

Compreendendo os estudos de metanálise na pesquisa clínica

Understanding meta-analysis studies in clinical research

Leonardo Roever¹

Recebido da Universidade Federal de Uberlândia.

RESUMO

Metanálise é uma técnica estatística utilizada para combinar resultados provenientes de diferentes estudos. Para isto, são descritas as medidas de efeito utilizadas em metanálise na área da saúde, bem como os modelos de efeito fixo e de efeitos aleatórios. O objetivo deste artigo foi apresentar, de forma metodológica, como realizar e interpretar uma metanálise.

Descritores: Metanálise; Medicina baseada em evidências; Bioestatística

ABSTRACT

Meta-analysis is a statistical technique used to combine results from different studies. For this, effect measures used in meta-analysis of health studies are described, as well as models of fixed effects and random effects. The aim of this paper is to methodologically present ways to perform and interpret a meta-analysis.

Keywords: Meta-analysis; Evidence-based medicine; Biostatistics

INTRODUÇÃO

A metanálise é utilizada para combinar resultados provenientes de diferentes estudos e, com isso, produzir estimativas que resumem o todo. Isto é denominado de estimativas metanalíticas. Para se considerar que o resultado de uma metanálise tenha significado aplicado, os estudos que compõem seus dados devem ser o resultado de uma revisão sistemática, que consiste em um conjunto de regras para identificar estudos sobre uma determinada questão e, assim, selecionar quais deles serão incluídos. No quadro 1, estão os elementos de uma revisão sistemática.⁽¹⁻¹⁴⁾

O objetivo deste artigo é apresentar os conceitos básicos e a forma metodológica para realização de uma metanálise, sem a utilização de fórmulas complexas, o que deve facilitar a compreensão do profissional de saúde.

A COMPOSIÇÃO DOS DADOS

Primeiramente, é necessário definir quais os resultados serão combinados para que, por sua vez, os estudos possam ser combinados, por meio de uma metanálise. Estes estudos podem comparar diferentes tecnologias (drogas, dispositivos, tratamentos, procedimentos etc.) e apresentam medidas do tamanho do efeito como a razão de chances, o risco relativo, a redução absoluta de risco e o número necessário para tratar.^(1,12-14)

MEDIDAS DE EFEITO

Para que estudos possam ser combinados em uma metanálise, é preciso escolher uma medida de efeito e calculá-la para cada estudo. Estas medidas são classificadas em quatro categorias: (I) medidas de efeito para descrever apenas um grupo com uma só variável binária, (II) medidas de efeito para comparar dois grupos por meio de variáveis binárias, (III) medidas de efeito para comparar dois grupos por meio de variáveis contínuas e (IV) medidas de efeito para grupos com dados categóricos ordinais.^(1,12-14)

Medidas de efeito para um grupo

São utilizadas quando o interesse é descrever um estudo ou estimar um efeito. As medidas de efeito mais utilizadas para um grupo são o *odds* e a incidência. Estimativas destas medidas são calculadas, e seu uso em metanálise requer também uma medida de variabilidade.^(1,12-14)

Odds ou chances

Por definição, *odds* é a razão entre a probabilidade de o evento ocorrer e a probabilidade do evento não ocorrer, o que é equivalente à razão entre o número de eventos e o número de não eventos.^(1,12-14)

Incidência

A incidência mede o número de novos casos surgidos em uma determinada população e um determinado intervalo de tempo.^(1,12-14)

1. Universidade Federal de Uberlândia, Uberlândia, MG, Brasil.

Data de submissão: 02/05/2016 – Data de aceite: 16/05/2016

Conflito de interesse: não há.

Endereço para correspondência:

Leonardo Roever

Rua Rafael Rinaldi, 431 – Martins

CEP: 38400-384 – Uberlândia, MG, Brasil

Fone: (34) 98803-9878 – E-mail: leonardoroever@hotmail.com

Medidas de efeito para comparação de dois grupos por meio de variáveis binárias

Os dados provenientes de estudos que comparam dois grupos quanto a uma variável binária (ocorrência ou não de um evento de interesse), geralmente são organizados em tabelas 2x2 (Quadro 2).⁽⁵⁾

As medidas de efeito mais comuns calculadas para este tipo de dados são *odds ratio*, risco relativo, redução absoluta de risco e número necessário para tratar (NNT).^(1,12-14)

Odds ratio ou razão de chances

O *odds ratio* expressa quantas vezes o *odds* de um grupo equivale ao *odds* do outro grupo. Um *odds ratio* igual a 1 indica que não existe diferença entre os *odds* dos dois grupos; um *odds ratio* menor que 1 indica que o grupo 1 tem *odds* menor que o grupo 2 e um *odds ratio* maior que 1 indica o contrário.^(1,12-14)

Odds ratio de Peto ou razão de chances de Peto

O *odds ratio* de Peto é uma forma alternativa para se estimar o *odds ratio*. A ideia e a interpretação são as mesmas, mas uma diferença relevante é que, no *odds ratio* de Peto, a, b, c ou d podem ser iguais a zero. O *odds ratio* de Peto é particularmente útil quando os tamanhos das amostras dos grupos não são muito diferentes.^(1,12-14)

Risco relativo

O risco relativo é definido como a razão entre a probabilidade do evento ocorrer no grupo 1 e a probabilidade do evento ocorrer no grupo 2. Um risco relativo igual a 1 indica que os dois grupos apresentam o mesmo risco, um risco relativo menor que 1 indica que o grupo 1 tem risco menor que o grupo 2 e um risco relativo maior que 1 indica o contrário.^(1,12-14)

Quadro 1. Elementos de uma revisão sistemática

Definir a questão clínica
Identificar todos os estudos completos sobre a questão, publicados ou não
Selecionar os estudos que preenchem padrões elevados de validade científica
Procurar por evidências de viés nos estudos selecionados
Descrever a qualidade científica dos estudos
Questionar se a qualidade está sistematicamente relacionada aos resultados do estudo
Descrever os estudos com uma figura (gráfico de floresta)
Decidir se os estudos são suficientemente semelhantes para justificar uma combinação
Se eles forem suficientemente semelhantes para serem combinados, calcular uma medida sumária de efeito e seu intervalo de confiança

Quadro 2. Apresentação dos dados dos estudos

Dados do estudo	Evento	Não evento	Tamanho dos grupos
Grupo 1	A	B	AB
Grupo 2	C	D	CD

Redução absoluta de risco

A redução absoluta de risco é definida como a diferença entre as probabilidades de ocorrer o evento nos dois grupos.^(1,12-14)

Número necessário para tratar

O NNT é o número de indivíduos que devem ser tratados a fim de prevenir um evento. O NNT é dado pelo inverso da redução absoluta de risco.^(1,12-14)

Medidas de efeito para comparação de dois grupos por meio de variáveis contínuas

Quando os dados de estudos são medidos em uma escala contínua, é comum aparecerem médias e desvios padrão para resumir as informações. Utilizando média e desvio padrão, as medidas de efeito mais comuns para dados contínuos são a diferença absoluta entre médias e a diferença padronizada entre médias.^(1,12-14)

Diferença absoluta entre médias

A diferença absoluta entre as médias quantifica a diferença entre a média do grupo 1 e a média do grupo 2. Esta medida deve ser utilizada quando todos os estudos da metanálise utilizam a mesma escala contínua.^(1,12-14)

Diferença padronizada entre médias

A diferença padronizada entre médias pode ser utilizada quando os dados são contínuos, mas usam diferentes escalas. Existem três maneiras diferentes de se estimar a diferença padronizada entre médias: d de Cohen, g ajustado de Hedges e Δ de Glass. Estes métodos diferem quanto ao uso do desvio padrão nos cálculos, e ao uso ou não de correção para evitar vício quando as amostras são pequenas (Quadro 3).^(1,12-14)

MODELOS DE EFEITO FIXO E ALEATÓRIOS

Em metanálise, são utilizados basicamente dois tipos de modelos: os de efeito fixo e os de efeitos aleatórios.^(1,12-14)

Efeito fixo

Pressupõe que o efeito de interesse é o mesmo em todos os estudos, e que as diferenças observadas entre eles são devidas à variabilidade interna de cada estudo.^(1,12-14)

Quadro 3. Maneiras diferentes de se estimar a diferença padronizada entre médias

Teste	Função
d de Cohen	O d de Cohen é a diferença média de cada estudo dividida pelo desvio padrão, criando uma medida padrão que possa ser comparável entre os estudos
g ajustado de Hedges	O g ajustado de Hedges é similar ao d de Cohen, porém inclui um ajuste para corrigir o vício de amostras pequenas
Δ de Glass	O Δ de Glass utiliza o desvio padrão do grupo controle para dividir a diferença de médias

Efeitos aleatórios

Pressupõe que o efeito de interesse não é o mesmo em todos os estudos. Estes estudos são parte de amostra aleatória de uma população hipotética de estudo, apesar dos efeitos dos estudos não serem considerados iguais, eles são conectados por meio de uma distribuição de probabilidade, geralmente suposta como normal.^(1,12-14)

MÉTODO DE MANTEL-HAENSZEL

Este método foi proposto para combinar *odds ratios* oriundo de diferentes estratos quando não são detectadas diferenças significativas entre eles, e é utilizado para combinar medidas de efeito oriundas de diferentes estudos quando a suposição é de que o efeito é o mesmo em todos eles, ou seja, a mesma suposição do modelo de efeito fixo. Quando comparado com o modelo de efeito fixo, o método de Mantel-Haenszel é mais robusto quando as amostras forem pequenas, visto que, nestas situações, produz estimativas melhores para os erros padrão das medidas de efeito.^(1,12-14)

MODELO DE EFEITOS ALEATÓRIOS

É utilizado quando se deseja combinar os dados que são parecidos ao ponto de ser coerente juntar a informação deles em uma medida resumo. Mas existem razões para que os estudos não sejam considerados idênticos quanto ao efeito do tratamento como, por exemplo, se o interesse for estudos em que está sendo testada a eficácia de uma droga, exercício ou dispositivo, pois, entre eles, pode haver diferença nos grupos de pessoas selecionadas, um grupo do estudo pode ser mais saudável, a faixa etária dos pacientes pode variar de estudo para estudo, a classe social pode ser diferente, entre outros fatores que podem influenciar para que o efeito do tratamento seja diferente entre os estudos. Quando isso acontece, estas diferenças devem ser investigadas, e alternativas como metanálise em subgrupos e metarregressão devem ser consideradas. A alternativa, nesta situação, é a utilização de um modelo de efeitos aleatórios, que considera a existência não apenas da variação dentro de cada estudo, mas também a variação entre os estudos, ou seja, ele considera que os efeitos dos estudos não são iguais, mas que são conectados por meio de uma distribuição de probabilidade, geralmente suposta como normal.^(1,12-14)

ESCOLHA DO MODELO

O modelo de efeito fixo é adequado quando acreditamos que o efeito de tratamento é idêntico entre os estudos, quando o objetivo for estimar um efeito de tratamento para uma população específica e não extrapolar para outras populações.^(1,12-14)

O modelo de efeitos aleatórios pode ser utilizado se combinar vários estudos que têm o mesmo objetivo, mas que não foram conduzidos da mesma maneira. Neste caso, é possível extrapolar para outras populações, o que torna a análise mais abrangente. Um ponto que deve ser verificado é que, se o número de estudos for muito pequeno, a estimativa da variância entre os estudos não terá uma boa precisão. Neste caso devem-se relatar os efeitos separados, e não como uma medida-resumo,

e utilizar um modelo de efeito fixo, e não fazer inferências para outras populações. Podemos também utilizar a abordagem Bayesiana, com a estimativa da variância entre os estudos podendo ser baseada também em dados externos aos estudos utilizados. Quanto à heterogeneidade, ela ajuda na escolha do modelo, mas não é determinante.^(1,12-14)

ABORDAGEM BAYESIANA

O método bayesiano na metanálise é uma técnica recente que ainda vem sendo explorada. Suas vantagens básicas para a preferência são: (I) a possibilidade de considerar o conhecimento prévio, (II) a flexibilidade da modelagem e a (III) naturalidade da interpretação dos resultados. No enfoque da metanálise, quando o número de estudos selecionados na revisão sistemática é pequeno, têm-se dificuldades de estimar as medidas de efeito metanalíticas por meio da teoria clássica, pois esta depende de pressupostos assintóticos. Nestes casos, a metanálise bayesiana também é sugerida como uma possível solução para obtenção de estimativas combinadas mais confiáveis. O quadro 4 mostra breve comparação entre inferência clássica e inferência Bayesiana.^(1,5,15)

A estatística bayesiana pode ser considerada um processo de diminuição de incerteza sobre o desconhecido, que se baseia em dados estatísticos e em evidências prévias.

METANÁLISE DE SUBGRUPOS

A metanálise é utilizada para se estimar o efeito médio do tratamento, mas, por mais bem feita que seja a revisão sistemática, ainda pode haver diferenças significativas entre os estudos: tratamento aplicado, qualidade do controle, estimador utilizado, etc.; diferenças que devem não apenas ser consideradas no planejamento da metanálise, mas também explicadas.^(1,12-16)

A metanálise de subgrupos é feita pela distinção de grupos de estudos. Trata-se de um método de segmentação de tal modo que exista grande homogeneidade dentro de cada grupo e heterogeneidade entre eles. É feita uma comparação aos estudos individuais, que têm um maior poder estatístico de detectar uma possível existência de subgrupos com efeitos diferentes, não detectados pelos estudos individuais, ou ainda refutar a alegação de efeitos dependentes de subgrupos. No entanto, o tamanho amos-

Quadro 4. Comparação entre inferência clássica e inferência bayesiana.⁽¹⁵⁻¹⁶⁾

Inferência clássica	Inferência Bayesiana
O parâmetro desconhecido é considerado uma constante fixa	O parâmetro desconhecido é considerado uma variável aleatória que segue uma distribuição <i>a priori</i>
É considerada somente a informação amostral por meio da verossimilhança	Podem-se considerar, <i>a priori</i> , informações de estudos anteriores, conhecimento pessoal etc.
Não se pode falar em probabilidade para as estimativas dos intervalos de confiança	Pode-se falar em probabilidade para as estimativas dos intervalos de credibilidade

tral em análise de subgrupos é frequentemente menor e, por isso, essas análises geralmente perdem em poder estatística.^(1,12-16)

Em um conjunto de estudos clínicos no qual exista a intervenção de diferentes drogas nos grupos de tratamento, o objetivo é realizar uma metanálise para a medida de efeito, mas deve-se considerar previamente uma análise de agrupamento para que esta heterogeneidade seja explicada. Com isto, deve-se definir o modelo de efeito a ser considerado em cada subgrupo: efeito fixo (quando não há heterogeneidade significativa dentro do subgrupo) ou efeitos aleatórios (quando há heterogeneidade). Nestes modelos, o efeito do tratamento é avaliado no subgrupo definido ou em um subgrupo complementar.^(1,12-16)

Na análise bayesiana, uma das vantagens sobre o enfoque clássico é a possibilidade de estimar o chamado intervalo de credibilidade fornecido por meio da distribuição *a posteriori*, e outra vantagem agregada à distribuição *a posteriori* é o cálculo de probabilidades. Com isto, é possível saber qual a probabilidade do *odds* tratamento ser maior que o *odds* controle, ou seja, além da estimativa pontual e do intervalo de credibilidade, também temos como resultados as probabilidades de ocorrências do *odds ratio*. No modelo de análise de subgrupos sob enfoque bayesiano, estas vantagens são estendidas para a diferença entre os subgrupos e, com isto, pode-se obter a estimativa para o intervalo de credibilidade e a probabilidade da diferença entre os subgrupos, o que não é possível no enfoque clássico.^(1,12-16)

MODELO DE EFEITO FIXO EM METANÁLISE DE SUBGRUPOS

O modelo de efeito fixo em metanálise de subgrupos possui as mesmas considerações vistas anteriormente. As diferenças entre os estudos são devidas apenas à variabilidade interna de cada um, isto é, a variabilidade entre os estudos é considerada nula.^(1,12-16)

MODELO DE EFEITO ALEATÓRIO EM METANÁLISE DE SUBGRUPOS

O modelo de efeitos aleatórios considera a variação dentro de cada estudo e a variação entre os estudos. Pode-se especificar o modelo de efeitos aleatórios para subgrupos da forma de um modelo hierárquico.^(1,12-16)

METARREGRESSÃO

A metarregressão é uma técnica utilizada para verificar se as características dos estudos empregados na análise de sensibilidade em metanálises permitem avaliar o impacto das covariáveis no resultado metanalisado, por meio de técnicas de regressão e, com isso, explicar a heterogeneidade existente entre estudos. A variável dependente é o tamanho do efeito, e as covariáveis são as características presentes nos estudos, por exemplo: idade, proporção de indivíduos, duração dos estudos, número de terapêuticas etc. Na metarregressão, o número de covariáveis a incluir está limitado ao número de estudos considerados na metanálise. Idealmente deve ser utilizada uma covariável por cada dez estudos.⁽¹⁾

HETEROGENEIDADE

Em estudos de metanálise, quando consideramos que a variabilidade entre os estudos não é apenas aleatória, dizemos que os estudos são heterogêneos. A avaliação sobre heterogeneidade tem papel fundamental na escolha do modelo de metanálise e deve ser realizada antes da escolha do modelo.^(1,12-16)

As maneiras mais utilizadas de se verificar a existência de heterogeneidade em metanálises são pelo teste Q de Cochran ou pela estatística I^2 de Higgins e Thompson. Quando nos referimos à heterogeneidade em medidas de efeito, pensamos em sua verdadeira variação, mas essa variação é constituída da verdadeira heterogeneidade e também do erro aleatório.^(1,12-16)

Separação da variação

Os efeitos observados variam pela heterogeneidade na medida de efeito e também pelo erro aleatório dentro do estudo. Para quantificar a heterogeneidade, é necessário separar a variação nesses dois componentes.^(1,12-16)

Para separar essas variações, é necessário calcular a variação total observada de estudo a estudo e estimar o quanto os efeitos observados devem variar supondo homogeneidade entre os estudos e, então, calcular o excesso de variação pela diferença da variação total pela variação, supondo homogeneidade – esta é a heterogeneidade.^(1,12-16)

Teste Q de Cochran

O teste Q de Cochran apresenta como hipótese nula a afirmação de que os estudos são homogêneos. Uma deficiência deste teste é possuir um baixo poder quando o número 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.^(1,12-16)

Estatística I^2

A estatística I^2 foi proposta como uma forma mais direta de verificar a existência de heterogeneidade por meio de proporção. Quando o valor for negativo, ele é igualado a zero. A significância de I^2 é equivalente a significância de Q. Na estatística I^2 uma escala em que um valor de I^2 próximo a 0% indica não heterogeneidade entre os estudos, próximo a 25% indica baixa heterogeneidade, próximo a 50% indica heterogeneidade moderada e próximo a 75% indica alta heterogeneidade entre os estudos.^(1,12-16)

Forest plot

O *forest plot*, ou gráfico floresta, é um gráfico que mostra informações individuais de cada estudo incluído na metanálise e também da medida metanalítica. Para cada estudo, o gráfico apresenta a medida de efeito e seu intervalo de confiança, sendo que a medida de efeito é representada por um símbolo que pode ser um quadrado, um círculo, ou outra figura dependendo do *software*. O tamanho deste símbolo é proporcional ao peso do estudo na metanálise, quanto maior o peso, maior o tamanho do símbolo da medida de efeito. Também é exibida em torno da medida de efeito uma linha horizontal, que é seu intervalo de confiança, sendo que, quanto maior esta linha, maior é a variabilidade dentro do estudo.^(1,12-17)

CONCLUSÃO

Este artigo abordou as questões básicas de uma metanálise, como as medidas de efeito, os modelos de efeito fixo e de efeitos aleatórios, como identificar heterogeneidade, análise bayesiana e de subgrupos entre os estudos e de como escolher o modelo a ser utilizado na área da saúde. A compreensão desta da metodologia e interpretação da metanálise é de fundamental importância em pesquisa na área da saúde.

REFERÊNCIAS

1. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. New York: John Wiley & Sons; 2009.
2. Brockwell SE, Gordon RI. A comparison of statistical methods for meta-analysis. *Statist Med.* 2001;20(6):825-40.
3. Rodrigues CL, Ziegelmann KP. Metanálise: um guia prático. *Rev HCPA.* 2010;30(4):436-47.
4. Egger M, Smith GD, Altman DG. Systematic reviews in health care: meta-analysis in context. *BMJ.* 2001;323(7304):101-5.
5. Fletcher RH, Fletcher SW. *Epidemiologia clínica: elementos essenciais.* Porto Alegre: Artmed; 2013.
6. Glass GV. Primary, secondary, and meta-analysis of research. *Educational Res.* 1976;5:3-8.
7. Hedges LV, Olkin I. *Statistical methods for meta-analysis.* New York: Academic Press; 1985.
8. Kulinskaya E, Morgenthaler S, Staudte RG. *Meta analysis: a guide to calibrating and combining statistical evidence.* New York: John Wiley & Sons; 2008.
9. Leandro G. *Meta-analysis in medical research: The handbook for the understanding and practice of meta-analysis.* London: BMJ Books; 2005.
10. Martinez EZ. Metanálise de ensaios controlados aleatorizados: aspectos quantitativos. *Medicina (Ribeirão Preto).* 2007;40(2): 223-35.
11. Normand SL. Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med.* 1999;18:321-59.
12. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research.* New York: John Wiley & Sons; 2000.
13. Whitehead A. *Meta-analysis of controlled clinical trials.* New York: John Wiley & Sons; 2002.
14. Berman NG, Parker RA. Meta-analysis: neither quick nor easy. *BMC Med Res Methodol.* 2002;2:10.
15. Ziegelmann PK. *Inferência Bayesiana e Teoria das Decisões.* Porto Alegre: Universidade Federal do Rio Grande do Sul. Instituto de Matemática. Departamento de Estatística; 2010.
16. Ziegelmann PK. Metanálise MTC: O uso combinado de evidência direta e indireta na comparação de múltiplos tratamentos. In: 19º SINAPE, 2010 Julho 26-30. São Pedro, SP, Brasil.
17. Ziegelmann PK. Aspectos conceituais e práticos sobre metanálise. *Rev Enferm UFPI.* 2013;2(Esp.):8-13.