudos que compõem a metanálise são homogêneos. Considerando $\hat{\theta}_j$ a medida de efeito para o j -ésimo estudo, a medida metanalítica θ_{MDL} de DerSimonian e Laird é estimada por

$$\hat{\theta}_{MDL} = \frac{\sum_{j=1}^J w_j^* \hat{\theta}_j}{\sum_{j=1}^J w_j^*}. \quad (9)$$

O peso w_j^* para o j -ésimo estudo é dado por

$$w_j^* = \frac{1}{\hat{\tau}^2 + w_j^{-1}},$$

onde

$$\hat{\tau}^2 = \frac{[Q - (J - 1)] \sum_{j=1}^J w_j}{\left(\sum_{j=1}^J w_j\right)^2 - \sum_{j=1}^J w_j^2}, \quad (10)$$

sendo Q dado em (6) (ver demonstração no apêndice). E ainda, w_j é o inverso da variância estimada para $\hat{\theta}_j$, $\hat{\theta}_M$ é a medida metanalítica estimada e J é o número de estudos contidos na metanálise.

Vamos considerar que $\hat{\tau}^2$ é uma estimativa amostral de τ^2 , uma medida da magnitude em que os efeitos de tratamento variam entre os estudos. Notar que todas as medidas metanalíticas mostradas na Seção 2 consideram $\tau^2 = 0$, ou seja, assumem que os J estudos envolvidos na metanálise são homogêneos. Se a expressão (10) trouxer um valor para $\hat{\tau}^2$ menor que zero, consideramos $\hat{\tau}^2 = 0$ e, conseqüentemente, $w_j^* = w_j$. Considerando que, assintoticamente, $\hat{\tau}^2 \sim \chi_{(J-1)}^2$, podemos usar esta estatística para testar a hipótese $H_0 : \tau^2 = 0$.

5 Exemplo: eficácia terapêutica do Interferon Alfa

Silva Filho et al. (2005) realizaram uma revisão sistemática com metanálise de 38 estudos clínicos aleatorizados e controlados (ver Tabela 2), sobre a eficácia terapêutica do Interferon Alfa no tratamento da hepatite crônica pelo vírus B. Em cada estudo os indivíduos foram alocados em um grupo que recebeu o tratamento (intervenção com o Interferon Alfa) ou em um grupo que não recebeu a intervenção (grupo controle). O evento de interesse neste estudo foi o clareamento do HBV-DNA (DNA do vírus da hepatite B). Notar que d_1 e d_0 representam as frequências de não clareamento do HBV-DNA nos grupos tratados e controles, respectivamente.

O risco relativo de cada estudo encontra-se na Figura 1. A estimativa do risco relativo metanalítico de Mantel-Haenszel (4) é 0,71 e seu respectivo intervalo de confiança 95% baseado em (5) é (0,67; 0,75). Como \widehat{RRM}_{MH} é menor que 1, observa-se uma evidência de que a ocorrência do evento é maior no grupo tratamento do que no grupo controle. A estatística Q é dada por 100,66, onde consideramos $\hat{\theta}_j = \ln \widehat{RR}_j$ na expressão (6). Como $Q > \chi_{37}^2$, ao nível de significância de 5%, rejeita-se a hipótese de homogeneidade entre os estudos na metanálise. O índice H^2 é dado por $H^2 = 2,72$. Como $H^2 > 1$ temos indícios de heterogeneidade entre os estudos. Com base em (8) temos $I^2 = 63\%$, indicando moderada heterogeneidade entre os estudos (ver Higgins et al., 2003).

Como $Q > J$, o erro padrão estimado para o logaritmo de H é

$$\widehat{ep}[\ln(H)] = \frac{1}{2} \frac{\ln(Q) - \ln(J - 1)}{\sqrt{2Q - (2J - 3)}} = \frac{1}{2} \frac{\ln(100,66) - \ln(38 - 1)}{\sqrt{2(100,66) - [2(38) - 3]}} = 0,04$$

e o intervalo de confiança 95% para o logaritmo de H é

$$\exp(\ln H \pm 1,96 \widehat{ep}[\ln(H)]) = \exp\left[\ln \sqrt{2,03} \pm 1,96(0,04)\right] \text{ ou } (1,51; 1,79).$$

Tabela 2 - Dados sobre a eficácia terapêutica do Interferon Alfa no tratamento da hepatite crônica pelo vírus B (Filho et al., 2005)

Estudo	Tratamento		Placebo		Estudo	Tratamento		Placebo	
	d_1	n_1	d_0	n_0		d_1	n_1	d_0	n_0
Alexander 1987	17	23	23	23	Mazzella 1988	7	33	16	33
Barbera 1986	1	9	7	9	Müller 1990	19	28	24	27
Barbera 1994	26	40	31	37	Pastore 1988	6	14	13	14
Brook 1989a	12	45	15	15	Pastore 1992	9	10	8	8
Brook 1989b	25	37	29	30	Perrilo 1990	60	62	40	43
Brunetto 1989	7	12	12	12	Porres 1988	11	17	5	6
Carreño 1987	8	10	10	10	Realdi 1990	26	39	35	40
Carreño 1991	12	16	7	8	Ruiz Moreno 1990	8	12	9	12
Carreño 1999	20	27	27	30	Ruiz Moreno 1991	12	24	10	12
Di Bisceglie 1993	14	25	21	22	Rumi 1993	19	31	21	25
Dusheiko 1986	10	14	5	5	Saracco 1989	7	33	16	31
Fattovich 1989	11	17	13	16	Sarin 1996	10	20	20	21
Fattovich 1992	11	30	24	30	Sokal 1998	57	70	66	74
Hadziyannis 1990	15	25	25	25	Tcherveniakova 1999	18	31	10	12
Hoofnagle 1988	21	31	13	14	Tomas 1994	94	179	45	54
Janssen 1999	44	61	50	57	Utili 1991	7	10	9	10
Lampertico 1997	15	21	21	21	Waked 1990	3	16	10	15
Lok 1986	16	34	13	18	Willians 1990	17	23	6	7
Lok 1988	16	34	13	18	Wong 1995	19	25	22	25

Substituindo esses limites em (8), temos o intervalo de confiança 95% para I^2 :

$$(56, 29\% ; 69, 09\%) .$$

Dado que o teste Q de Cochran e as estatísticas I^2 e H^2 sugerem a presença de heterogeneidade entre os estudos, consideraremos ainda a medida metanalítica de DerSimonian e Laird. Das expressões (10) e (9), temos

$$\hat{\tau}^2 = 0,04 \quad \text{e} \quad \hat{\theta}_{MDL} = 0,80.$$

Observamos que, neste exemplo, $\hat{\theta}_{MDL}$ é próximo de \widehat{RRM}_{MH} , mesmo na presença de alguma heterogeneidade entre os estudos que compõem a metanálise.

6 Representação gráfica: o *forest plot*

Os resultados de uma metanálise são geralmente apresentados em um gráfico chamado *forest plot*, que mostra informações individuais dos estudos incluídos na metanálise, informações sobre a variabilidade de cada estudo e a medida de efeito metanalítica (Lewis e Clarke, 2001; Moja et al., 2007). Sua origem se dá nos últimos anos da década de 1970, quando Freiman et al. (1978) exibiu resultados de

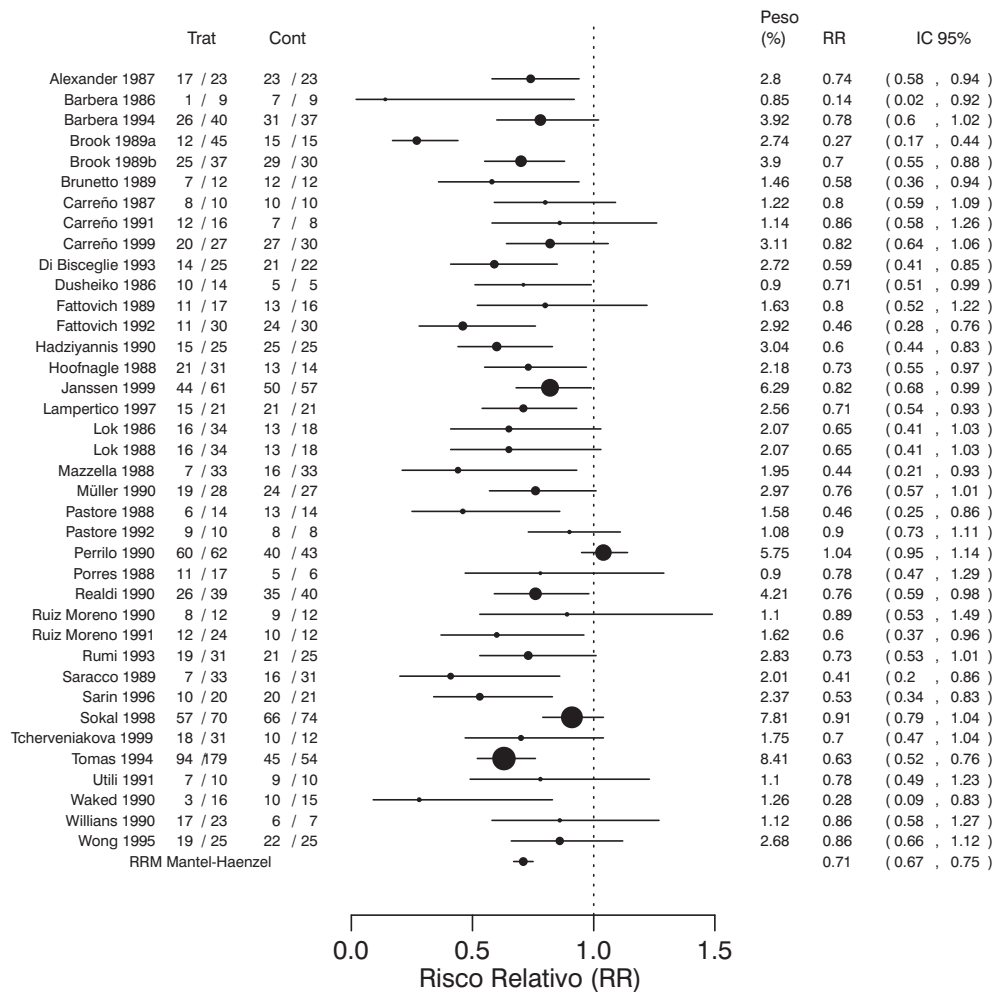


Figura 1 - Forest Plot dos dados sobre a eficácia terapêutica do Interferon Alfa no tratamento da hepatite crônica pelo vírus B (Filho et al., 2005).

diversos estudos com linhas horizontais mostrando o intervalo de confiança de cada estudo com a respectiva estimativa pontual. Curiosamente, este estudo não era uma metanálise. Lewis e Ellis(1982) propuseram um gráfico similar para metanálise e acrescentaram a estimativa da medida de efeito metanáltica.

O *forest plot* para os dados do exemplo anterior é mostrado na Figura 1. Na primeira coluna, à esquerda, são listados os estudos, rotulados pelo nome do primeiro autor e o ano da publicação. A ordem dos estudos é arbitrária, sendo que os estudos podem ser listados, por exemplo, segundo uma ordem alfabética ou de acordo com o

ano da publicação. As estimativas pontuais da medida de efeito de cada estudo são representadas por símbolos cujo tamanho é proporcional ao respectivo peso. Assim, o estudo de Tomas 1994 é aquele com maior contribuição à medida metanalítica, observando que ele possui o maior tamanho amostral dentre todos os estudos. À direita da figura, aparecem os pesos dos estudos, as medidas de efeito e os seus respectivos intervalos de confiança. Geralmente, o peso de cada estudo é expresso em porcentagens, representando a contribuição de cada um dos estudos à medida metanalítica. A medida metanalítica de Mantel-Haenzel é representada na parte inferior da figura. Os intervalos de confiança para cada estudo são representados por linhas horizontais, enquanto uma linha vertical separa as regiões onde o risco relativo é menor ou maior que 1. Observa-se que, com a exceção de um único estudo, todas as medidas de risco relativo encontram-se à esquerda desta linha vertical, evidenciando que o tratamento geralmente traz uma redução do risco do evento.

7 Vieses em metanálises

As metanálises são sujeitas a vários tipos de vieses. Dentre os mais importantes, encontram-se aqueles relacionados à seleção dos estudos na revisão sistemática, quando um ou mais estudos que poderiam trazer contribuições substanciais à medida metanalítica não são incluídos na análise (ver Egger e Smith, 1998). Possíveis causas desta não inclusão são:

- (a) Tais estudos foram publicados em revistas que não estão indexadas nas bases de dados escolhidas pelos pesquisadores;
- (b) Tais estudos foram publicados em línguas que os pesquisadores não têm domínio (por exemplo, artigos publicados em importantes revistas médicas da China no idioma deste país serão obviamente incompreensíveis aos pesquisadores brasileiros que desconhecem o mandarim);
- (c) Tais estudos foram publicados em revistas que não estão disponíveis *on-line* ou nas bibliotecas que os pesquisadores têm acesso.

Outro tipo de viés bastante discutido na literatura é o chamado viés de publicação (*publication bias*), conseqüente da prática de muitos periódicos médicos aceitarem mais facilmente para publicação os estudos de intervenção que mostraram efeitos benéficos (Dickersin et al., 1987; Easterbrook et al., 1991; Thornton e Lee, 2000). Isto torna indisponíveis aos pesquisadores os resultados dos estudos que não são publicados, e os resultados das metanálises tenderão a trazer medidas superestimadas do efeito do tratamento em estudo. O gráfico do funil (*funnel plot*) é uma ferramenta simples e bastante utilizada para a detecção de viés de publicação (Terrin et al. 2005). Trata-se de um gráfico da dispersão que traz no eixo das abscissas as medidas de efeito estimadas para cada um dos estudos que compõem a metanálise, e no eixo das ordenadas o respectivo peso de cada estudo

(ou o tamanho amostral). Quando o gráfico exibe uma disposição assimétrica de pontos, há evidências de que o viés de publicação está presente na metanálise (Egger et al., 1997). Na ausência do viés de verificação, a disposição dos pontos no gráfico assemelha-se ao formato de uma pirâmide ou funil invertido. Testes de hipóteses para a detecção de evidências de viés de publicação foram propostos por Dear e Begg (1992) e Begg e Mazumdar (1994).

8 Discussão

Perante uma questão clínica relevante sobre intervenções terapêuticas onde não existe um consenso dos estudos publicados, ou quando não há uma comprovação adequada da eficácia de um procedimento, a metanálise pode trazer importantes informações para as decisões sobre condutas, desde que alguns critérios sejam atendidos (Martinez, 2007). Estes critérios incluem uma questão clínica adequadamente formulada, uma busca completa e abrangente dos estudos em fontes confiáveis, a garantia da qualidade metodológica dos estudos e a utilização de ferramentas estatísticas adequadas. O presente artigo trouxe uma introdução aos aspectos estatísticos da metanálise e pretende ser útil a pesquisadores, com formação estatística ou não, que estão iniciando sua carreira na pesquisa clínica ou que têm algum interesse nestes métodos.

Outro aspecto importante da metanálise é o uso de modelos de regressão, que podem incorporar a heterogeneidade entre os estudos através de efeitos aleatórios. Estes modelos podem ainda introduzir covariáveis que podem estar associadas às medidas de efeito, como o ano da publicação do estudo, o tipo da intervenção, dosagens e a severidade da doença. Modelos de regressão em metanálises, em um contexto clássico e Bayesiano, serão discutidos em um segundo artigo.

Agradecimentos

A pesquisa recebeu apoio da Fundação de Amparo ao Ensino, Pesquisa e Assistência (FAEPA) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HC-FMRP), Universidade de São Paulo – USP.

MAZIN, S. C.; MARTINEZ, E. Z. Statistical methods in meta-analysis I: An introduction. *Rev. Bras. Biom.*, São Paulo, v.27, n.2, p.139-160, 2009.

- **ABSTRACT:** *Meta-analysis is a statistical method for combining the results from different studies in a synthesis measure. In the present article, we describe some aspects of the meta-analysis as the measures of effect, the measures of heterogeneity among the studies, the presentation of results and the bias that can occur when conducting a meta-analysis.*
- **KEYWORDS:** *Meta-analysis; clinical trials; parameter estimation; statistics in medicine.*

Referências

- ATALLAH, A. N. Systematic reviews of medical literature and meta-analyses. *Sao Paulo Med. J.*, São Paulo, v.114, n.5, p.1285-1287, 1996.
- ATALLAH, A. N. Revisão sistemática da literatura e metanálise. In: ATALLAH, A. N.; CASTRO, A. A. (Ed.). *Medicina baseada em evidências: fundamentos da pesquisa clínica*. São Paulo: Lemos-Editorial; 1998. p.42-48.
- BEGG, C. B.; MAZUMDAR, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*, Washington, v.50, n.4, p.1088-1101, 1994.
- BIGGERSTAFF, B. J.; TWEEDIE, R. L. Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Stat. Med.*, New York, v.16, n.7, p.753-768, 1997.
- BLAND, J. M.; ALTMAN, D. G. Statistics notes. The odds ratio. *Br. Med. J.*, Baarn, v.320, n.7247, p.1468; 2000.
- COCHRAN, W. G. The combination of estimates from different experiments. *Biometrics*, Washington, v.10, p.101-129, 1954.
- COCHRANE A. L. Foreword. In: CHALMERS, I.; ENKIN, M.; KEIRSE, M. J. N. C. (Ed.). *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press, 1989.
- COOK, D. J.; MULROW, C. D.; HAYNES, R. B. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann. Intern. Med.*, Philadelphia, v.126, n.5, p.376-380, 1997.
- COUTINHO, E. Revisão sistemática e meta-análise em odontologia. In: LUIZ, R. R.; COSTA, A. J. L.; NADANOVSKY, P. (Ed.). *Epidemiologia e bioestatística na pesquisa odontológica*. São Paulo: Atheneu, 2005. p.397-403.
- DEAR, K. B. G.; BEGG, C. B. An approach for assessing publication bias prior to performing a meta-analysis. *Stat. Sci.*, Hayward, v.7, n.2, p.237-245, 1992.
- DERSIMONIAN, R.; LAIRD, N. Meta-analysis in clinical trials. *Control. Clin. Trials*, New York, v.7, n.3, p.177-188, 1986.
- DICKERSIN, K.; CHAN, S.; CHALMERS, T. C.; SACKS, H. S.; SMITH, H. Publication bias and clinical trials. *Controlled Clinical Trials*, v.8, n.4, p.343-353, 1987.
- EASTERBROOK, P. J.; BERLIN, J. A.; GOPALAN, R.; MATTEWS, D. R. Publication bias in clinical research. *Lancet*, London, v.337, n.8746, p.867-872, 1991.
- EGGER, M.; SMITH, G. D. Meta-analysis: potentials and promise. *Br. Med. J.*, Baarn, v.315, n.7119, p.1371-1374, 1997.
- EGGER, M.; SMITH, G. D. Meta-analysis bias in location and selection of studies. *Br. Med. J.*, Baarn, v.316, n.7124, p.61-66, 1998.
- EGGER, M.; SMITH, G. D.; SCHNEIDER, M.; MINDER, C. E. Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.*, Baarn, v.315, n.7109, p.629-634, 1997.

- ENGELS, E. A.; SCHMID, C. H.; TERRIN, N.; OLKIN, I.; LAU, J. Heterogeneity and statistical significance in meta-analysis. *Statistics in analysis: an empirical study of 125 meta-analyses*. *Stat. Med.*, New York, v.19, n.13, p.1707-1728, 2000.
- FREIMAN, J. A.; CHALMERS, T. C.; SMITH, H.; KUEBLER, R. R. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. *N. Engl. J. Med.*, Boston, v.299, n.13, p.690-694, 1978.
- GLASS, G. V.; MCGAW, B.; SMITH, M. L. *Meta-analysis in social research*. Beverly Hills: Sage, 1981. 280p.
- GLASS, G. V. Primary, secondary, and meta-analysis of research. *Educ. Res.*, Thousand, v.5, p.3-8, 1976.
- HAINES, S. J.; WALTERS, B. C. What is metanalysis? *Surg. Neurol.*, New York, v.44, n.6, p.581-582, 1995.
- HENMI, M.; COPAS, J. B.; EGUCHI, S. Confidence intervals and P-values for meta-analysis with publication bias. *Biometrics*, Washington, v.63, n.2, p.475-482, 2007.
- HIGGINS, J. P. T.; THOMPSON, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.*, New York, v.21, n.11, p.1539-1558, 2002.
- HIGGINS, J. P. T.; THOMPSON, S. G.; DEEKS, J. J.; ALTMAN, D. G. Measuring inconsistency in meta-analyses. *Br. Med. J.*, Baarn, v.327, n.7414, p.557-560, 2003.
- HOLCOMB Jr, H. L.; CHAIWORAPONGSA, T.; LUKE, D. A.; BURGDORF, K. D. An odd measure of risk: use and misuse of the odds ratio. *Obstet. Gynecol.*, Philadelphia, v.98, n.4, p.685-688, 2001.
- HUEDO-MEDINA, T. B.; SÁNCHEZ-MECA, J.; MARÍN-MARTINEZ, F. ; BOTELLA, J. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index. *Psychol. Methods*, Washington, v. 11, n.2, p.193-206, 2006.
- KULIK, J. A.; KULIK, C. L. C. Meta-analysis in education. *Int. J. Educ. Res.*, Ruston, v.13, p.221-340, 1989.
- LEWIS, S.; CLARKE, M. Forest plots: trying to see the wood and the trees. *Br. Med. J.*, Baarn, v.322, n.7300, p.1479-1480, 2001.
- LEWIS, J. A.; ELLIS, S. H. A statistical appraisal of post-infarction beta-blocker trials. *Primary Cardiology*, supl.1, p.31-37, 1982.
- MANTEL, N.; HAENSZEL, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Nat. Cancer Inst.*, Oxford, v.22, n.4, p.719-748, 1959.
- MARTINEZ, E. Z. Metanálise de ensaios clínicos controlados aleatorizados: aspectos quantitativos. *Rev. Med.*, Ribeirão Preto, v.40, n.2, p.223-235, 2007.
- MEDEIROS L. R; STEIN A. Níveis de evidência e graus de recomendação da medicina baseada em evidências. *Rev. AMRIGS*, Porto Alegre, v.46, p.43-46, 2002.

- MOJA, L.; MOSCHETTI, I.; LIBERATI, A.; GENSINI, G. F.; GUSINU, R. Understanding systematic reviews: the meta-analysis graph (also called 'forest plot'). *Intern. Emerg. Med.*, v.2, n.2, p.140-142, 2007.
- MOREIRA, R. O.; PAPELBAUM, M.; APPOLINARIO, J. C.; MATOS, A. G.; COUTINHO, W. F.; MEIRELLES, R. M. R.; ELLINGER, V. C. M.; ZAGURY, L. Diabetes mellitus e depressão: uma revisão sistemática. *Arq. Bras. Endocrinol. Metabol.*, São Paulo, v.47, n.1, p.19-29, 2003.
- MULROW, C. D. Systematic reviews: rationale for systematic reviews. *Br. Med. J.*, Baarn, v.309, n.6954, p.597-599, 1994.
- PEARSON, K. Report on certain enteric fever inoculation statistics. *Br. Med. J.*, Baarn, v.3, p.1243-1246, 1904.
- PETTITTI, D. B. *Meta-analysis, decision analysis, and cost-effectiveness analysis*. New York: Oxford University Press, 1994. 246p.
- POCOCK, S. *Clinical trials: a practical approach*. Chichester: Wiley, 1983. 266p.
- ROBINS, J.; GREENLAND, S.; BRESLOW, N. E. A general estimator for the variance of the Mantel-Haenszel odds ratio. *Am. J. Epidemiol.*, Oxford, v.124, n.5, p.719-723, 1986.
- ROTHMAN, K. J.; GREENLAND, S. *Modern epidemiology*. 2nd. ed.. Philadelphia: Lippincott-Raven, 1998. 737p.
- SCHECHTMAN, E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat. Which of these should we use? *Value Health*, Malden, v.5, n.5, p.431-436, 2002.
- SILVA FILHO, C. R.; SACONATO, H.; CONTERNO, L. A.; MARQUES, I.; ATALLAH, A. N. Avaliação da qualidade de estudos clínicos e seu impacto nas metanálises. *Rev. Saúde Pública*, São Paulo, v.39, n.6, p.865-873, 2005.
- SIDIK, K.; JONKMAN, J. N. A comparison of heterogeneity variance estimators in combining results of studies. *Stat. Med.*, New York, v 26, n.9, p.1964-1981, 2007.
- SIMON, S. D. Understanding the odds ratio and the relative risk. *J. Androl.*, Lawrence, v.22, n.4, p.533-536, 2001.
- SONG, F.; SHELDON, T. A.; SUTTON, A. J.; ABRAMS K. R.; JONES, D. J. Methods for exploring heterogeneity in meta-analysis. *Eval. Health Prof.*, London, v.24, n.2, p.126-151. 2001.
- TERRIN, N.; SCHMID, C. H; LAU, J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J. Clin. Epidemiol.*, New York, v.58, n.9, p.894-901, 2005.
- THORNTON, A.; LEE, P. Publication bias in meta-analysis: its causes and consequences. *J. Clin. Epidemiol.*, New York, v.53, n.2, p.207-216, 2000.
- XU, H.; PLATT, R. W.; LUO, Z. C.; WEI, S.; FRASER, W. D. Exploring heterogeneity in meta-analyses: needs, resources and challenges. *Paediatr. Perinat. Epidemiol.*, Oxford, v.22, sup.1, p.18-28, 2008.

YUSUF, S.; PETO, R.; LEWIS, J.; COLLINS, R.; SLEIGHT, P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog. Cardiovascular Dis.*, Orlando, v.27, n.5, p.335-371, 1985.

WHITEHEAD, A. *Meta-analysis of controlled clinical trials*. Chichester: John Wiley & Sons, 2002. 336p.

WOOLF, B. On estimating the relationship between blood groups and disease. *Ann. Hum. Genet.*, London, v.19: p.251-253, 1955.

Recebido em 28.01.2009.

Aprovado após revisão em 30.06.2009.

Apêndice

Estimação de τ^2 a partir da estatística Q

Uma maneira usual de estimar τ^2 é igualando o valor observado de Q (equação (6)) com sua esperança $E(Q)$ (DerSimonian e Laird, 1986; Higgins e Thompson, 2002). Observar que

$$\begin{aligned} Q &= \sum_{j=1}^J w_j (\hat{\theta}_j - \hat{\theta}_M)^2 = \sum_{j=1}^J w_j \left[(\hat{\theta}_j - \theta_M) - (\hat{\theta}_M - \theta_M) \right]^2 \\ &= \sum_{j=1}^J w_j \left[(\hat{\theta}_j - \theta_M)^2 - 2(\hat{\theta}_j - \theta_M)(\hat{\theta}_M - \theta_M) + (\hat{\theta}_M - \theta_M)^2 \right] \\ &= \sum_{j=1}^J w_j (\hat{\theta}_j - \theta_M)^2 - \left(\sum_{j=1}^J w_j \right) (\hat{\theta}_M - \theta_M)^2. \end{aligned}$$

A esperança $E(Q)$ é dada por

$$\begin{aligned} E(Q) &= \sum_{j=1}^J w_j E \left[(\hat{\theta}_j - \theta_M)^2 \right] - \left(\sum_{j=1}^J w_j \right) E \left[(\hat{\theta}_M - \theta_M)^2 \right] \\ &= \sum_{j=1}^J w_j \text{Var}(\hat{\theta}_j) - \left(\sum_{j=1}^J w_j \right) \text{Var}(\hat{\theta}_M). \end{aligned}$$

Uma expressão para a variância de Q é apresentada por Biggerstaff e Tweedie (1997). Considerando

$$\hat{\theta}_M = \frac{\sum_{j=1}^J w_j \hat{\theta}_j}{\sum_{j=1}^J w_j} \sim N \left(\theta_M; \frac{1}{\sum_{j=1}^J w_j} + \frac{\sum_{j=1}^J w_j^2}{\left(\sum_{j=1}^J w_j \right)^2} \tau^2 \right)$$

(ver Biggerstaff e Tweedie, 1997), temos

$$\begin{aligned} E(Q) &= \sum_{j=1}^J w_j (\tau^2 + w_j^{-1}) - \left(\sum_{j=1}^J w_j \right) \left(\frac{1}{\sum_{j=1}^J w_j} + \frac{\sum_{j=1}^J w_j^2}{\left(\sum_{j=1}^J w_j \right)^2} \tau^2 \right) \\ &= \tau^2 \sum_{j=1}^J w_j + J - 1 - \frac{\sum_{j=1}^J w_j^2}{\sum_{j=1}^J w_j} \tau^2 \\ &= J - 1 + \left(\sum_{j=1}^J w_j - \frac{\sum_{j=1}^J w_j^2}{\sum_{j=1}^J w_j} \right) \tau^2. \end{aligned}$$

Ao igualarmos o valor observado de Q com $E(Q)$, temos um estimador de τ^2 dado por

$$\hat{\tau}^2 = \frac{Q - (J - 1)}{\sum_{j=1}^J w_j - \frac{\sum_{j=1}^J w_j^2}{\sum_{j=1}^J w_j}},$$

que corresponde à equação (10) (ver Higgins e Thompson, 2002). Se assumirmos $\tilde{\sigma}^2 = w_j^{-1}$, temos

$$\hat{\tau}^2 = \tilde{\sigma}^2 \left(\frac{Q}{J-1} - 1 \right),$$

que corresponde à equação (7).