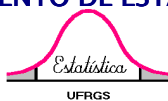




UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA



Metanálise: um guia prático

Autora: Caroline Legramanti Rodrigues
Orientadora: Professora Dra. Patrícia Klarmann Ziegelmann

Porto Alegre, 28 de junho de 2010.

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Departamento de Estatística

Metanálise: um guia prático

Autora: Caroline Legramanti Rodrigues

Monografia apresentada para obtenção
do grau de Bacharel em Estatística.

Banca Examinadora:
Professora Dra. Patrícia Klarmann Ziegelmann
Professora Dra. Luciana Neves Nunes

Porto Alegre, 28 de junho de 2010.

Dedico este trabalho aos meus pais, Enio e Jovilde, que sempre colocaram a educação dos filhos acima de qualquer outra coisa e que me ensinaram que o conhecimento é o que eu posso adquirir de mais valioso, pois ninguém jamais poderá tirá-lo de mim.

“O mundo está nas mãos daqueles que tem a coragem de sonhar e de correr o risco de viver seus sonhos.” (Paulo Coelho)

Agradecimentos

À professora Patrícia pela paciência e dedicação não só na orientação deste trabalho, mas também por ter sido minha tutora no início do curso, pelo ano em que fui monitora de suas disciplinas e pelas ótimas aulas de Inferência Bayesiana. Será sempre uma referência para mim.

À professora Luciana por ter aceitado fazer parte da minha banca e por ser uma das professoras mais próxima dos alunos que eu conheço.

Ao amor da minha vida, Jorge, por ter aguentado meu mau humor quando eu ia mal em alguma disciplina, minha falta de tempo, principalmente neste último semestre, e por não tocar guitarra ou bateria quando eu precisava estudar.

À minha família que sempre apostou e torceu por mim.

Aos colegas e amigos que fiz durante estes anos de Estatística, em especial à Simone Lumertz que se tornou minha amiga inseparável desde o primeiro dia de aula; à Julia Bastiani que tentou (sem sucesso) me ensinar como cursar Estatística sem se estressar, acho que eu fui quem ensinou ela a ficar um pouco estressada; à Lisiane Moura por ser a melhor professora particular que eu já tive e ao Gilberto Beuren por me fazer rir, mesmo quando o momento não era apropriado. E mais que especial à minha querida amiga Marjana Peixoto que me acompanhou nos melhores e piores momentos do curso, mesmo agora que ela já está formada, ela está sempre me consolando e motivando, certamente a amizade dela é o que eu vou levar de melhor destes anos de UFRGS.

Ao Renato Carlson que foi meu professor de Estatística na Unisinos, quando eu fazia Administração, o culpado (no bom sentido) por eu ter mudado de curso e que tem acompanhado minha trajetória na Estatística.

Aos meus chefes e colegas da Cientec e da Siqueira Campos, estágios por onde passei.

Aos meus chefes e colegas de CRM da Renner que me contrataram, mesmo sabendo que eu ainda não estava formada e que precisava sair no meio do trabalho pra ir à aula.

A todos os meus amigos de fora da UFRGS que me apoiaram, mesmo os que estão longe, mas sempre torcendo por mim.

A todos os professores do departamento de Estatística, até mesmo aos poucos e maus professores que me fizeram dar mais valor ainda aos bons.

Resumo

Este trabalho tem como objetivo servir como um guia prático para pessoas que queiram conduzir uma metanálise. Para isto, são descritas as medidas de efeito mais comumente utilizadas em metanálises na área da saúde, bem como os modelos de efeito fixo e de efeitos aleatórios. Alguns conceitos de heterogeneidade e de como detectá-la também são apresentados. Para mostrar um passo a passo de como conduzir uma metanálise são apresentados três exemplos, cada um com diferentes tipos de estudos a serem combinados, sendo o primeiro com estudos para comparar dois grupos com dados binários, o segundo para comparar dois grupos com dados contínuos e o terceiro para um grupo com dados binários, estes exemplos foram feitos com os softwares STATA e R.

Abstract

The goal of this paper is to serve as a practical guide for people who want to conduct a meta-analysis. For this, the most commonly effect measures used in meta-analysis of medical area are described. Fixed effect models and random effects models are described too. Some concepts of heterogeneity and how to detect it are also presented. To show a step by step how to conduct a meta-analysis are presented three examples, each one with different types of studies to be combined, the first with studies to compare two groups with binary data, the second to compare two groups with continuous data and the third for a group with binary data, these examples were done with STATA and R.

Sumário

1. Introdução.....	9
2. Medidas de efeito.....	11
2.1. Medidas de efeito para um grupo.....	11
2.1.1. <i>Odds</i> ou chances.....	12
2.1.2. Incidência.....	12
2.2. Medidas de efeito para comparação de dois grupos através de variáveis binárias.....	12
2.2.1. <i>Odds ratio</i> ou razão de chances.....	13
2.2.2. <i>Odds ratio</i> de Peto ou razão de chances de Peto.....	14
2.2.3. Risco relativo.....	14
2.2.4. Redução absoluta de risco.....	15
2.2.5. Número necessário para tratar.....	15
2.3. Medidas de efeito para comparação de dois grupos através de variáveis contínuas.....	16
2.3.1. Diferença absoluta entre médias.....	16
2.3.2. Diferença padronizada entre médias.....	17
3. Modelos de efeito fixo e modelos de efeitos aleatórios.....	18
3.1. Modelo de efeito fixo.....	18
3.1.1. Estimação da medida metanalítica θ_M	19
3.1.2. Método de Mantel-Haenszel.....	19
3.2. Modelo de efeitos aleatórios.....	22
3.2.1. Estimação da medida metanalítica θ_M^*	23
3.2.2. Método de DerSimonian e Laird.....	24
4. Heterogeneidade.....	25
4.1. Separação da variação.....	26
4.2. Teste Q de Cochran.....	26
4.3. Estatística I^2	27

5. <i>Forest plot</i>	27
6. Escolha do modelo.....	28
7. Metanálise no STATA.....	28
7.1. Comando <i>metan</i>	29
7.2. Exemplo: comparação de dois grupos com dados binários.....	30
7.3. Exemplo: comparação de dois grupos com dados contínuos.....	36
7.4. Exemplo: um grupo com dados binários.....	41
8. Metanálise no R.....	46
8.1. Pacote <i>metafor</i>	46
8.2. Exemplo: comparação de dois grupos com dados binários.....	47
8.3. Exemplo: comparação de dois grupos com dados contínuos.....	49
9. Considerações finais.....	51
Referências Bibliográficas.....	52

1. Introdução

A pesquisa é algo recorrente em todas as áreas e, hoje em dia, com a facilidade dos meios de comunicação, é possível encontrar inúmeros estudos sobre a mesma questão de pesquisa. Com uma infinidade de estudos disponíveis surgem alguns problemas como, quais estudos escolher para a pesquisa, como tirar conclusões se alguns divergem entre si, quais têm as melhores conclusões, entre outras dúvidas que podem surgir. Uma maneira de resolver estes problemas é combinar os resultados dos estudos.

Uma maneira de se pensar quando se deseja combinar vários estudos é simplesmente juntar os dados dos estudos e analisar como se fossem informações retiradas de apenas um amplo estudo. Porém, esta maneira desconsidera completamente a variabilidade e o tamanho de cada estudo, que tem importante influência no resultado e deveriam ser considerados. A maneira mais apropriada é utilizar técnicas estatísticas adequadas. Estas técnicas serão detalhadas neste trabalho e englobam o que é conhecido na literatura como metanálise. A metanálise é um tipo de análise de dados em que os resultados de vários estudos, que abordam a mesma questão de pesquisa são combinados, gerando assim, estimativas que resumem o todo chamadas de estimativas metanalíticas.

Os estudos que compõem uma metanálise são escolhidos através de uma revisão sistemática. Uma revisão sistemática consiste de um conjunto de regras para identificar estudos sobre uma determinada questão e, em seguida, selecionar quais deles serão incluídos ou não na metanálise. Para que a revisão sistemática seja útil, são necessários muitos estudos sobre a questão, mas que não tenham tanta concordância entre si a ponto da questão já estar respondida. Na Tabela 1.1 estão os elementos de uma revisão sistemática (Fletcher & Fletcher, 2006).

Tabela 1.1 - Elementos de uma Revisão Sistemática

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

Não são recentes as tentativas de unir os resultados de diferentes estudos. A primeira metanálise da história é atribuída ao estatístico Karl Pearson. Em 1904, Pearson combinou, através de correlações, os dados de cinco estudos para examinar o efeito preventivo de inoculações contra febre entérica (Pearson, 1904). Os motivos relatados por Pearson para a combinação de estudos são, ainda hoje, as principais razões para o uso da metanálise. O principal argumento é que muitos dos estudos são pequenos para permitir uma conclusão confiável, fazendo com que o tamanho do erro seja grande e o poder do estudo seja baixo.

Em 1931 e 1932, Tippet e Fisher, apresentaram métodos para a combinação de valores p (Whitehead, 2002). Em 1938, Yates e Cochran consideraram a combinação de estimativas a partir de diferentes experimentos agrícolas (Yates & Cochran, 1938), esses autores desdobraram essa fonte de variação para cada genótipo, determinando uma regressão linear da produtividade em relação à média de todos eles, por ambiente. A primeira metanálise para avaliar o efeito de uma intervenção terapêutica foi publicada em 1955 (Whitehead, 2002). Na década de 1970, a metanálise passou a ser usada nas ciências sociais, principalmente em pesquisas de educação. Porém, o termo “metanálise” ainda não era utilizado. Foi em 1977, que o termo “metanálise” foi utilizado pela primeira vez em um artigo intitulado “Primary, secondary and meta-analysis of research” pelo psicólogo Gene Glass (Glass, 1977).

Na área médica, a metanálise passou a ser utilizada com mais frequência na década de 1980. Questões chave da medicina passaram a ser respondidas com a metanálise, como o tratamento de doenças do coração e de câncer (Whitehead, 2002).

Em 1993, uma rede de clínicos, epidemiologistas e outros profissionais da saúde fundaram o Cochrane Collaboration, que tem por objetivo elaborar, manter e divulgar revisões sistemáticas da área da saúde. Desde sua fundação em Oxford, o Centro Cochrane tem crescido rapidamente e diversos centros foram fundados na Europa, América do Norte e Latina, África e Austrália com centenas de pessoas colaborando em grupos de revisão. No Brasil, o Centro Cochrane foi inaugurado em 1996 em São Paulo. Para conhecer mais sobre o Centro Cochrane e utilizar seus materiais de pesquisa, é possível acessar a página na internet <http://www.centrocochranedobrasil.org.br/>, na página são encontrados links para os outros centros e suas bibliotecas virtuais.

O uso da metanálise vem crescendo nas mais diversas áreas, como em ciências sociais, agricultura, ecologia, veterinária, entre outras; mas certamente, é a medicina a área que mais utilizou e utiliza metanálises até hoje (Whitehead, 2002).

O objetivo deste trabalho é apresentar os conceitos básicos de uma metanálise e servir de guia para pessoas que queiram conduzir uma com o uso dos softwares STATA e R, já que existem poucos materiais neste formato em português. Para o leitor que queira aprofundar mais seus conhecimentos em metanálise sugere-se Sutton, 2000; Borestein, 2009; Egger, 2009, entre outros (ver referências bibliográficas). O foco deste trabalho será voltado à área da saúde, mas lembrando que uma metanálise pode ser conduzida em qualquer outra área. Para este propósito no Capítulo 2 serão descritas algumas das medidas de efeito mais comumente utilizadas em metanálises na área da saúde, no Capítulo 3 serão abordados os modelos de efeito fixo e de efeitos aleatórios, no Capítulo 4 serão introduzidas algumas noções de heterogeneidade e de como detectá-la, no Capítulo 5 será apresentado o *forest plot*, no Capítulo 6 será dado um breve relato de como escolher o melhor modelo para conduzir uma metanálise, no Capítulo 7 será mostrado como conduzir uma metanálise utilizando o software STATA e no Capítulo 8 como conduzir uma metanálise no software R.

2. Medidas de efeito

Para que estudos possam ser combinados em uma metanálise, é necessário escolher uma medida de efeito e calculá-la para cada estudo, com isso eles poderão ser combinados através dessa medida.

Medidas de efeito podem ser classificadas de diferentes maneiras. Neste trabalho elas serão classificadas em quatro categorias: medidas de efeito para descrever apenas um grupo com uma só variável binária, medidas de efeito para comparar dois grupos através de variáveis binárias, medidas de efeito para comparar dois grupos através de variáveis contínuas e medidas de efeito para grupos com dados categóricos ordinais. Neste trabalho serão brevemente descritas algumas medidas de efeito para as três primeiras categorias. Para um estudo mais aprofundado destas medidas e das medidas de efeito para dados categóricos veja Sutton, 2000.

2.1. Medidas de efeito para um grupo

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.

2.1.1. Odds ou chances

Por definição *odds* é a razão entre a probabilidade de o evento ocorrer e a probabilidade do evento não ocorrer que, matematicamente, é equivalente a razão entre o número de eventos e o número de não eventos. Por questões teóricas, geralmente o *odds* é calculado em escala logarítmica. O logaritmo do *odds* de um estudo j é estimado por

$$\ln(\widehat{Odds})_j = \ln\left(\frac{a_j}{b_j}\right) \quad (2.1)$$

onde a_j é o número de indivíduos que sofreu o evento no estudo j e b_j é o número de indivíduos que não sofreu o evento no estudo j , com $j = 1, 2, \dots, J$.

A variância do logaritmo do *odds* é estimada por

$$\widehat{\text{var}}(\ln(\widehat{Odds}))_j = \frac{1}{a_j} + \frac{1}{b_j} \quad (2.2)$$

2.1.2. Incidência

A incidência mede o número de novos casos surgidos numa determinada população e num determinado intervalo de tempo. Ela também é usualmente calculada em escala logarítmica. O logaritmo da incidência de um estudo j é estimado por

$$\ln(\widehat{Incid})_j = \ln\left(\frac{d_j}{q_j}\right)$$

onde d_j são os eventos observados no estudo j em q_j unidades de tempo.

A variância do logaritmo da incidência é estimada por

$$\widehat{\text{var}}(\ln(\widehat{Incid}))_j = \frac{1}{d_j}$$

2.2. Medidas de efeito para comparação de dois grupos através 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 2 x 2, conforme a seguir.

Tabela 2.1 - Dados de um estudo j

	Eventos	Não eventos	Tamanho do grupo
Grupo 1	a_j	b_j	n_{1j}
Grupo 2	c_j	d_j	n_{2j}

Os dois grupos comparados (Grupo 1 e Grupo 2) podem ser o grupo tratado e o grupo controle de um ensaio clínico ou o grupo exposto e o grupo não exposto a determinado fator de um estudo longitudinal, ou ainda, eventos e não eventos podem ser os casos e os controles em um estudo caso-controle.

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.

2.2.1. Odds ratio ou razão de chances

O *odds ratio* expressa quantas vezes o *odds* de um grupo equivale ao *odds* do outro grupo. O *odds ratio* de um estudo j é estimado por

$$\widehat{OR}_j = \frac{a_j / b_j}{c_j / d_j} = \frac{a_j d_j}{b_j c_j}$$

onde a_j , b_j , c_j e d_j são as células da Tabela 2.1, com $j = 1, 2, \dots, J$. Por questões teóricas o *odds ratio* pode ser estimado em escala logarítmica, ou seja,

$$\ln(\widehat{OR}_j) = \ln\left(\frac{a_j d_j}{b_j c_j}\right)$$

O *odds ratio* varia de 0 a infinito. 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.

Uma maneira usual de se calcular a estimativa da variância do logaritmo do *odds ratio* é utilizar o método de Woolf, onde

$$\widehat{\text{var}}(\ln(\widehat{OR}_j)) = \frac{1}{a_j} + \frac{1}{b_j} + \frac{1}{c_j} + \frac{1}{d_j} \quad (2.3)$$

Note que na expressão (2.3) se algum dos eventos for igual a zero (a_j , b_j , c_j ou $d_j = 0$) ela é indefinida. Uma maneira usual de corrigir esse problema é somar 0,5 em cada célula da

Tabela 2.1. Neste caso, a estimativa do *odds ratio* também deve ser feita com os dados de todas as caselas somados de 0,5 (Sutton, 2000).

2.2.2. Odds ratio de Peto ou razão de chances de Peto

O *odds ratio* de Peto é uma forma alternativa de 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_j , b_j , c_j ou d_j podem ser iguais a zero.

O *odds ratio* de Peto é particularmente útil quando os tamanhos das amostras dos grupos não são muito diferentes.

O *odds ratio* de Peto do estudo j é estimado por

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

onde O_j é o número de respostas observadas no grupo tratamento, E_j é o número de respostas esperadas no grupo tratamento e V_j é uma estimativa da variância da diferença entre O_j e E_j , ou seja,

$$O_j = a_j$$

$$E_j = \frac{n_{1j}(a_j + c_j)}{n_{1j} + n_{2j}}$$

$$V_j = \frac{n_{1j}n_{2j}(a_j + c_j)(b_j + d_j)}{(n_{1j} + n_{2j})^2(n_{1j} + n_{2j} - 1)}$$

2.2.3. 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.

O risco relativo de um estudo j é estimado por

$$\widehat{RR}_j = \frac{a_j/n_{1j}}{c_j/n_{2j}} = \frac{a_j n_{2j}}{c_j n_{1j}}$$

Assim como o *odds ratio*, o risco relativo também é calculado muitas vezes na escala logarítmica por questões teóricas, ou seja,

$$\ln(\widehat{RR}_j) = \ln\left(\frac{a_j n_{2j}}{c_j n_{1j}}\right)$$

O risco relativo varia de 0 a infinito. 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.

Assim como no *odds ratio*, é usual calcular a estimativa da variância do logaritmo do risco relativo, estimada por

$$\widehat{\text{var}}(\ln(\widehat{RR}_j)) = \frac{1}{a_j} - \frac{1}{n_{1j}} + \frac{1}{c_j} - \frac{1}{n_{2j}} \quad (2.4)$$

Note que na expressão (2.4) se algum dos eventos a_j ou c_j for igual a zero, ela é indefinida. Neste caso, o problema também pode ser corrigido com a soma de 0,5 em cada célula da Tabela 2.1. Assim, a estimativa do risco relativo também deve ser feita com os dados de todas as caselas somados de 0,5.

2.2.4. Redução absoluta de risco

A redução absoluta de risco é definida como a diferença entre as probabilidades de ocorrer o evento nos dois grupos.

A redução absoluta do risco em um estudo j é estimada por

$$\widehat{RAR}_j = \frac{a_j}{n_{1j}} - \frac{c_j}{n_{2j}}$$

E sua variância é estimada por

$$\widehat{\text{var}}(\widehat{RAR}_j) = \frac{a_j b_j}{n_{1j}^3} + \frac{c_j d_j}{n_{2j}^3}$$

2.2.5. Número necessário para tratar

O número necessário para tratar é utilizado quando os grupos são tratamento e controle e pode ser interpretado como o número de indivíduos que devem ser tratados com o tratamento a fim de prevenir um evento.

O número necessário para tratar é dado pelo inverso da redução absoluta de risco

$$\widehat{NNT}_j = \frac{1}{\widehat{RAR}_j}$$

Assim, a variância também é estimada pelo inverso da variância da redução absoluta do risco

$$\widehat{\text{var}}(\widehat{NNT}_j) = \frac{1}{\widehat{\text{var}}(\widehat{RAR}_j)}$$

2.3. Medidas de efeito para comparação de dois grupos através de variáveis contínuas

Quando os dados de estudos são medidos em uma escala contínua, é comum aparecerem médias e desvios padrão nos relatórios 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.

2.3.1. 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. Note que ela só deve ser utilizada quando todos os estudos da metanálise utilizam a mesma escala contínua.

A diferença absoluta entre as médias de um estudo j é estimada por

$$\hat{D}_j = \bar{x}_{1j} - \bar{x}_{2j}$$

onde \bar{x}_{1j} é a média do grupo 1 e \bar{x}_{2j} é a média do grupo 2.

Quando se assume que as variâncias das duas populações são iguais, a variância da diferença do estudo j é estimada por

$$\widehat{\text{var}}(\hat{D}_j) = s_{p_j}^2 \left(\frac{1}{n_{1j}} + \frac{1}{n_{2j}} \right)$$

onde n_{1j} é o tamanho da amostra do grupo 1, n_{2j} é o tamanho da amostra do grupo 2 e $s_{p_j}^2$ é a variância conjunta dos dois grupos, dada por

$$s_{p_j}^2 = \frac{(n_{1_j} - 1)s_{1_j}^2 + (n_{2_j} - 1)s_{2_j}^2}{n_{1_j} + n_{2_j} - 2} \quad (2.5)$$

onde $s_{1_j}^2$ e $s_{2_j}^2$ são as variâncias dos grupos 1 e 2, respectivamente.

Quando não se assume que as variâncias das duas populações são iguais, então a variância da diferença do estudo j é estimada por

$$\widehat{\text{var}}(\hat{D}_j) = \frac{s_1^2}{n_{1_j}} + \frac{s_2^2}{n_{2_j}}$$

2.3.2. Diferença padronizada entre médias

A diferença padronizada entre médias pode ser utilizada quando os estudos vem de populações com dados contínuos, mas que usam diferentes escalas, por exemplo, diferentes testes psicológicos. Existem três maneiras diferentes de se estimar a diferença padronizada entre médias, são elas: d de Cohen, g ajustado de Hedges e Δ de Glass. Estes métodos diferem quanto ao uso do desvio padrão nos cálculos e quanto ao uso ou não de correção para evitar vício quando as amostras são pequenas.

d de Cohen

O d de Cohen é a diferença média de cada estudo dividida pelo desvio padrão, criando assim uma medida padrão que possa ser comparável entre os estudos. O d de Cohen é estimado por

$$d_j = \frac{\bar{x}_{1_j} - \bar{x}_{2_j}}{\sqrt{s_{p_j}^2}}$$

onde $s_{p_j}^2$ é a variância dada em (2.5).

A variância de d_j é estimada por

$$\widehat{\text{var}}(d_j) = \frac{n_{1_j} + n_{2_j}}{n_{1_j} n_{2_j}} + \frac{d_j^2}{2(n_{1_j} + n_{2_j} - 2)}$$

g ajustado de Hedges

O g ajustado de Hedges é similar ao d de Cohen, porém ele inclui um ajuste para corrigir o vício de amostras pequenas. O g é definido como

$$g_j = \frac{\bar{x}_{1j} - \bar{x}_{2j}}{\sqrt{s_{p_j}^2}} \left(1 - \frac{3}{4(n_{1j} + n_{2j}) - 9} \right)$$

A variância de g_j é estimada por

$$\widehat{\text{var}}(g_j) = \frac{n_{1j} + n_{2j}}{n_{1j} n_{2j}} + \frac{g_j^2}{2(n_{1j} + n_{2j} - 3,94)}$$

Δ de Glass

O Δ de Glass utiliza o desvio padrão do grupo controle para dividir a diferença de médias. O Δ de Glass é dado por

$$\Delta_j = \frac{\bar{x}_{1j} - \bar{x}_{2j}}{\sqrt{s_{2j}^2}}$$

A variância de Δ_j é estimada por

$$\widehat{\text{var}}(\Delta_j) = \frac{n_{1j} + n_{2j}}{n_{1j} n_{2j}} + \frac{\Delta_j^2}{2(n_{2j} - 1)}$$

3. Modelos de efeito fixo e modelos de efeitos aleatórios

Em metanálise são utilizados basicamente dois tipos de modelos, os modelos de efeito fixo e os modelos de efeitos aleatórios. Estes dois modelos são descritos nas Seções 3.1 e 3.2, respectivamente.

3.1. Modelo de efeito fixo

Em um modelo de efeito fixo assume-se que o efeito de interesse é o mesmo em todos os estudos e que as diferenças entre eles são devidas apenas à variabilidade interna de cada estudo. Seja J o número de estudos de uma metanálise e Y_j o efeito observado no estudo j com $j=1, 2, \dots, J$, o modelo de efeito fixo é dado por

$$Y_j = \theta_M + \varepsilon_j$$

onde ε_j é o erro aleatório do estudo j e θ_M é o efeito comum a todos os estudos chamado de medida metanalítica. Os erros aleatórios são supostos terem distribuição normal com média 0 e variância σ_j^2 conhecida ($\varepsilon_j \sim N(0, \sigma_j^2)$), é por este motivo que medidas como *odds ratio*, risco relativo, *odds* e incidência são calculados em escala logarítmica. Note que a variância conhecida σ_j^2 é a variância do efeito de interesse estimada no estudo j .

3.1.1. Estimação da medida metanalítica θ_M

O estimador de máxima verossimilhança de θ_M é dado por

$$\hat{\theta}_{MV} = \frac{\sum_{j=1}^J \frac{Y_j}{\sigma_j^2}}{\sum_{j=1}^J \frac{1}{\sigma_j^2}} = \frac{\sum_{j=1}^J W_j Y_j}{\sum_{j=1}^J W_j} \quad (3.1)$$

Note que $\hat{\theta}_{MV}$ é uma média ponderada dos efeitos observados (Y_1, Y_2, \dots, Y_j), sendo o peso (W_j) dado a cada estudo j igual ao inverso da variância do estudo j .

A variância da medida metanalítica $\hat{\theta}_{MV}$ é estimada por

$$\hat{\sigma}_{\theta_{MV}}^2 = \frac{1}{\sum_{j=1}^J W_j}$$

Um intervalo de confiança assintótico para θ_M é dado por

$$LI_{\theta_{MV}} = \hat{\theta}_{MV} - z_{\alpha/2} \times \sqrt{\hat{\sigma}_{\theta_{MV}}^2}$$

e

$$LS_{\theta_{MV}} = \hat{\theta}_{MV} + z_{\alpha/2} \times \sqrt{\hat{\sigma}_{\theta_{MV}}^2}$$

onde $z_{\alpha/2}$ é o valor que acumula a probabilidade $(1 - \alpha/2)$ na distribuição Normal padrão.

3.1.2. Método de Mantel-Haenszel

Uma maneira alternativa de calcular a medida metanalítica é através do método de Mantel-Haenszel. Este método foi originalmente proposto para combinar *odds ratios* oriundos

de diferentes estratos quando não são detectadas diferenças significativas entre eles. Por exemplo, combinar a razão entre o *odds* de infarto do miocárdio entre indivíduos com pressão arterial elevada e o *odds* de infarto do miocárdio entre indivíduos sem pressão arterial elevada de dois estratos: indivíduos com até 60 anos e indivíduos com mais de 60 anos. Nesta situação, se o efeito (medido em *odds ratio*) da pressão arterial elevada no infarto do miocárdio não foi significativamente diferente nos dois estratos de idade considerados, a combinação destes efeitos através do método de Mantel-Haenszel avalia o efeito do fator pressão arterial elevada no desfecho infarto do miocárdio controlado pela covariável idade.

No contexto de metanálise, o método de Mantel-Haenszel é utilizado para combinar medidas de efeito oriundas de diferentes estudos quando a suposição é de que o efeito é o mesmo em todos os 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. A medida metanalítica calculada através do método de Mantel-Haenszel é estimada por

$$\hat{\theta}_{MH} = \frac{\sum_{j=1}^J W_j Y_j}{\sum_{j=1}^J W_j} \quad (3.2)$$

Note que as expressões (3.1) e (3.2) são as mesmas, o que muda no Mantel-Haenszel é a maneira de calcular os pesos W_j e, conseqüentemente, a estimativa da variância de $\hat{\theta}_{MH}$. No método de Mantel-Haenszel, o cálculo do peso W_j depende da medida de efeito a ser combinada.

Mantel-Haenszel para combinar *odds ratios*: quando o interesse for combinar *odds ratios* utilizando o método de Mantel-Haenszel, o peso W_j de cada estudo é calculado da seguinte forma

$$W_j = \frac{b_j c_j}{n_{1j} + n_{2j}}$$

A variância da medida metanalítica $\hat{\theta}_{MH}$ é, então, estimada por

$$\hat{\sigma}_{\hat{\theta}_{MH}}^2 = \frac{1}{2} \left(\frac{E}{R^2} + \frac{F+G}{R+S} + \frac{H}{S^2} \right)$$

onde

$$R = \sum_{j=1}^J \frac{a_j d_j}{n_{1j} + n_{2j}}$$

$$S = \sum_{j=1}^J \frac{b_j c_j}{n_{1j} + n_{2j}}$$

$$E = \sum_{j=1}^J \frac{(a_j + d_j) a_j d_j}{(n_{1j} + n_{2j})^2}$$

$$F = \sum_{j=1}^J \frac{(a_j + d_j) b_j c_j}{(n_{1j} + n_{2j})^2}$$

$$G = \sum_{j=1}^J \frac{(b_j + c_j) a_j d_j}{(n_{1j} + n_{2j})^2}$$

$$H = \sum_{j=1}^J \frac{(b_j + c_j) b_j c_j}{(n_{1j} + n_{2j})^2}$$

Mantel-Haenszel para combinar riscos relativos: para combinar riscos relativos, o peso de cada estudo j é dado por

$$W_j = \frac{c_j n_{1j}}{n_{1j} + n_{2j}}$$

A variância da medida metanalítica $\hat{\theta}_{MH}$ é, então, estimada por

$$\hat{\sigma}_{\hat{\theta}_{MH}}^2 = \frac{P}{T \times U}$$

onde

$$P = \sum_{j=1}^J \frac{n_{1j} n_{2j} (a_j + c_j) - a_j c_j (n_{1j} + n_{2j})}{(n_{1j} + n_{2j})^2}$$

$$T = \sum_{j=1}^J \frac{a_j n_{2j}}{n_{1j} + n_{2j}}$$

$$U = \sum_{j=1}^J \frac{c_j n_{1j}}{n_{1j} + n_{2j}}$$

Mantel-Haenszel para combinar reduções absolutas de risco: para combinar reduções absolutas de risco, o peso de cada estudo j é dado por

$$W_j = \frac{n_{1j} n_{2j}}{n_{1j} + n_{2j}}$$

A variância da medida metanalítica $\hat{\theta}_{MH}$ é, então, estimada por

$$\hat{\sigma}_{\hat{\theta}_{MH}}^2 = \frac{L}{M^2}$$

onde

$$L = \sum_{j=1}^J \frac{a_j b_j n_{2j}^3 + c_j d_j n_{1j}^3}{n_{1j} n_{2j} (n_{1j} + n_{2j})^2}$$

$$M = \sum_{j=1}^J \frac{n_{1j} n_{2j}}{n_{1j} + n_{2j}}$$

3.2. Modelo de efeitos aleatórios

Quando se decide combinar estudos em uma metanálise é porque foi feita uma revisão sistemática e foram escolhidos estudos 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 um medicamento, entre eles pode haver diferença nos grupos de pessoas selecionadas, um grupo pode ser mais saudável em um estudo do que em outro, 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 isto acontece, estas diferenças devem ser investigadas e alternativas como metanálise em subgrupos e meta regressão devem ser consideradas. Outra alternativa é a utilização de um modelo de efeitos aleatórios. Este modelo 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 através de uma distribuição de probabilidade, geralmente suposta normal. O modelo de efeitos aleatórios é dado por

$$Y_j = \theta_M + \zeta_j + \varepsilon_j$$

onde ε_j é o erro aleatório do estudo j , ζ_j é o efeito aleatório de cada estudo j e θ_M é a medida metanalítica. No modelo de efeitos aleatórios existe a suposição de que os erros aleatórios têm distribuição normal com média 0 e variância conhecida σ_j^2 ($\varepsilon_j \sim N(0, \sigma_j^2)$) e os efeitos aleatórios tem distribuição normal com média 0 e variância desconhecida τ^2 ($\zeta_j \sim N(0, \tau^2)$). Como ζ_j e ε_j são independentes, tem-se que $Y_j \sim N(\theta_M, \tau^2 + \sigma_j^2)$. O parâmetro τ^2 representa a variabilidade entre os estudos e deve ser estimado.

3.2.1. Estimação da medida metanalítica θ_M^*

A notação usada para modelos de efeitos aleatórios será a mesma usada anteriormente para modelos de efeito fixo, para diferenciar um do outro um asterisco (*) indicará que é um modelo de efeitos aleatórios. Note que o modelo de efeitos aleatórios possui dois parâmetros desconhecidos: a medida metanalítica θ_M^* e a variância entre os estudos τ^2 .

O estimador de máxima verossimilhança de θ_M é dado por

$$\hat{\theta}_{MV}^* = \frac{\sum_{j=1}^J \frac{Y_j}{\sigma_j^2 + \hat{\tau}^2}}{\sum_{j=1}^J \frac{1}{\sigma_j^2 + \hat{\tau}^2}} = \frac{\sum_{j=1}^J W_j^* Y_j}{\sum_{j=1}^J W_j^*} \quad (3.3)$$

Nos modelos de efeito fixo a cada estudo é atribuído um peso que pode ser o inverso da variância. Em modelos de efeitos aleatórios o peso também pode ser o inverso da variância, mas nesse caso é somada a variância dentre os estudos σ_j^2 e a variância entre os estudos $\hat{\tau}^2$.

O peso para modelos de efeitos aleatórios é dado por

$$W_j^* = \frac{1}{\hat{V}_{Y_j}}$$

onde \hat{V}_{Y_j} é a variância estimada de Y_j e é dada por

$$\hat{V}_{Y_j} = \sigma_j^2 + \hat{\tau}^2$$

A variância da medida metanalítica é estimada por

$$\hat{\sigma}_{\hat{\theta}_{MV}}^{2*} = \frac{1}{\sum_{j=1}^J W_j^*}$$

Analisando a expressão (3.3) é possível notar que a estimativa de θ_{MV}^* depende da estimativa de τ^2 .

O estimador de máxima verossimilhança de τ^2 é dado por

$$\hat{\tau}^2 = \frac{\sum_{j=1}^J \frac{(Y_j - \hat{\theta}_{MV}^*)^2 - \sigma_j^2}{(\sigma_j^2 + \tau_j^2)^2}}{\sum_{j=1}^J \frac{1}{(\sigma_j^2 + \tau_j^2)^2}} \quad (3.4)$$

Como σ_j^2 pode ser superior a $(Y_j - \hat{\theta}_{MV}^*)^2$ para um dado j , consideraremos o estimador de máxima verossimilhança $\hat{\tau}_{MV}^2$ de τ^2 dado por

$$\hat{\tau}_{MV}^2 = \begin{cases} 0, & \text{se } \hat{\tau}^2 \leq 0 \\ \hat{\tau}^2, & \text{caso contrário} \end{cases}$$

Analisando a expressão (3.4) é possível notar que a estimativa de τ^2 depende de θ_{MV}^* , portanto estas estimativas só podem ser obtidas por métodos iterativos (Brockwell e Gordon, 2001).

3.2.2. Método de DerSimonian e Laird

Baseado na suposição de que os estudos incluídos na metanálise são homogêneos (apresentam a mesma medida de efeito) nem sempre é realística, DerSimonian e Laird propuseram estimar a medida metanalítica através do estimador pelo método dos momentos, ou seja,

$$\hat{\theta}_{DSL}^* = \frac{\sum_{j=1}^J \frac{Y_j}{\sigma_j^2 + \hat{\tau}^2}}{\sum_{j=1}^J \frac{1}{\sigma_j^2 + \hat{\tau}^2}} \quad (3.5)$$

É possível notar que o estimador de θ_M^* pelo método dos momentos (3.5) é igual ao estimador de máxima verossimilhança (3.3). Ambos dependem dos efeitos observados Y_j ,

suas variâncias conhecidas σ_j^2 e da variância entre os estudos τ^2 desconhecida. Na seção anterior foi mostrado como estimar τ^2 pelo método da máxima verossimilhança, uma alternativa para estimar τ^2 é utilizar o método dos momentos proposta por DerSimonian e Laird. Neste caso,

$$\hat{\tau}^2 = \frac{Q - gl}{C} \quad (3.6)$$

onde,

$$Q = \sum_{j=1}^J W_j Y_j^2 - \frac{\left(\sum_{j=1}^J W_j Y_j \right)^2}{\sum_{j=1}^J W_j}, \quad (3.7)$$

$$C = \sum_{j=1}^J W_j - \frac{\sum_{j=1}^J W_j^2}{\sum_{j=1}^J W_j},$$

$$W_j = \frac{1}{\sigma_j^2 + \hat{\tau}^2}, \quad (3.8)$$

$$gl = J - 1 \quad (3.9)$$

onde J é o número total de estudos.

Note que $\hat{\tau}^2$ na expressão (3.6) e W_j na expressão (3.8) dependem um do outro para serem estimados, portanto, estas estimativas só podem ser obtidas através de métodos iterativos.

4. Heterogeneidade

Em metanálises, quando a variabilidade entre os estudos não é apenas aleatória, dizemos que os estudos são heterogêneos. A avaliação sobre heterogeneidade tem papel importante na escolha do modelo de metanálise e, portanto, deve ser realizada antes da escolha do modelo.

As maneiras mais usuais 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 na sua verdadeira variação. No entanto, essa variação é constituída da verdadeira heterogeneidade e também do erro aleatório.

4.1. Separação da variação

Os efeitos observados variam por dois motivos. Um deles é a verdadeira heterogeneidade na medida de efeito, já o outro é o erro aleatório dentro do estudo. Para quantificar a heterogeneidade é necessário separar a variação nestes dois componentes.

Para separar estas variações é necessário calcular a variação total observada de estudo a estudo, 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.

Para obtermos a variação total, usamos a estatística Q, que é definida como,

$$Q = \sum_{j=1}^J W_j (Y_j - \hat{\theta}_M)^2, \quad (4.1)$$

onde W_j é o peso do estudo j , Y_j é a medida de efeito do estudo j e $\hat{\theta}_M$ é a estimativa para a medida metanalítica. A equação (4.1) é equivalente a (3.7).

O próximo passo é determinar o valor esperado de Q assumindo que os estudos são homogêneos. Neste caso ele é simplesmente os graus de liberdade (gl), definido em (3.9).

O excesso de variação é a diferença entre Q e os graus de liberdade,

$$Q - gl,$$

que é a parte da variação atribuída à verdadeira diferença dos efeitos de estudo a estudo.

4.2. Teste Q de Cochran

O teste Q de Cochran apresenta como hipótese nula a afirmação de que os estudos que compõem a metanálise são homogêneos. Sob esta hipótese nula, Q segue uma distribuição qui-quadrado com $J - 1$ graus de liberdade, onde J é o número total de estudos que compõem a metanálise. A estatística Q do teste de Cochran é o Q definido em (4.1), Cochran propôs esta estatística que posteriormente foi utilizada por DerSimonian e Laird para estimar τ^2 , conforme visto na seção 3.2.1.

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 (Higgins e Thompson, 2002).

4.3. Estatística I^2

A estatística I^2 foi proposta por Higgins e Thompson como uma forma mais direta de verificar a existência de heterogeneidade através de proporção.

A estatística I^2 é dada por

$$I^2 = \frac{Q - (J - 1)}{Q} \times 100\%$$

Note que a estatística I^2 pode variar de valores negativos até 100%. Quando o valor for negativo ele é igualado a 0. A significância de I^2 é equivalente a significância de Q.

Higgins et al. (2003) sugerem 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.

5. *Forest plot*

O *forest plot* 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, por exemplo, 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 o seu intervalo de confiança, sendo que quanto maior esta linha, maior é a variabilidade dentro do estudo.

Exemplos de *forest plots* serão apresentados nos Capítulos 7 e 8.

6. Escolha do modelo

A escolha de qual modelo é o mais adequado a ser utilizado não é trivial. Não existe uma regra de escolha, mas alguns critérios devem ser considerados.

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.

Já o modelo de efeitos aleatórios pode ser utilizado quando o pesquisador combina vários estudos que tem o mesmo objetivo, mas que não foram conduzidos da mesma maneira, por exemplo. Neste caso, é possível extrapolar para outras populações, o que torna a análise mais abrangente.

Embora o modelo de efeitos aleatórios possa parecer melhor por ser mais abrangente, devem ser tomados alguns cuidados. Se o número de estudos for muito pequeno, a estimativa da variância entre os estudos não terá uma boa precisão. Para resolver este problema existem várias opções, algumas delas são: relatar os efeitos separados e não como uma medida resumo, já que com um número pequeno de estudos não se deve fazer conclusões sobre a medida de efeito e seu intervalo de confiança; outra opção seria utilizar um modelo de efeito fixo, lembrando das suas limitações de não fazer inferências para outras populações; e uma terceira opção seria usar uma abordagem Bayesiana, onde a estimativa da variância entre os estudos é baseada em dados não só dos estudos utilizados, mas em dados externos.

Quanto à heterogeneidade, ela ajuda na escolha do modelo, mas não é determinante, maiores detalhes poderão ser vistos na Seção 7.3, cujo exemplo serve para mostrar que não devemos apenas olhar o teste de heterogeneidade e conduzir uma metanálise de olhos fechados.

7. Metanálise no STATA

O STATA é um software de estatística que possui vários métodos metanalíticos disponíveis. As análises no STATA são feitas através de sua linha de comando. Alguns destes comandos devem ser instalados.

Para metanálise os comandos disponíveis são *metan* que é o comando usado para fazer metanálises, *metareg* que é usado para fazer modelos de regressão em metanálises, *metacum*

que faz metanálises cumulativas e *metainf* que é usado para verificar a influência individual de um estudo na metanálise (Egger, 2001).

Nos exemplos que serão apresentados, ainda neste capítulo, será usado o comando *metan*. É possível obter maiores informações sobre os outros comandos citados no menu *Help* do STATA, clicando em *Help*, em seguida em *Contents* e na janela que abrirá, em *Command* digitar o nome do comando que se quer obter mais informações.

7.1. Comando *metan*

Para instalar o comando *metan*, abra o STATA clique no menu *Help*. Em seguida clique na opção *SJ and User-written Programs*, na janela que abrir clique em *Stata Journal* e procure o *sj8-1*, clique em *sbe24_2* e então clique no local indicado para instalar e o *metan* já estará disponível no STATA. Depois de instalado o comando *metan*, é possível obter informações sobre ele conforme descrito anteriormente.

Com o comando *metan* é possível fazer metanálises baseadas em modelos de efeito fixo ou modelos de efeitos aleatórios.

- Opções para modelos de efeito fixo:
 - *fixed*: é o default. Ele utiliza o método de Mantel-Haenszel para variáveis binárias e o método do inverso da variância para variáveis contínuas;
 - *fixedi*: utiliza o método do inverso da variância. Está disponível apenas para variáveis binárias;
 - *peto*: utiliza o método de Peto para calcular *odds ratios*. Este também está disponível apenas para variáveis binárias.
- Opções para modelos de efeitos aleatórios:
 - *random*: utiliza o método de DerSimonian e Laird com estimação de heterogeneidade pelo método de Mantel-Haenszel;
 - *randomi*: utiliza o método de DerSimonian e Laird com estimação de heterogeneidade pelo método do inverso da variância, este não está disponível para variáveis contínuas.
- Opções para medidas de efeito:
 - *rr*: risco relativo;

- *or*: odds ratio;
- *rd*: redução absoluta do risco;
- *cohen, hedges, glass*: são os três métodos diferentes de calcular a diferença padronizada entre médias;
- *nostandard*: diferença absoluta entre médias.

7.2. Exemplo: comparação de dois grupos com dados binários

A Tabela 7.1 apresenta dados de 22 ensaios clínicos randomizados sobre o uso de estreptoquinase intravenosa na prevenção de mortes por infarto do miocárdio. Os dados deste exemplo foram retirados de Egger, 2001.

Tabela 7.1 – Estreptoquinase intravenosa no infarto do miocárdio

Estudo	Ano da publicação	Grupo tratado		Grupo controle	
		Morte	Total	Morte	Total
Fletcher	1959	1	12	4	11
Dewar	1963	4	21	7	21
1st European	1969	20	83	15	84
Heikinheimo	1971	22	219	17	207
Italian	1971	19	164	18	157
2nd European	1971	69	373	94	357
2nd Frankfurt	1973	13	102	29	104
1st Australian	1973	26	264	32	253
NHLBI SMIT	1974	7	53	3	54
Valere	1975	11	49	9	42
Frank	1975	6	55	6	53
UK Collaborative	1976	48	302	52	293
Klein	1976	4	14	1	9
Austrian	1977	37	352	65	376
Lasierra	1977	1	13	3	11
N German	1977	63	249	51	234
Witchitz	1977	5	32	5	26
2nd Australian	1977	25	112	31	118
3rd European	1977	25	156	50	159
ISAM	1986	54	859	63	882
GISSI-1	1986	628	5860	758	5852
ISIS-2	1988	791	8592	1029	8595

Etapa 1: a primeira etapa para se realizar uma metanálise destes dados no STATA é digitar os dados no editor de dados. Para isto, vá no menu *Data* e clique em *Data editor*. Isto fará abrir uma planilha onde serão digitados os dados conforme Figura 7.1. Também podem ser importados diretamente arquivos em outros formatos como em ASCII ou Excel.

	estudo	anodapubli-ão	mortet	totalt	mortec	totalc
1	Fletcher	1959	1	12	4	11
2	Dewar	1963	4	21	7	21
3	1st European	1969	20	83	15	84
4	Heikinheimo	1971	22	219	17	207
5	Italian	1971	19	164	18	157
6	2nd European	1971	69	373	94	357
7	2nd Frankfurt	1973	13	102	29	104
8	1st Australian	1973	26	264	32	253
9	NHLBI SMIT	1974	7	53	3	54
10	Valere	1975	11	49	9	42
11	Frank	1975	6	55	6	53
12	UK Collaborative	1976	48	302	52	293
13	Klein	1976	4	14	1	9
14	Austrian	1977	37	352	65	376
15	Lasierra	1977	1	13	3	11
16	N German	1977	63	249	51	234
17	Witczitz	1977	5	32	5	26
18	2nd Australian	1977	25	112	31	118
19	3rd European	1977	25	156	50	159
20	ISAM	1986	54	859	63	882
21	GISSI-1	1986	628	5860	758	5852
22	ISIS-2	1988	791	8592	1029	8595

Figura 7.1 – Editor de dados do STATA

Na primeira coluna foram digitados os nomes dos estudos, na segunda o ano de publicação, as colunas *mortet* e *mortec* apresentam o número indivíduos que morreram no grupo tratamento e no grupo controle, respectivamente. Finalmente, as colunas *totalt* e *totalc* apresentam os totais de indivíduos de cada grupo.

Etapa 2: para análises de estudos que comparam dois grupos quanto a uma variável binária, o comando *metan* requer variáveis que expressem o número de indivíduos que respondeu e que não respondeu em cada um dos grupos. Como na Tabela 7.1 são expressos apenas o número de mortes e o número total de cada grupo, é necessário que sejam criadas variáveis com os indivíduos que sobreviveram em cada grupo, fazendo a diferença do total pelo número de mortes através do comando *generate* digitado na linha de comando, conforme segue:

generate sobreviveramt = totalt - mortet

generate sobreviveramc = totalc - mortec

Note que os nomes *sobreviveramt* e *sobreviveramc* foram escolhidos pelo analista podendo ser qualquer nome de sua escolha. Note, ainda, que as duas novas variáveis aparecem automaticamente na planilha de dados como mostra a Figura 7.2.

	estudo	anodapubli-o	mortet	totalt	mortec	totalc	sobreviver-t	sobreviver~c
1	Fletcher	1959	1	12	4	11	11	7
2	Dewar	1963	4	21	7	21	17	14
3	1st European	1969	20	83	15	84	63	69
4	Heikinheimo	1971	22	219	17	207	197	190
5	Italian	1971	19	164	18	157	145	139
6	2nd European	1971	69	373	94	357	304	263
7	2nd Frankfurt	1973	13	102	29	104	89	75
8	1st Australian	1973	26	264	32	253	238	221
9	NHLBI SMIT	1974	7	53	3	54	46	51
10	Valere	1975	11	49	9	42	38	33
11	Frank	1975	6	55	6	53	49	47
12	UK Collaborative	1976	48	302	52	293	254	241
13	Klein	1976	4	14	1	9	10	8
14	Austrian	1977	37	352	65	376	315	311
15	Lasierna	1977	1	13	3	11	12	8
16	N German	1977	63	249	51	234	186	183
17	Witchitz	1977	5	32	5	26	27	21
18	2nd Australian	1977	25	112	31	118	87	87
19	3rd European	1977	25	156	50	159	131	109
20	ISAM	1986	54	859	63	882	805	819
21	GISSI-1	1986	628	5860	758	5852	5232	5094
22	ISIS-2	1988	791	8592	1029	8595	7801	7566

Figura 7.2 – Editor de dados do STATA depois do uso do comando *generate*

Etapa 3: o primeiro passo da metanálise é definir a medida de efeito que será utilizada. Neste exemplo será utilizado o risco relativo, mas poderíamos escolher outra medida para dados binários dependendo dos delineamentos dos estudos envolvidos. Dada a medida de efeito escolhida, o segundo passo é verificar sua variabilidade entre e dentre os estudos. Para isto, indicamos o *forest plot* e o teste de heterogeneidade utilizando a estatística I^2 . Na linha de comando digitamos:

```
metan mortet sobreviveramt mortec sobreviveramc, rr label(namevar=estudo)
```

label(namevar=estudo) é uma opção para que o nome dos estudos apareça na análise e no gráfico. Esta opção não é uma obrigatoriedade, se ela não for usada, os estudos serão numerados para serem identificados.

Este comando gera o resultado da metanálise com modelo de efeito fixo pelo método de Mantel-Haenszel. Lembre que este é o modelo padrão do STATA e que se alguma outra opção de modelo, descrita na seção anterior, fosse desejável esta deveria estar no comando.

Este resultado é apresentado na Figura 7.4. O respectivo *forest plot* também é produzido conforme apresentado na Figura 7.3.

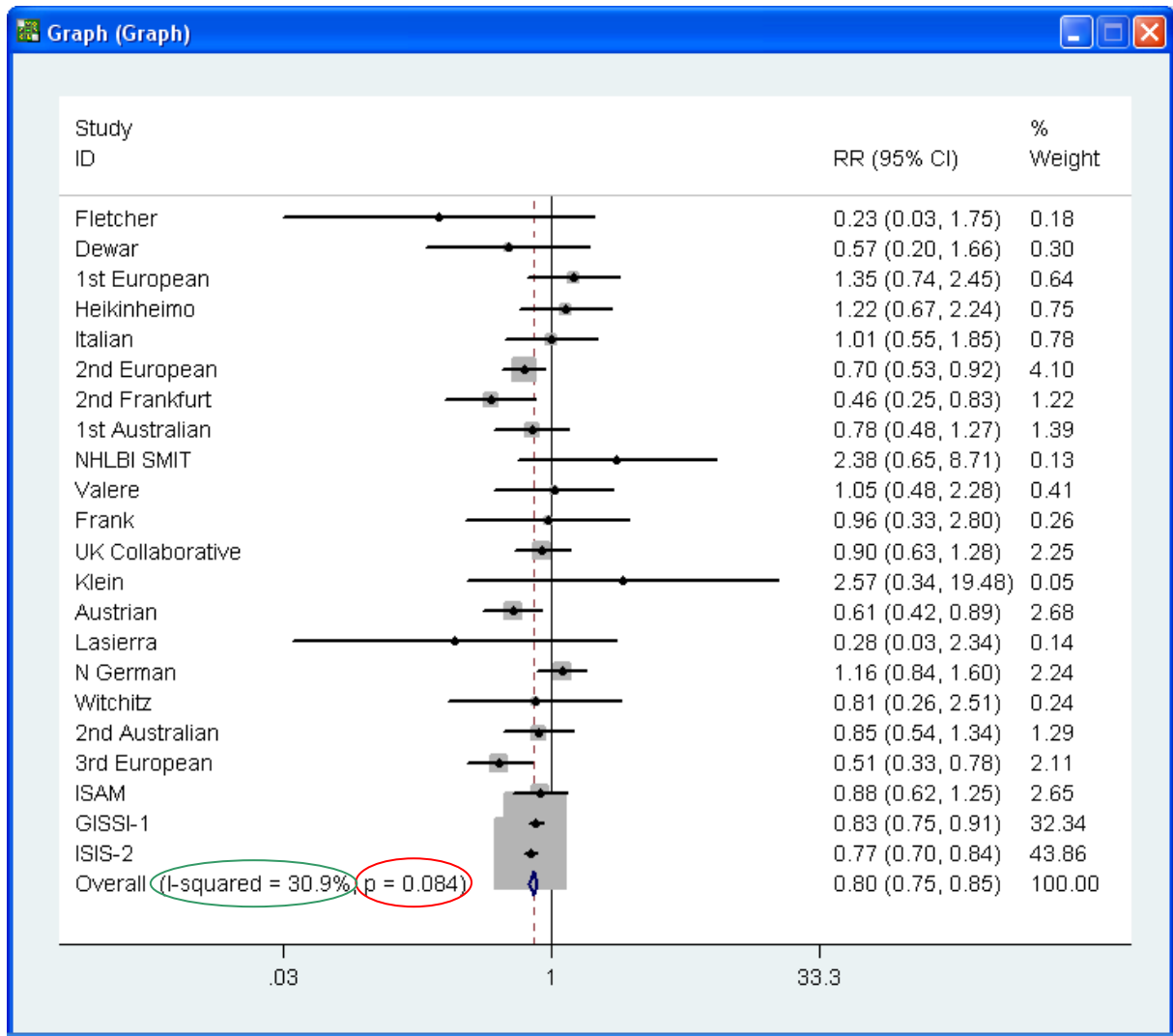


Figura 7.3 – *Forest plot* no STATA para estreptoquinase intravenosa no infarto do miocárdio

Analisando o gráfico e o teste Q que apresentou um valor $p = 0,084$ (circulado em vermelho), que aceita a hipótese nula de homogeneidade, isto é, que não se têm evidências para afirmar que os estudos são heterogêneos, podemos perceber que não existe heterogeneidade ou ela é muito baixa ($I^2 = 30,9\%$, circulado em verde), com isso podemos optar por uma metanálise com modelo de efeito fixo.

Etapa 4: a última etapa é a análise dos resultados do modelo escolhido. Este exemplo ilustra a situação em que o modelo escolhido é o mesmo modelo padrão utilizado pelo

STATA que é o modelo de efeito fixo pelo método de Mantel-Haenszel. O método de Mantel-Haenszel foi escolhido pois alguns estudos apresentam tamanhos de amostra muito pequenos.

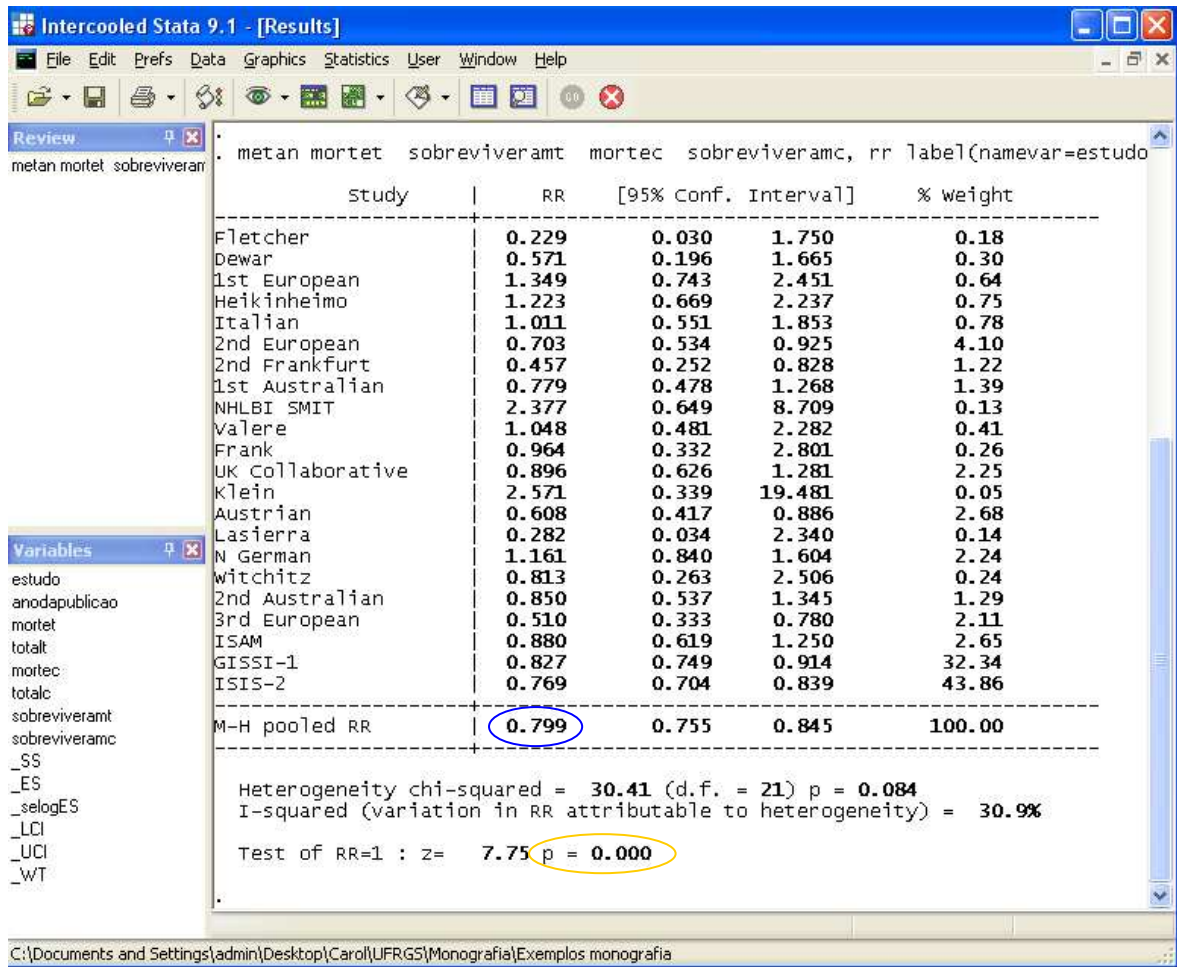


Figura 7.4 – Metanálise no STATA para estreptoquinase intravenosa no infarto do miocárdio

Assim como na Figura 7.3, a saída de resultados do STATA mostra o risco relativo, seu intervalo de confiança e o peso atribuído em percentual para cada estudo, assim como, a medida metanalítica (circulada em azul) e seu intervalo de confiança. O teste de heterogeneidade que aparece na Figura 7.3 também aparece na saída Figura 7.4. Na última linha da saída de resultados, aparece ainda o teste para o risco relativo, que tem como hipótese nula RR=1 (os riscos do grupo tratamento e controle são iguais), com um valor p=0,000 (circulado em amarelo), a hipótese nula foi rejeitada, o risco não é o mesmo para o grupo tratamento e o grupo controle.

Tanto a saída da análise quanto o gráfico mostram que existe o efeito da estreptoquinase na prevenção de morte por infarto do miocárdio. A metanálise é dominada pelos estudos GISSI-1 e ISIS-2, que contribuem com 76,2% do peso da análise.

Observações: note que no canto inferior esquerdo da Figura 7.4 estão as variáveis do editor de dados e que depois que rodamos o comando *metan*, seis novas variáveis foram criadas. São elas:

- *_SS*: tamanho da amostra do estudo;
- *_ES*: medida de efeito;
- *_selogES*: erro padrão dos logaritmos;
- *_LCI*: limite inferior do intervalo de confiança da medida de efeito;
- *_UCI*: limite superior do intervalo de confiança da medida de efeito;
- *_WT*: peso do estudo em percentual.

No editor de dados, essas variáveis aparecem em seis novas colunas, conforme Figura 7.5 a seguir.

	totalt	mortec	totalc	sobreviver~t	sobreviver~c	_SS	_ES	_selogES	_LCI	_UCI	_WT
1	12	4	11	11	7	23	.2291667	1.037187	.0300121	1.749875	.1779447
2	21	7	21	17	14	42	.5714286	.5455447	.1961523	1.664679	.2984281
3	83	15	84	63	69	167	1.349398	.3044893	.7429476	2.450878	.6356595
4	219	17	207	197	190	426	1.223207	.3080275	.6688157	2.237142	.7451696
5	164	18	157	145	139	321	1.010501	.3093867	.5510442	1.853051	.7841209
6	373	94	357	304	263	730	.7025555	.1401748	.5337816	.924693	4.095298
7	102	29	104	89	75	206	.4570656	.3032928	.2522407	.8282126	1.224342
8	264	32	253	238	221	517	.7786458	.2489399	.4780154	1.268347	1.393269
9	53	3	54	46	51	107	2.377358	.6624228	.6489919	8.708633	.1267025
10	49	9	42	38	33	91	1.047619	.3972436	.4809157	2.282116	.4132081
11	55	6	53	49	47	108	.9636363	.5443194	.3315795	2.80052	.2605325
12	302	52	293	254	241	595	.895568	.1825921	.6261464	1.280918	2.250427
13	14	1	9	10	8	23	2.571429	1.03318	.3394144	19.48133	.0519005
14	352	65	376	315	311	728	.6080419	.1921228	.4172519	.8860714	2.679763
15	13	3	11	12	8	24	.2820513	1.079584	.0339926	2.3403	.1385559
16	249	51	234	186	183	483	1.160879	.1648978	.840283	1.603792	2.241786
17	32	5	26	27	21	58	.8125	.5747073	.2634096	2.506196	.2352143
18	112	31	118	87	87	230	.8496544	.2342113	.536885	1.344632	1.287133
19	156	50	159	131	109	315	.5096154	.2174866	.3327496	.7804903	2.111328
20	859	63	882	805	819	1741	.8800932	.1791469	.619496	1.250313	2.650371
21	5860	758	5852	5232	5094	11712	.827365	.050696	.7491083	.9137969	32.33757
22	8592	1029	8595	7801	7566	17187	.7689759	.0447583	.704392	.8394814	43.86127

Figura 7.5 – Editor de dados do STATA depois do uso do comando *metan*

7.3. Exemplo: comparação de dois grupos com dados contínuos

Na Tabela 7.2 estão os dados de 9 ensaios clínicos randomizados que tem por objetivo comparar dois anestésicos diferentes usados em pequenas cirurgias quanto ao tempo de recuperação dos pacientes, no qual quanto mais rápida a recuperação, melhor é o anestésico. O tempo de recuperação é o tempo em minutos contado do momento em que os gases anestésicos são desligados até o momento em que o paciente abre os olhos. Para atender a suposição de normalidade dos dados, os tempos foram transformados para escala logarítmica. Os dados deste exemplo foram retirados de Whitehead, 2002.

O número total de pacientes submetidos a cada um dos tipos de anestésicos, a média do tempo de recuperação e o desvio padrão de cada grupo são apresentadas na Tabela 7.2.

Tabela 7.2 – Tempo (em escala logarítmica) de recuperação de pacientes submetidos aos anestésicos A e B

Estudo	Anestésico A			Anestésico B		
	Nº de pacientes	Média	Desvio Padrão	Nº de pacientes	Média	Desvio Padrão
1	4	1,141	0,967	5	0,277	0,620
2	10	2,165	0,269	10	1,519	0,913
3	17	1,790	0,795	17	1,518	0,849
4	8	2,105	0,387	9	1,189	1,061
5	7	1,324	0,470	10	0,456	0,619
6	11	2,369	0,401	10	1,550	0,558
7	10	1,074	0,670	12	0,265	0,502
8	5	2,583	0,409	4	1,370	0,934
9	14	1,844	0,848	19	2,118	0,749

Etapa 1: os dados foram digitados no STATA conforme Figura 7.6. Na coluna *estudo* estão os números dos estudos, n_a e n_b são o número total de pacientes submetidos ao uso do anestésico A e ao anestésico B, respectivamente, M_A e M_B são as médias do tempo de recuperação (em escala logarítmica) dos pacientes submetidos ao uso dos anestésicos A e B e DP_A e DP_B são os desvios padrão dos tempos de recuperação dos pacientes que usaram os anestésicos A e B.

	estudo	n_a	M_A	DP_A	n_b	M_B	DP_B
1	1	4	1.141	.967	5	.277	.62
2	2	10	2.165	.269	10	1.519	.913
3	3	17	1.79	.795	17	1.518	.849
4	4	8	2.105	.387	9	1.189	1.061
5	5	7	1.324	.47	10	.456	.619
6	6	11	2.369	.401	10	1.55	.558
7	7	10	1.074	.67	12	.265	.502
8	8	5	2.583	.409	4	1.37	.934
9	9	14	1.844	.848	19	2.118	.749

Figura 7.6 – Editor de dados do STATA

Etapa 2: o primeiro passo da metanálise é definir a medida de efeito que será usada, neste exemplo usaremos uma medida de efeito para dados contínuos e a medida usada será a diferença absoluta entre médias, pois todos os estudos mediram o tempo de recuperação dos pacientes na mesma escala.

O segundo passo é verificar a variabilidade entre e dentro dos estudos, para isto, devemos olhar o *forest plot* e o teste de heterogeneidade.

Na linha de comando digitamos:

```
metan n_a M_A DP_A n_b M_B DP_B, nostandard
```

Lembrando que *nostandard* é a opção para o uso da diferença absoluta de médias.

O *forest plot* gerado por esse comando é de uma metanálise com modelo de efeito fixo pelo método do inverso da variância, já que não foi definido nenhum outro e este é o padrão do STATA para dados contínuos.

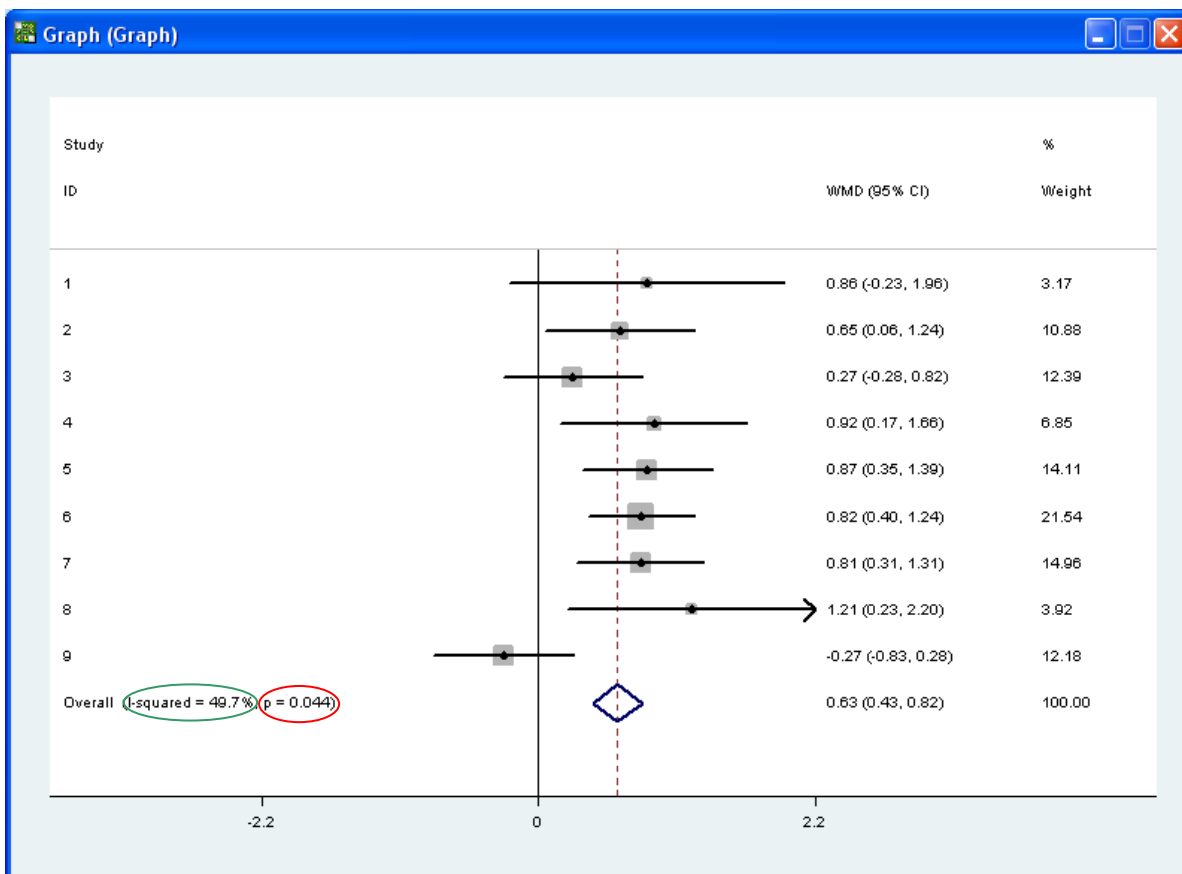


Figura 7.7 – *Forest plot* no STATA para diferença absoluta das médias dos anestésicos A e B com modelo de efeito

Analisando o gráfico e o teste Q que apresentou um valor $p = 0,044$ (circulado em vermelho), que rejeita a hipótese nula de homogeneidade, o que significa que temos evidências para afirmar que os estudos são heterogêneos, podemos verificar que a heterogeneidade pode ser considerada moderada através da estatística $I^2 = 49,7\%$ (circulada em verde), com isso podemos optar por uma metanálise com modelo de efeito aleatório ou indo além na observação do gráfico, podemos verificar que apenas o estudo 9 está bem diferente dos demais. O pesquisador poderia desconfiar de que há alguma falha no estudo 9 ou que ele tem características distintas dos demais e que isso passou despercebido pela revisão sistemática, o pesquisador deve investigar melhor e se descobrir alguma divergência, pode retirar este estudo da análise ou se não encontrar nada de anormal ou muito diferente, o estudo deve permanecer na metanálise. Vale salientar que não encontrando evidências de que o estudo tem características distintas dos demais, ele não pode ser retirado da metanálise só porque apresentou resultados diferentes.

Apenas para verificar o que aconteceria se retirássemos o estudo 9 da nossa metanálise, geramos o *forest plot* sem ele, conforme Figura 7.8 a seguir.

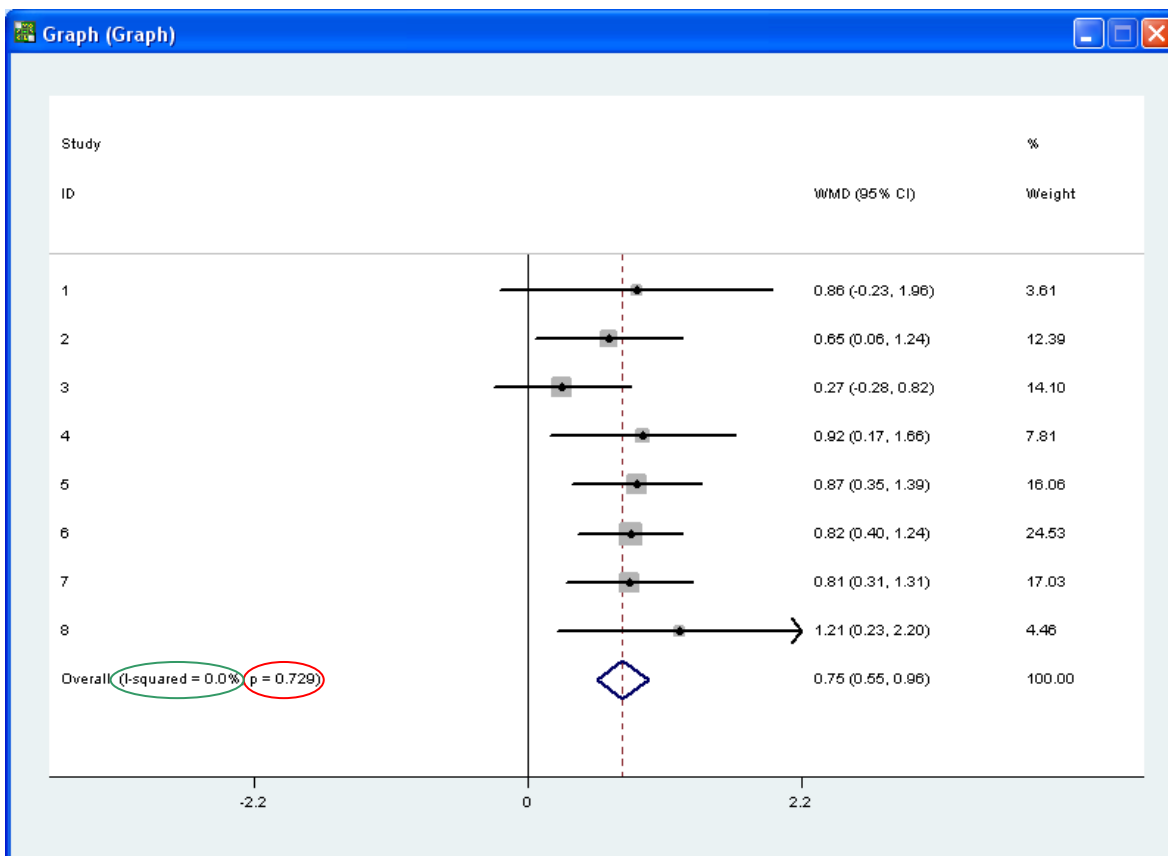


Figura 7.8 – *Forest plot* no STATA para diferença absoluta das médias dos anestésicos A e B com modelo de efeito fixo sem o estudo 9

Podemos perceber com este novo gráfico e com a estatística I^2 que agora os estudos são homogêneos e que o estudo 9 era o que causava a heterogeneidade.

A partir de agora devemos escolher, conduzir uma metanálise com modelo de efeito fixo, pois o peso do estudo 9 é baixo e poderíamos pensar que ele não interfere muito nos dados da metanálise, ou conduzir a metanálise com um modelo de efeitos aleatórios com o estudo 9. Ou ainda, se fosse comprovado que o estudo 9 apresenta características distintas dos demais, ele poderia ser retirado e a metanálise conduzida com modelo de efeito fixo, já que os demais estudos são homogêneos.

Neste exemplo vamos conduzir a metanálise sem tirar o estudo 9, pois não temos conhecimento suficiente de que ele difere dos demais, portanto não podemos retirá-lo e utilizar um modelo de efeitos aleatórios, pois existe uma heterogeneidade moderada e achamos conveniente usar este modelo. Que fique claro para o leitor que outra escolha poderia ter sido feita, é uma escolha que cabe ao especialista da área.

Etapa 3: neste exemplo o modelo escolhido para a metanálise com modelo de efeitos aleatórios, usamos o comando *metan* da seguinte maneira:

```
metan n_a M_A DP_A n_b M_B DP_B, nostandard random
```

Um novo *forest plot* é gerado conforme Figura 7.9, a diferença dele para o primeiro da Figura 7.7, é que neste já estamos conduzindo a metanálise com modelo de efeitos aleatórios, o que faz com que os pesos sejam calculados de maneira diferente.

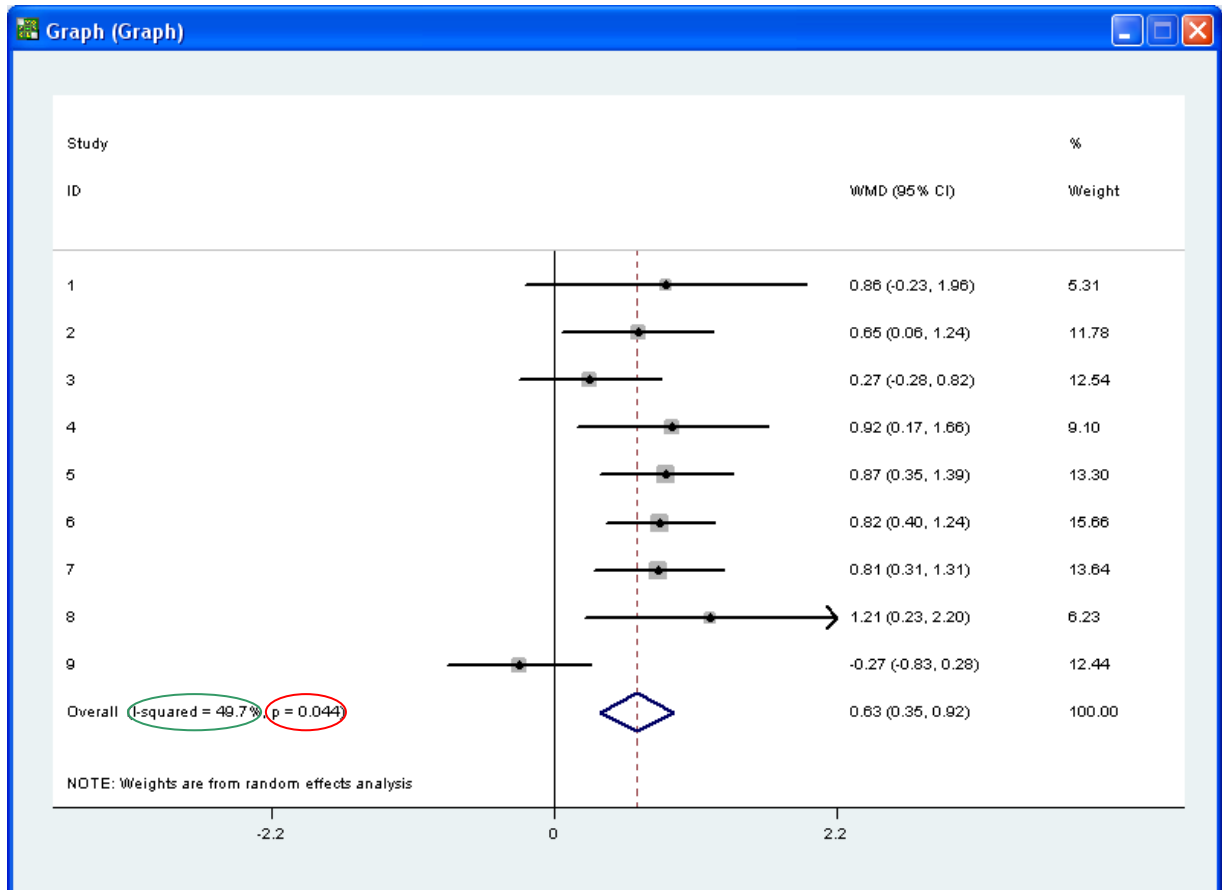


Figura 7.9 – *Forest plot* no STATA para diferença absoluta das médias dos anestésicos A e B com modelo de efeitos aleatórios

A saída de análises do STATA está na Figura 7.10, a seguir:

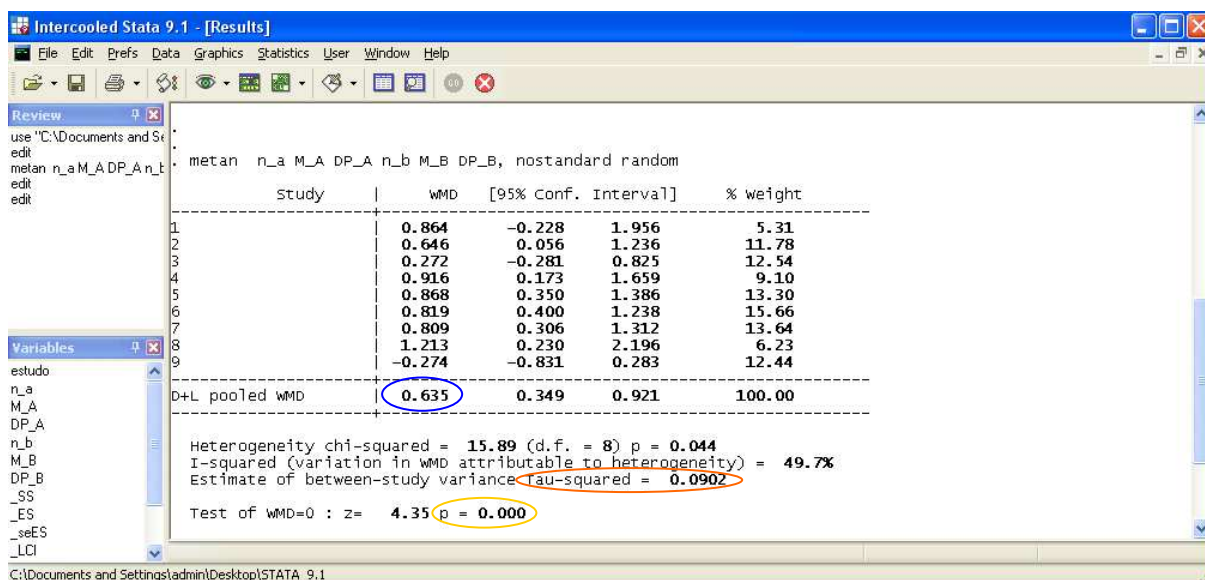


Figura 7.10 – Metanálise no STATA para diferença das médias dos anestésicos A e B

A saída do STATA mostra a diferença absoluta entre médias (WMD), seu intervalo de confiança e o peso atribuído em percentual para cada estudo, assim como, a medida metanalítica (circulada em azul) e seu intervalo de confiança. O teste de heterogeneidade Q e a estatística I^2 também aparecem, e já foram analisados no *forest plot*. Note que uma estimativa para o τ^2 aparece (circulada em laranja), pois este é um modelo de efeitos aleatórios e ele possui variabilidade entre os estudos. Na última linha, aparece ainda o teste para a diferença absoluta entre médias, que tem como hipótese nula $WMD=0$ (as médias dos dois tratamentos são iguais), com um $p=0,000$ (circulado em amarelo), a hipótese nula foi rejeitada, a média do tempo de recuperação dos dois anestésicos não é igual.

Tanto a saída da análise quanto o gráfico mostram que o tempo de recuperação dos pacientes submetidos ao anestésico A é maior que o tempo de recuperação dos pacientes submetidos ao anestésico B.

7.4. Exemplo: um grupo com dados binários

A Tabela 7.3 apresenta dados de 7 estudos sobre a eficácia no uso de um determinado diurético para controle da pressão arterial. Em cada estudo todos os pacientes apresentavam pressão alta, foram medicados com o diurético e sua pressão medida após uma hora. No grupo evento foram contados os casos que apresentaram pressão normal após o uso do diurético, ou seja, o número de pacientes nos quais o uso do diurético foi eficaz e no grupo não evento são

os pacientes que permaneceram com pressão alta após o uso do diurético. Foi considerada alta, a pressão maior que 140/90 mmHg.

Tabela 7.3 – Diminuição da pressão arterial com uso de um diurético

Estudo	Evento	Não evento
1	231	25
2	265	33
3	62	9
4	114	11
5	157	27
6	106	16
7	163	15

Etapa 1: os dados foram digitados no STATA conforme figura 7.11.



Figura 7.11 – Editor de dados do STATA

Na primeira coluna foram digitados os números dos estudos, na segunda a variável *evento* que é número de pacientes que apresentou pressão normal após o uso do diurético e *naoevento* é a variável que representa o número de pacientes que permaneceu com pressão alta após o uso do diurético.

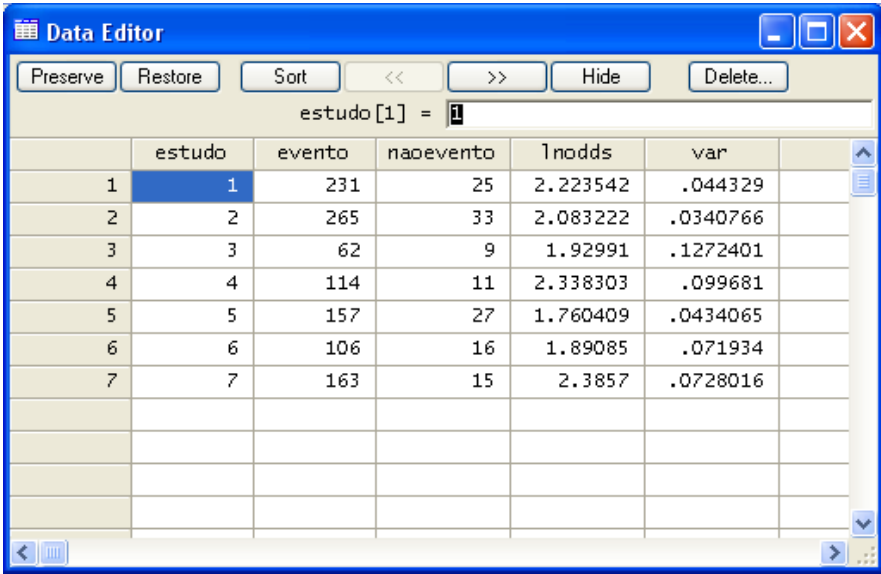
Etapa 2: para análises de estudos para um grupo com dados binários, o comando *metan* requer as medidas de efeito de cada estudo e suas respectivas variâncias. Neste exemplo vamos usar *odds* como medida de efeito, portanto, vamos gerar no STATA essas

variáveis através do comando *generate*. Para gerar o logaritmo do *odds* usaremos a expressão (2.1) e para sua variância a expressão (2.2).

```
generate lnodds = ln(evento/naoevento)
```

```
generate var = 1/evento+1/naoevento
```

As duas novas variáveis aparecem automaticamente na planilha de dados como mostra a Figura 7.12.



The screenshot shows the STATA Data Editor window with a dataset containing 7 observations. The columns are 'estudo', 'evento', 'naoevento', 'lnodds', and 'var'. The data is as follows:

	estudo	evento	naoevento	lnodds	var
1	1	231	25	2.223542	.044329
2	2	265	33	2.083222	.0340766
3	3	62	9	1.92991	.1272401
4	4	114	11	2.338303	.099681
5	5	157	27	1.760409	.0434065
6	6	106	16	1.89085	.071934
7	7	163	15	2.3857	.0728016

Figura 7.12 – Editor de dados do STATA depois do uso do comando *generate*

Etapa 3: podemos agora gerar o *forest plot* para verificar a presença ou não de heterogeneidade entre os estudos. Quando os dados são de um grupo o comando *metan* é utilizado da seguinte maneira:

```
metan lnodds var, eform
```

No comando *metan* primeiro é colocada a medida de efeito, seguida da sua variância. A opção *eform* é utilizada para os dados saírem na escala original e não em logaritmo, para as medidas que não utilizam a escala logarítmica, esta opção não é necessária. O *forest plot* gerado é apresentado na Figura 7.13.

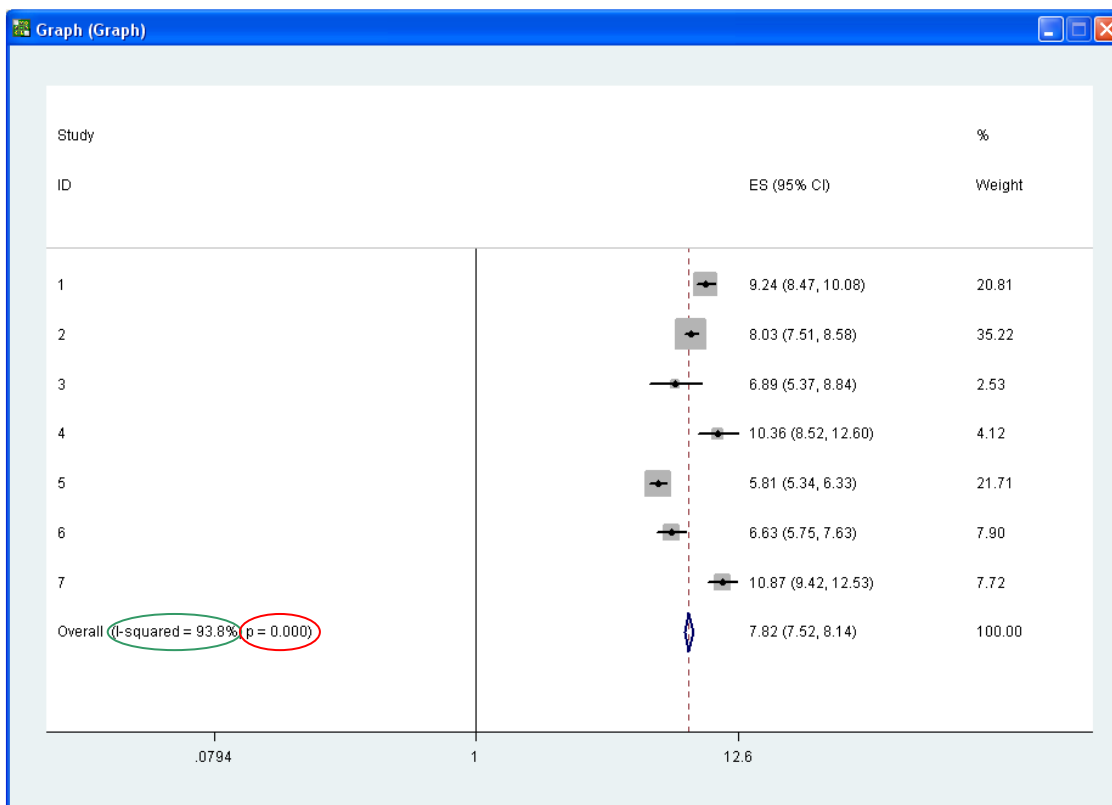


Figura 7.13 – *Forest plot* no STATA para pressão arterial com modelo de efeito fixo

Através da Figura 7.13 é possível analisar o *forest plot* e verificar que o teste Q apresentou um valor $p = 0,000$ (circulado em vermelho), que rejeita a hipótese nula de homogeneidade, o que significa que se tem evidências de que os estudos são heterogêneos. Através da estatística $I^2 = 93,8\%$ (circulada em verde), pode-se perceber que existe uma alta heterogeneidade entre os estudos. Com isso podemos optar por uma metanálise com modelo de efeitos aleatórios.

Etapa 4: para gerar uma metanálise com modelo de efeitos aleatórios para um grupo com variável binária, o comando é o seguinte:

metan lnodds var, randomi eform

Um novo *forest plot* é gerado com o modelo de efeitos aleatórios, conforme Figura 7.14.

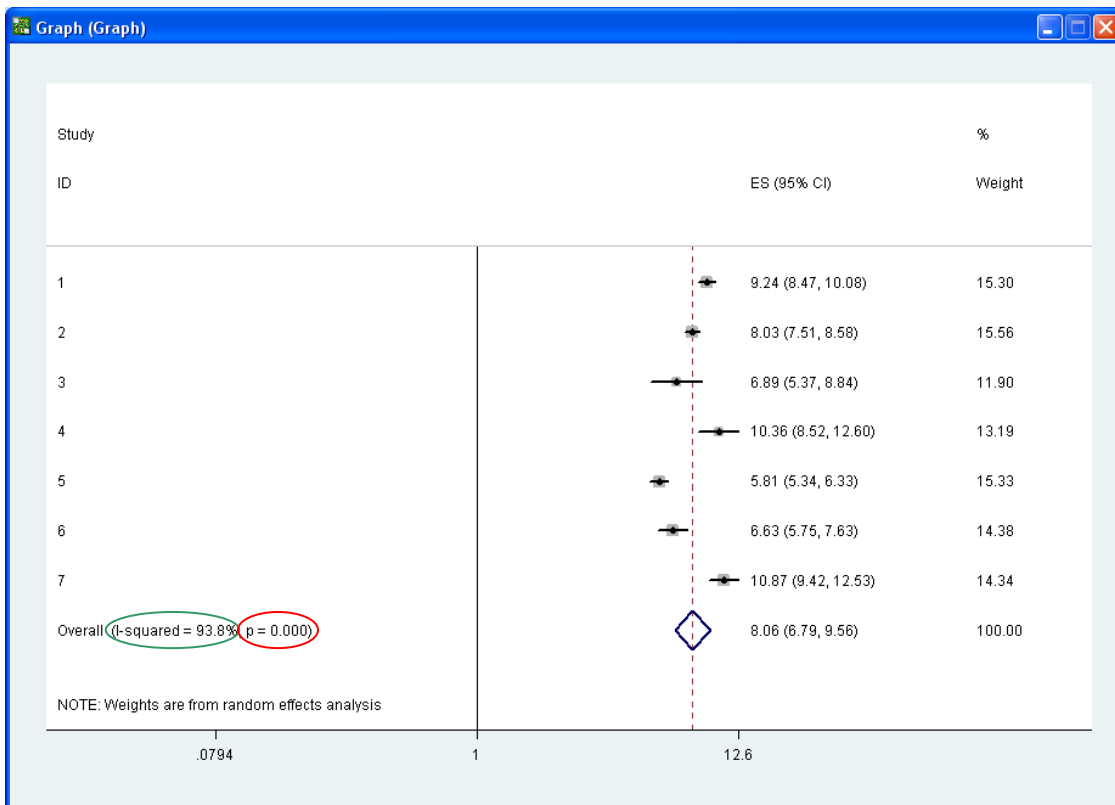


Figura 7.14 – Forest plot no STATA para pressão arterial com modelo de efeitos aleatórios

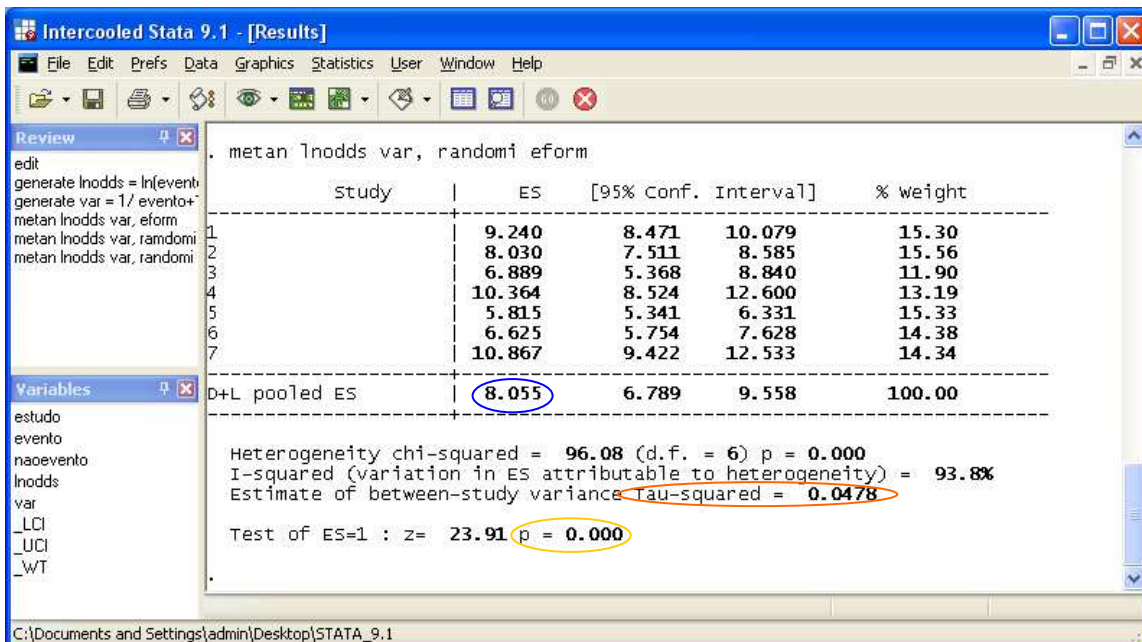


Figura 7.15 – Metanálise no STATA para pressão arterial

A saída do STATA mostra a medida de efeito, neste caso o *odds*, de cada estudo, seu intervalo de confiança e o peso atribuído em percentual para cada estudo, assim como, a medida metanalítica (circulada em azul) e seu intervalo de confiança. O teste de heterogeneidade Q e a estatística I² também aparecem, e já foram analisados no *forest plot*. A estimativa para o τ^2 aparece (circulada em laranja), pois este é um modelo de efeitos aleatórios e ele possui variabilidade entre os estudos. Na última linha da saída de resultados, aparece ainda o teste para a medida de efeito, neste caso o *odds*, que tem como hipótese nula *odds*=1 (a probabilidade de ocorrer o evento e não ocorrer é a mesma). Com um p=0,000 (circulado em amarelo), a hipótese nula foi rejeitada, ou seja, a probabilidade de ocorrer o evento é significativamente diferente da probabilidade de não ocorrer o evento. Baseados nos 7 estudos da metanálise estima-se que o *odds* de baixar a pressão seja 8,055, para pacientes que fazem o tratamento 8,055 a cada 1 paciente tem o efeito desejado (baixar a pressão).

8. Metanálise no R

O R é um software estatístico programável, ele é livre e pode ser obtido diretamente do site www.r-project.org. O R possui alguns pacotes básicos que já vem instalados juntamente com o programa e outros que são colaborações que o software recebe. Estes devem ser instalados separadamente, como é o caso do pacote que será utilizado neste trabalho chamado *metafor* que é um dos pacotes disponíveis atualmente para metanálises. Existem outros pacotes para metanálise no R que podem ser baixados como *meta*, *metacor* e *metaMA*.

Para instalar um pacote que não vem previamente instalado no R, no menu *Pacotes* clique em *Instalar pacote(s)* selecione o *CRAN mirror* que é o local de onde será baixado o pacote, na janela *Packages* que irá abrir, procure o pacote *metafor*, ele será instalado. Para começar a usar o *metafor*, ou qualquer outro pacote que for instalado separadamente, é necessário carregá-lo, no menu *Pacotes* clique na opção *Carregar pacote* procure o pacote *metafor*, então ele será carregado e estará pronto para ser usado.

8.1. Pacote *metafor*

O pacote *metafor* possui funções para conduzir metanálises no R. Estão disponíveis modelos de efeito fixo e efeitos aleatórios. Os métodos de Mantel-Haenszel e de Peto estão disponíveis para modelos de efeito fixo.

Alguns gráficos como forest, funnel e radial plots também podem ser feitos com o pacote *metafor*. Também é possível fazer modelos de meta-regressão com este pacote.

Os exemplos do capítulo anterior serão feitos novamente neste capítulo apenas para mostrar como conduzir uma metanálise no R, por isso serão mostrados apenas os comandos e as saídas do R, pois as conclusões são as mesmas, o exemplo da Seção 7.4 não será feito com o R, pois com o comando *metafor* não é possível conduzir metanálises para um grupo com dados binários.

8.2. Exemplo: comparação de dois grupos com dados binários

Este exemplo se refere ao mesmo exemplo da Seção 7.2, os dados utilizados serão os mesmos e não serão repetidos nesta seção.

Os dados da Tabela 7.1 são introduzidos no R da seguinte forma:

```
a=c(1,4,20,22,19,69,13,26,7,11,6,48,4,37,1,63,5,25,25,54,628,791)
```

```
na=c(12,21,83,219,164,373,102,264,53,49,55,302,14,352,13,249,32,112,156,859,  
5860,8592)
```

```
c=c(4,7,15,17,18,94,29,32,3,9,6,52,1,65,3,51,5,31,50,63,758,1029)
```

```
nc=c(11,21,84,207,157,357,104,253,54,42,53,293,9,376,11,234,26,118,159,882,5852,  
8595)
```

```
b=na-a
```

```
d=nc-c
```

onde a e c são os indivíduos que morreram no grupo tratamento e controle respectivamente, na e nc são todos os indivíduos do grupo tratamento e controle, b e d são os indivíduos que sobreviveram no grupo tratamento e controle.

Os dados também podem ser importados de outros arquivos em formato txt. Para importar dados de Excel, por exemplo, o arquivo deve ser salvo em formato “Texto (separado por tabulações)” e no R deve ser digitado o seguinte comando

```
dados = read.table("C:/exemplo/dados.txt")
```

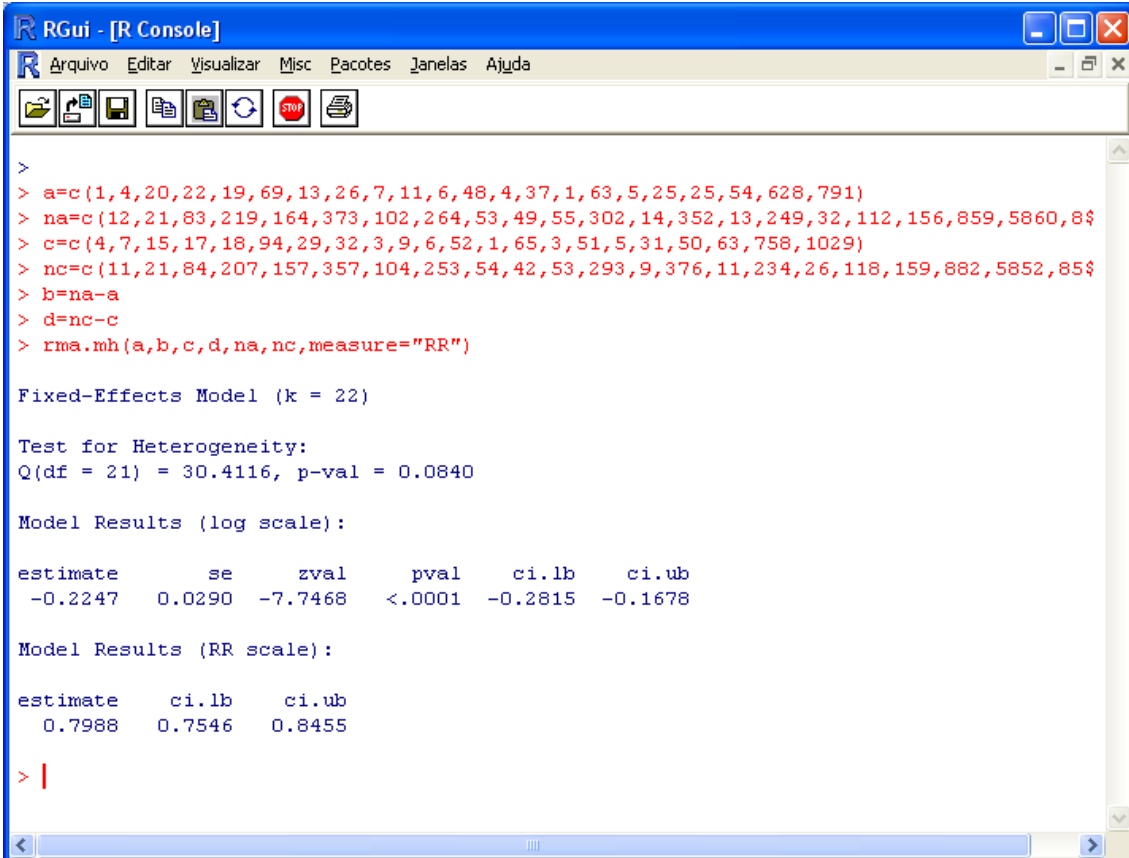
O programa irá ler o arquivo *dados* do caminho especificado, neste caso o arquivo estava no diretório C, na pasta exemplo.

Para conduzir uma metanálise com modelo de efeito fixo com o método de Mantel-Haenszel conforme feito no STATA anteriormente, o comando no R é o seguinte:

```
rma.mh(a,b,c,d,na,nc,measure="RR")
```


A medida que está sendo usada também deve ser indicada, neste caso RR (risco relativo).

A saída do R é a seguinte:



```
>
> a=c(1,4,20,22,19,69,13,26,7,11,6,48,4,37,1,63,5,25,25,54,628,791)
> na=c(12,21,83,219,164,373,102,264,53,49,55,302,14,352,13,249,32,112,156,859,5860,85)
> c=c(4,7,15,17,18,94,29,32,3,9,6,52,1,65,3,51,5,31,50,63,758,1029)
> nc=c(11,21,84,207,157,357,104,253,54,42,53,293,9,376,11,234,26,118,159,882,5852,855)
> b=na-a
> d=nc-c
> rma.mh(a,b,c,d,na,nc,measure="RR")

Fixed-Effects Model (k = 22)

Test for Heterogeneity:
Q(df = 21) = 30.4116, p-val = 0.0840

Model Results (log scale):

estimate      se      zval      pval      ci.lb      ci.ub
-0.2247    0.0290   -7.7468    <.0001   -0.2815   -0.1678

Model Results (RR scale):

estimate      ci.lb      ci.ub
0.7988    0.7546    0.8455

> |
```

Figura 8.1 – Metanálise com modelo de efeito fixo no R

O *forest plot* no R é gerado pelo seguinte comando:

```
forest(rma.mh(a,b,c,d,na,nc,measure="RR"))
```

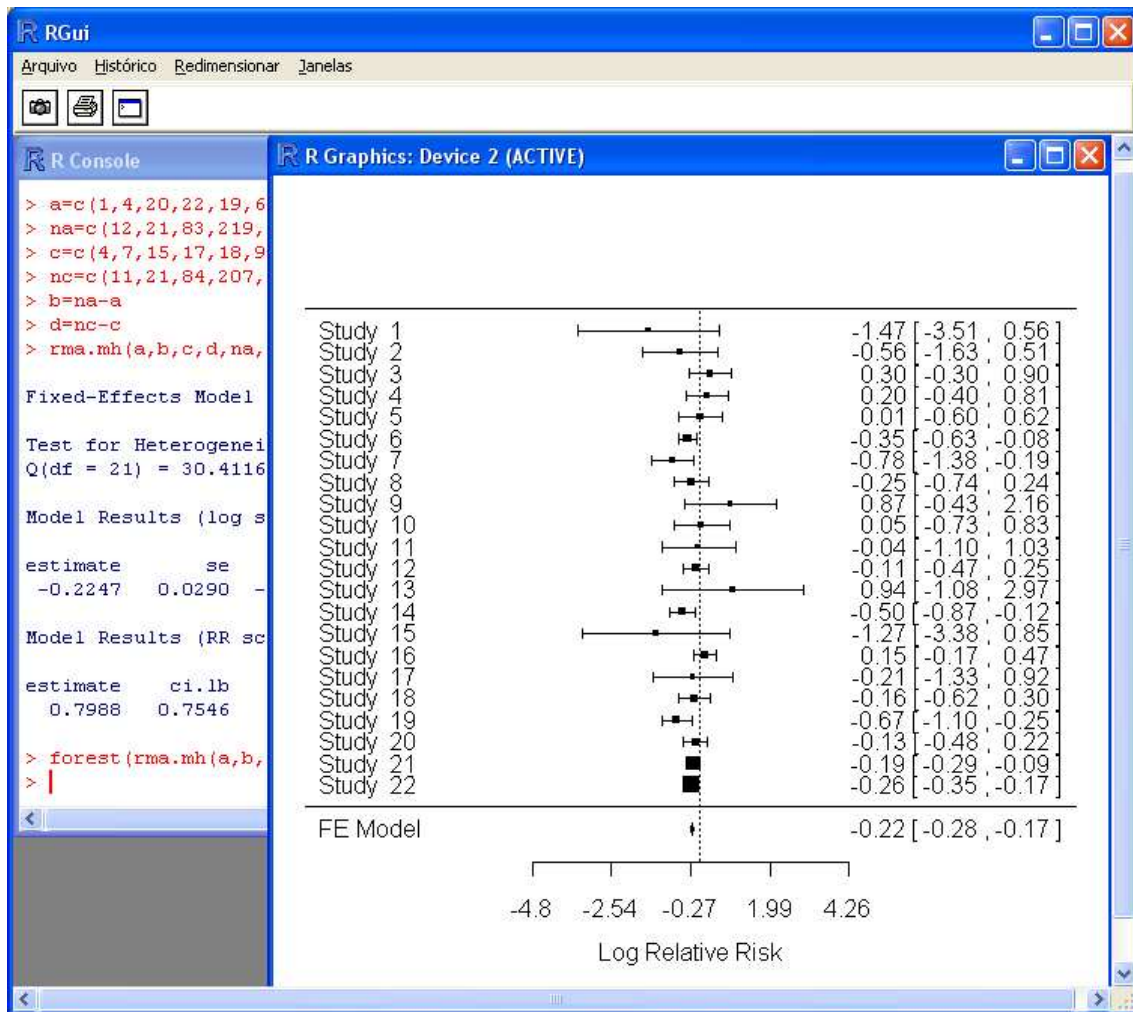


Figura 8.2 – Forest plot no R

Note que as medidas de efeito e os intervalos de confiança não são os mesmos dados pelo STATA na Figura 7.3, pois para o risco relativo, o R apresenta os dados em escala logarítmica e o STATA na escala original.

8.3. Exemplo: comparação de dois grupos com contínuos

Este exemplo se refere ao mesmo exemplo da Seção 7.3, os dados utilizados serão os mesmos e não serão repetidos nesta seção.

Os dados da Tabela 7.2 são introduzidos no R da seguinte forma:

$nA=c(4,10,17,8,7,11,10,5,14)$

$mA=c(1.141,2.165,1.790,2.105,1.324,2.369,1.074,2.583,1.844)$

$dpA=c(0.967,0.269,0.795,0.387,0.470,0.401,0.670,0.409,0.848)$

$nB=c(5,10,17,9,10,10,12,4,19)$

$mB=c(0.277,1.519,1.518,1.189,0.456,1.550,0.265,1.370,2.118)$

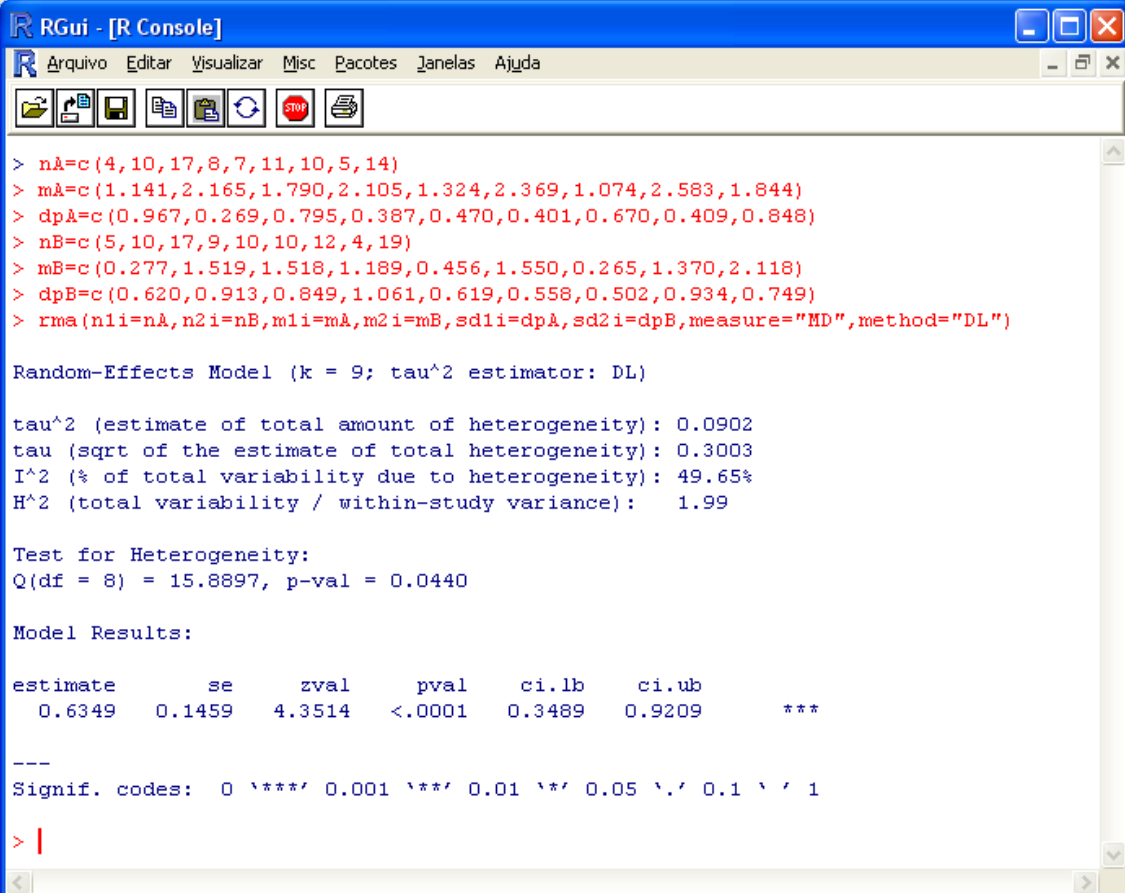
$dpB=c(0.620,0.913,0.849,1.061,0.619,0.558,0.502,0.934,0.749)$

onde nA e nB são o número total de pacientes submetidos ao uso do anestésico A e ao anestésico B, respectivamente, mA e mB são as médias do tempo de recuperação (em escala logarítmica) dos pacientes submetidos ao uso dos anestésicos A e B e dpA e dpB são os desvios padrão dos tempos de recuperação dos pacientes que usaram os anestésicos A e B.

Para conduzir uma metanálise com modelo de efeitos aleatórios conforme feito no STATA anteriormente, pelo método de DerSimonian e Laird, como medida de efeito a diferença absoluta entre médias, o comando no R é o seguinte:

```
rma(n1i=nA,n2i=nB,m1i=mA,m2i=mB,sd1i=dpA,sd2i=dpB,measure="MD",method="DL")
```

A saída do R é a seguinte:



```
RGui - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda
[Icons]
> nA=c(4,10,17,8,7,11,10,5,14)
> mA=c(1.141,2.165,1.790,2.105,1.324,2.369,1.074,2.583,1.844)
> dpA=c(0.967,0.269,0.795,0.387,0.470,0.401,0.670,0.409,0.848)
> nB=c(5,10,17,9,10,10,12,4,19)
> mB=c(0.277,1.519,1.518,1.189,0.456,1.550,0.265,1.370,2.118)
> dpB=c(0.620,0.913,0.849,1.061,0.619,0.558,0.502,0.934,0.749)
> rma(n1i=nA,n2i=nB,m1i=mA,m2i=mB,sd1i=dpA,sd2i=dpB,measure="MD",method="DL")

Random-Effects Model (k = 9; tau^2 estimator: DL)

tau^2 (estimate of total amount of heterogeneity): 0.0902
tau (sqrt of the estimate of total heterogeneity): 0.3003
I^2 (% of total variability due to heterogeneity): 49.65%
H^2 (total variability / within-study variance): 1.99

Test for Heterogeneity:
Q(df = 8) = 15.8897, p-val = 0.0440

Model Results:

estimate      se      zval      pval      ci.lb      ci.ub      ***
  0.6349     0.1459     4.3514    <.0001     0.3489     0.9209

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> |
```

Figura 8.3 – Metanálise com modelo de efeitos aleatórios no R

O *forest plot* é feito com o comando:

```
forest(rma(n1i=nA,n2i=nB,m1i=mA,m2i=mB,sd1i=dpA,sd2i=dpB,measure="MD",  
method="DL"))
```

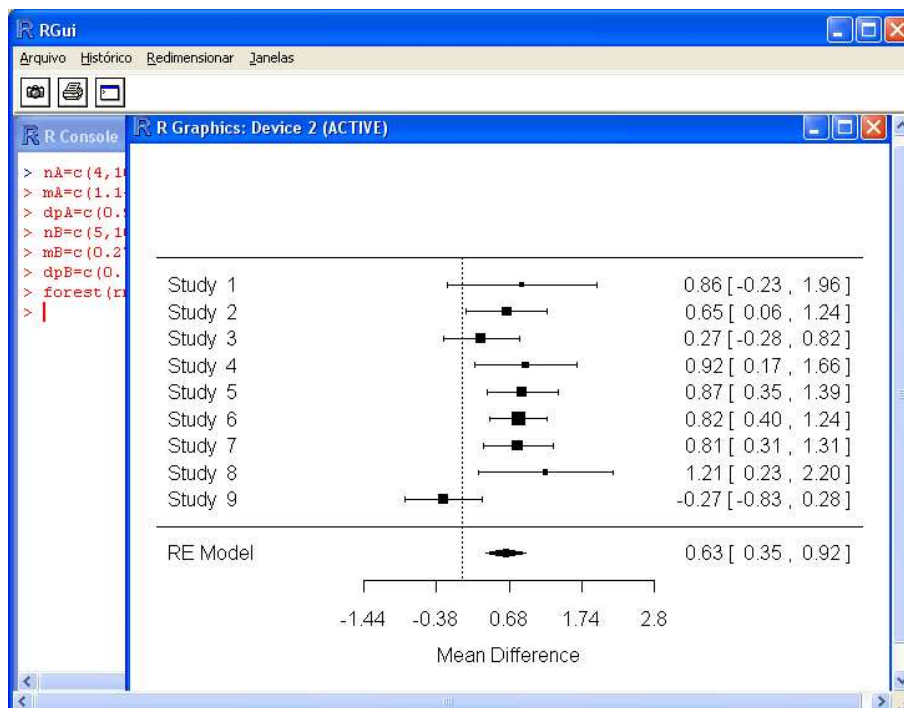


Figura 8.4 – *Forest plot* no R

9. Considerações finais

Esta monografia abordou as questões básicas de uma metanálise, como as medidas de efeito mais utilizadas na área da saúde e os modelos de efeito fixo e de efeitos aleatórios que foram, também, descritos. Noções de como identificar heterogeneidade entre os estudos e de como escolher o modelo a ser utilizado foram abordadas. Este trabalho atingiu o objetivo de servir como um guia prático para pessoas que queiram conduzir uma metanálise básica, já que não existem materiais neste formato em português.

Mostrar a utilização de dois softwares foi um ponto relevante deste trabalho, pois nem todo mundo tem acesso ao STATA que é um software pago, já o R é um software livre e qualquer um pode instalá-lo.

Existem outras técnicas utilizadas em metanálises que não foram abordadas neste trabalho como, por exemplo, abordagem bayesiana, análise de subgrupo e meta-regressão. As duas últimas são duas importantes formas de explorar a heterogeneidade, informações sobre estas técnicas podem ser vistas em Sutton, 2000.

Referências Bibliográficas

BORENSTEIN, M.; HEDGES, L. V.; HIGGINS, J. P. T., ROTHSTEIN, H. R. Introduction to Meta-Analysis, John Wiley & Sons, 2009.

BROCKWELL, S. E.; GORDON, R. A comparison of statistical methods for meta-analysis. *Statistics in Medicine*, v. 20, n. 6, p. 825-840, 2001.

EGGER, M.; SMITH, G. D.; ALTMAN, D. G. Systematic reviews in health care: Meta-analysis in context. *BMJ*, 2001.

FLETCHER, R. H.; FLETCHER, S. W. *Epidemiologia clínica: elementos essenciais*; tradução Roberta Marchiori Martins. Porto Alegre, Artmed, 2006.

GLASS, G. V. Primary, secondary, and meta-analysis of research. *Educational Researcher*. V. 5, p. 3-8, 1976.

HEDGES, L. V.; OLKIN I. *Statistical Methods for Meta-Analysis*. Academic Press. 1985.

KULINSKAYA, E.; MORGENTHALER, S.; STAUDTE R. G. *Meta Analysis: A guide to calibrating and combining statistical evidence*. John Wiley & Sons, 2008.

LEANDRO, G. *Meta-analysis in medical research: The handbook for the understanding and practice of meta-analysis*. London: BMJ Books, 2005.

MARTINEZ, E. Z. *Metanálise de ensaios controlados aleatorizados: aspectos quantitativos*. *Revista Medicina (Ribeirão Preto)*, v. 40, n.2 p. 223-235, 2007.

MARTINEZ, E. Z.; MAZIN S. C. *Modelos Estatísticos em Metanálises*. 54ª RBRAS, 13º SEAGRO, São Carlos, 2009.

NORMAND, S. L. T. *Meta-analysis: formulating, evaluating, combining, and reporting*. *Statistics in Medicine*, v. 18, p. 321-359, 1999.

NUNES, L. N. *Meta-análise*. Porto Alegre, 1997.

PEARSON, K. Report on certain enteric fever inoculation statistics. *British Medical Journal*, v 3, p.1243-1246, 1904.

SUTTON, A. J.; ABRAMS, K. R.; JONES, D. R.; SHELDON, T. A.; SONG, F. *Methods for Meta-Analysis in Medical Research*. John Wiley & Sons. 2000.

VIECHTBAUER, W. *Meta-Analysis Package for R*, 2009.

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

YATES, F.; COCHRAN, W. G. The analysis of groups of experiments. *The Journal of Agricultural Science*, v. 28, n. 4, p. 556-580, 1938.