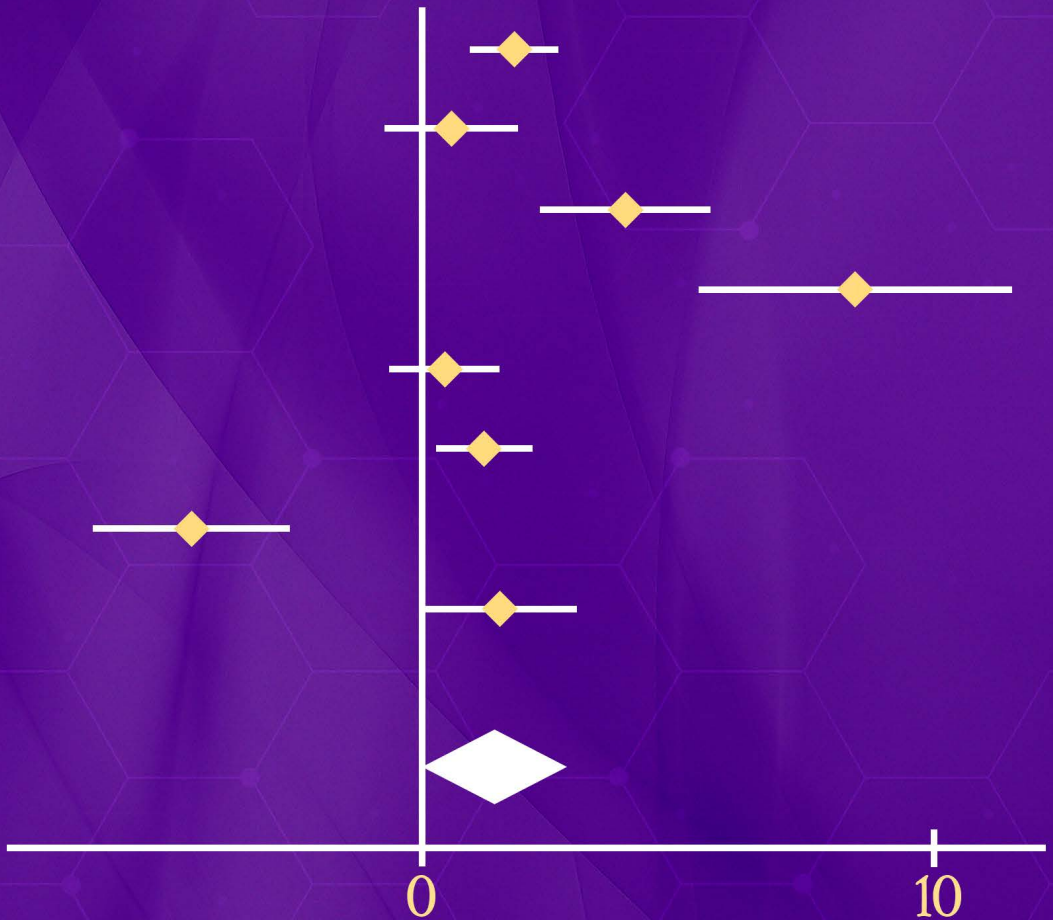


Elirez Bezerra da Silva

METANÁLISE NA ÁREA DO EXERCÍCIO FÍSICO E DA SAÚDE

O que é. Como fazer. Como redigir.
Como ler e interpretar



METANÁLISE

na área do exercício físico e da saúde

O que é

Como fazer

Como redigir

Como ler e interpretar



Pedro & João
editores

Elirez Bezerra da Silva

METANÁLISE

na área do exercício físico e da saúde

O que é

Como fazer

Como redigir

Como ler e interpretar



Pedro & João
editores

Copyright © Elirez Bezerra da Silva

Todos os direitos garantidos. Qualquer parte desta obra pode ser reproduzida, transmitida ou arquivada desde que levados em conta os direitos do autor.

Elirez Bezerra da Silva

Metanálise na área do exercício físico e da saúde. O que é. Como fazer. Como redigir. Como ler e interpretar. São Carlos: Pedro & João Editores, 2022. 82p. 16 x 23 cm.

ISBN: 978-85-7993-971-6 [Digital]

1. Metanálise. 2. Exercício físico. 3. Saúde. Título.

CDD – 310/370

Capa: Petricor Design

Ficha Catalográfica: Hélio Márcio Pajeú – CRB - 8-8828

Diagramação: Diany Akiko Lee

Editores: Pedro Amaro de Moura Brito & João Rodrigo de Moura Brito

Conselho Científico da Pedro & João Editores:

Augusto Ponzio (Bari/Itália); João Wanderley Geraldi (Unicamp/ Brasil); Hélio Márcio Pajeú (UFPE/Brasil); Maria Isabel de Moura (UFSCar/Brasil); Maria da Piedade Resende da Costa (UFSCar/Brasil); Valdemir Miotello (UFSCar/Brasil); Ana Cláudia Bortolozzi (UNESP/Bauru/Brasil); Mariangela Lima de Almeida (UFES/ Brasil); José Kuiava (UNIOESTE/Brasil); Marisol Barenco de Mello (UFF/Brasil); Camila Caracelli Scherma (UFFS/Brasil); Luis Fernando Soares Zuin (USP/Brasil).



Pedro & João Editores

www.pedroejoaoeditores.com.br

13568-878 – São Carlos – SP

2022

Agradecimentos

Ao Exército brasileiro por ter me autorizado a fazer os cursos de mestrado e doutorado.

A Escola de Educação Física do Exército pela minha formação em educação física e por ter me proporcionado praticar ensino, pesquisa e desporto por quase 20 anos.

Ao Instituto de Pesquisa da Capacitação Física do Exército pela oportunidade de desenvolver pesquisas na área da aptidão física e saúde.

Ao Instituto Brasileiro de Medicina e Reabilitação pela minha formação em fisioterapia.

A Universidade Gama Filho (UGF) por ter me acolhido para os cursos de mestrado e doutorado.

Ao Prof. Dr. Paulo Sérgio Chagas Gomes por ter sido meu orientador nos dois cursos e ter aberto a janela para o mundo científico.

Ao Programa de Pós-graduação em Ciências do Exercício e Esporte da Universidade do Estado do Rio de Janeiro (PPGCEE / UERJ) pelo meu trabalho como professor visitante e professor pesquisador.

Ao Programa de Mestrado Profissional em Gestão do Trabalho para a Qualidade do Ambiente Construído da Universidade Santa Úrsula pelo meu trabalho como professor permanente.

Aos meus alunos de mestrado, doutorado e integrantes do Grupo de Pesquisa em Ciências do Exercício e da Saúde (www.gpces.com.br) pela oportunidade em ensinar e aprender.

À Marcelle Brandão pela ideia.

À Deus por mais este dia...

Nestes meus quase 30 anos de docência e pesquisa, orientando futuros mestres e doutores, pude vivenciar diferentes formas de fazer pesquisas, entre elas, a metanálise. Este livro é uma síntese desta vivência de orientação de 20 dissertações de mestrado e 08 teses de doutorado com enfoque nos métodos quantitativos para investigação na área da saúde, nas quais estão incluídas 19 revisões sistemáticas com e sem metanálises publicadas em periódicos de QUALIS A1 a B1.

Por isso mesmo, o livro é prático. A primeira parte do livro é introdutória. A segunda parte do livro mostra como fazer uma metanálise. Para isto, faz uso de dois aplicativos para metanálise, o *RevMan* e o JAMOVI, disponíveis gratuitamente na internet para *download*, juntamente com quatro exemplos verdadeiros de metanálises desenvolvidas por alunos de mestrado e/ou doutorado, sob minha orientação e dois exemplos fictícios de metanálises, que eu criei para melhor explicar o assunto. Os exemplos utilizados neste livro mostram os resultados de metanálises de média, diferença de médias (DM), diferença de médias padronizada (DMP), risco relativo (RR), diferença de riscos (DR) e *odds ratio* (OR), utilizando o *RevMan* e o JAMOVI, mas mostram também como são feitos esses cálculos pelos dois aplicativos. A terceira parte do livro destina-se a uma discussão interessante sobre significância estatística e significância clínica, que eu prefiro chamar de relevância clínica. A quarta parte do livro apresenta boas recomendações para a redação de metanálises. A quinta parte finaliza o livro dando boas dicas para a leitura rápida, dinâmica e crítica de metanálises.

Desejo a você uma boa leitura.

Elirez Bezerra da Silva

Acesso ao currículo LATTES do autor:

<http://lattes.cnpq.br/7486340493431857>

Endereço eletrônico do autor:

elirezsilva@cosmevelho.com.br

Prefácio

Em 30 de janeiro de 2020 o surto do novo corona vírus foi declarado pela Organização Mundial da Saúde uma Emergência de Saúde Pública de Importância Internacional. Desde então, até bem recentemente, praticamente todas as instituições educacionais e/ou de pesquisa foram obrigadas a obedecer ao distanciamento social. Tal fato provocou, em todo o meio científico, interromper e/ou mudar uma série de projetos de pesquisa que inicialmente se caracterizavam por modelos experimentais (e suas variantes). Grande parte destas mudanças que todos nós experimentamos foi continuar estudando o mesmo tópico, mas, em vez de estudos experimentais optar por revisões sistemáticas, muitas vezes com metanálise. A estatística do número de publicações, que o autor apresenta na sua introdução é um sinal claro da importância ainda maior dada aos estudos de metanálise. Como bem coloca o autor desta obra, o termo metanálise foi introduzido “há cerca de 40 anos depois que ela foi utilizada pela primeira vez para verificar a eficácia de um tratamento na década de 1950”. Na impossibilidade do contato com outras pessoas, alunos de pós-graduação e orientadores acadêmicos valorizaram ainda mais os estudos de revisão sistemática e metanálise. A quantidade destes estudos aumentou exponencialmente em quase todas as áreas do conhecimento, o foi acompanhado pela área de ciências do exercício e do esporte. Muito se aprendeu sobre metanálise e os progressos também foram muitos e significativos. O próprio estudo Atividade física para adultos: Guia de Atividade Física para a População Brasileira (Rev Bras Ativ Fis Saúde, 2021, 10.12820/rbafs.26e0215), um marco importante na Educação Física no Brasil, baseou-se na análise de um número grande de estudos de metanálise realizados em vários países, incluindo Brasil, nos quais meus colegas e eu tivemos de nos debruçar para a elaboração do Guia. A importância e a relevância de estudos cujo objetivo é ajudar profissionais de diversas áreas do conhecimento a tomar decisões a respeito de intervenções, há muito já se tornou uma referência na área de ciências do exercício, do esporte e da saúde. Entretanto, algo que parece simples de ser feito, na verdade envolve grande complexidade nas suas diversas fases de elaboração. Desde a escolha da pergunta a ser respondida, bases de dados onde serão recuperadas as informações, seleção dos documentos a serem

analisados, análise e interpretação dos resultados e redação do documento, a jornada parece interminável. Isto tudo sem falar na importância que é a leitura de uma metanálise. Esta obra veio numa boa hora, pois se insere de maneira apropriada no atual cenário científico, principalmente nas áreas propostas neste documento. No meu entender o professor Elirez é talhado para esta tarefa e aqui está a prova irrefutável. Para mim não é uma novidade, pois há muito o professor Elirez tem demonstrado sua competência em suas aulas e seminários, desde os seus tempos de mestrando, doutorando e, com certeza, há seis anos como professor do Programa de Pós-graduação Stricto Senso em Ciências do Exercício e do Esporte da Universidade do Estado do Rio de Janeiro. Suas apresentações sempre foram extremamente organizadas, com abordagem simples, porém profunda e muitas vezes bastante complexas, sempre com qualidade ímpar, caracterizada também pela objetividade e precisão cirúrgicas que são suas características. Esta obra, que ele compartilha conosco, foi recebida por mim com muita alegria, por que não dizer com muito orgulho também. Elirez nos apresenta um documento com excelente conteúdo, didaticamente bem construído, prático, com exemplos de várias publicações de sua autoria, e o melhor, nas áreas de educação física, fisioterapia e afins. Os exemplos compartilhados são fundamentais no entendimento do processo, tornando a tarefa do(a) leitor(a) bastante simplificada. Com certeza, esta publicação será um sucesso, ou melhor, um dos livros de cabeceira de muitos futuros alunos de pós-graduação e, vislumbro ainda, uma referência que será recomendada por professores de metodologia da pesquisa e estatística como eu. Parabênico meu colega pela excelente iniciativa e produto final de muito alto nível.

Um forte abraço.

Paulo Sergio Chagas Gomes,

Ph.D.

Departamento de Ciência da Atividade Física, Instituto de Educação Física
Desportos da Universidade do Estado do Rio de Janeiro

Sumário

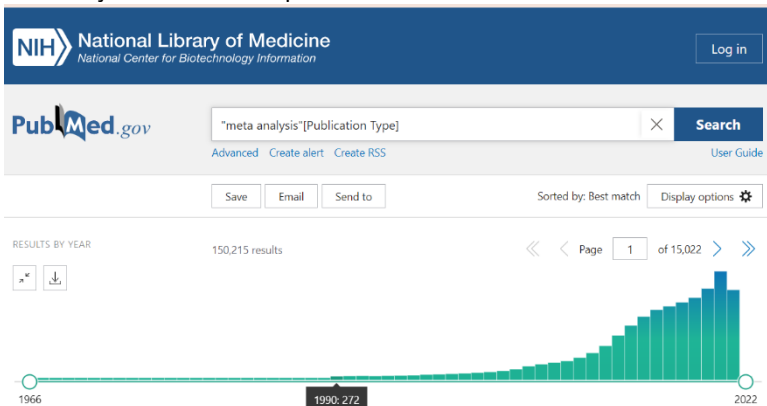
Parte 1 - INTRODUÇÃO	10
Parte 2 - FAZENDO A METANÁLISE	15
Justificativa para elaboração da metanálise	16
Definição do objetivo	16
Estabelecimento dos critérios de inclusão dos estudos	17
Criação da estratégia de busca e seleção dos estudos	18
Avaliação da qualidade metodológica e vieses dos estudos incluídos	22
Extração dos dados dos estudos incluídos	24
Metanálise dos dados extraídos	25
Tipo de desfecho e medidas de efeito	26
Modelo de análise e heterogeneidade	29
Peso dos estudos e resultado da metanálise	32
Viés de publicação	33
Média	36
Diferença de médias	43
Diferença de médias padronizada	46
Risco relativo	51
Diferença de risco	55
<i>Odds ratio</i>	57
Avaliação do nível de evidência da metanálise	62
Parte 3 - SIGNIFICÂNCIA ESTATÍSTICA E/OU SIGNIFICÂNCIA CLÍNICA	67
Parte 4 - REDIGINDO A METANÁLISE	72
Parte 5 - LENDO E INTERPRETANDO A METANÁLISE	76
REFERÊNCIAS	80

INTRODUÇÃO

O termo “metanálise” surgiu há cerca de 40 anos depois que ela foi utilizada pela primeira vez para verificar a eficácia de um tratamento na década de 1950. A partir de 1980, passou a ser um tipo de pesquisa frequente e importante na área de saúde para a tomada de decisões e estabelecimento de diretrizes, mas tem sido utilizada também em outras áreas como a educação, finanças e ecologia.

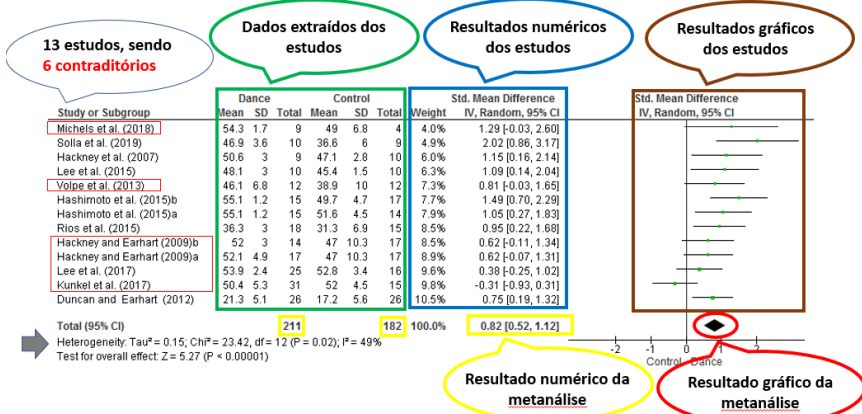
Para se fazer uma metanálise, é necessário somente um computador, internet e conhecimento. A produção de metanálises aumentou exponencialmente com a comercialização da internet a partir da década de 90. Antes de 1990, sem a comercialização da internet, que existia para servir somente aos órgãos governamentais e educacionais dos países, a produção foi praticamente nula. Em 1990 havia 272. Em 2000 havia 848. Em 2010 havia 4691. Em 2020 havia 20104, totalizando hoje 150215 metanálises publicadas na base de dados PubMed (Figura 1). Hoje em dia, computador e internet são bem mais fáceis de se ter. O mais difícil é o conhecimento. Por isso mesmo, o objetivo deste livro é ajudá-lo a saber o que é, como fazer, como redigir, como ler e interpretar uma metanálise.

Figura 1 – Aumento exponencial da produção de metanálises com a comercialização da internet a partir de 1990



Metanálise é uma técnica estatística que cria um resultado a partir dos resultados de outros estudos, que tiveram os mesmos objetivos e foram revisados sistematicamente. No meio acadêmico, as revisões sistemáticas que fazem uso desta técnica estatística passaram a ser nominadas como metanálises.

Figura 2 – Forest plot da metanálise do controle postural de participantes com a doença de Parkinson que realizaram um programa de atividades de dança versus controle. Mean = média; SD = desvio-padrão; Total = quantidade de participantes; Weight = peso do estudo; Std. Mean Difference = diferença de médias padronizada; IV = método inverso da variância; Random: modelo de efeitos randômicos; 95% CI = intervalos de confiança 95%; Heterogeneity = heterogeneidade; I^2 = inconsistência; Test for overall effect = teste do efeito geral (da metanálise)



Fonte: Almeida H. (2020)

A Figura 2, chamada de *forest plot*, ajuda a entender sumariamente o que é uma metanálise e como ela cria um novo resultado: Foram incluídos nesta metanálise 13 estudos, 7 deles mostraram a eficácia da prática da dança para melhorar o controle postural de Parkinsonianos, enquanto 6 deles (inclusos nos retângulos vermelhos) mostraram a ineficácia. Na cor verde, os resultados de controle postural dos Parkinsonianos do “grupo dança” e “grupo controle” extraídos de cada um dos 13 estudos incluídos na metanálise. Na cor azul, os resultados numéricos dos estudos decorrentes da análise dos 13 resultados de controle postural extraídos dos estudos, com os correspondentes resultados gráficos na cor marrom. Na cor amarela, o

resultado numérico da metanálise decorrentes desses 13 resultados numéricos dos estudos, com o correspondente resultado gráfico da metanálise na cor vermelha.

Para melhor compreensão do processo de metanálise que será mostrado mais a frente, é importante saber diferenciar “dados extraídos dos estudos”, “resultados dos estudos” e “resultado da metanálise”. Dados extraídos dos estudos são os resultados de controle postural contidos nos estudos, que foram extraídos e incluídos para serem metanalisados. Resultados dos estudos são a medida de efeito (no caso da Figura 2, a medida de efeito foi a diferença de média padronizada – *Std. Mean Difference*), o intervalo de confiança 95% e o peso de cada estudo. Resultado da metanálise é a medida de efeito e intervalo de confiança 95% da metanálise. Os métodos estatísticos, modelos de análises, medidas de efeitos e heterogeneidade serão abordados mais a frente neste livro.

A replicação ou reprodutibilidade do resultado de uma pesquisa é uma característica do conhecimento científico, ou seja, diversos estudos realizados e publicados que tiveram o mesmo objetivo, deveriam replicar ou reproduzir o mesmo resultado. Entretanto, nem sempre isto ocorreu. Pelo contrário, estudos com o mesmo objetivo apresentaram resultados diferentes e, às vezes, contraditórios. Ora, como tomar uma decisão acertada e confiável em uma situação desta? Neste caso, esta dúvida será dirimida pela metanálise, que considerando esses resultados diferentes e/ou contraditórios, criará um novo resultado, apontando a decisão a ser tomada com mais precisão e confiança.

Toda metanálise deve ser precedida de uma boa revisão, que deve conter as seguintes características básicas: sistemática, específica, generalizada e crítica.

Sistemática porque segue os mesmos passos durante sua elaboração. Metanálises realizada aqui, ou no Japão, ou nos EUA, ou na Alemanha serão diferentes em relação aos autores e objetivos estabelecidos, mas muito semelhantes em relação ao planejamento, ao método, ou aos passos a serem dados para a sua elaboração. Normalmente, a sistematização exige justificar a elaboração da metanálise, definir o objetivo, estabelecer os critérios de inclusão dos estudos, criar a estratégia de busca e selecionar os estudos, avaliar a qualidade metodológica e vieses dos estudos incluídos, extrair os dados

dos estudos incluídos, metanalisar os dados extraídos e avaliar o nível de evidência da metanálise.

Específica porque delimita o assunto a ser revisado, definindo claramente qual o tratamento e qual a resposta dos participantes (nas ciências da saúde esta resposta pode ser chamada de desfecho ou ainda variável resposta ou ainda quadro clínico ou ainda doença) a serem estudados. Se necessário, pode delimitar ainda algumas características dos participantes, o grupo de comparação e o tipo de estudo. O grupo de comparação é aquele comparado com o grupo tratamento (redundância necessária), que pode realizar outro tratamento ou servir de controle. Sendo controle, nenhum tratamento ele fará. Há controles que são chamados de “*sham*” (palavra em inglês que significa falso), ou seja, controle que fará um tratamento que não produzirá alguma resposta dos participantes. Por isso “*sham*”, falso, sem efeito. O tipo de estudo, também chamado de desenho do estudo, é a forma como o estudo foi planejado. Há diferentes tipos de estudo, sendo os mais frequentes na área da saúde o experimental controlado randomizado, experimental pragmático, pré-experimental, coorte, caso controle, transversal, de regressão simples, de regressão múltipla, de confiabilidade e de validade. Na medicina, o experimental controlado randomizado e o de validade são chamados de ensaio clínico e de acurácia diagnóstica, respectivamente. Tipos de estudo não é escopo deste livro, mas é muito importante saber as características e limitações de cada tipo de estudo e isto pode ser resolvido lendo um bom livro de metodologia da pesquisa científica. Importante saber também que a metanálise não deve misturar os tipos de estudo, ou seja, a metanálise deve conter estudos que tenham o mesmo desenho, tipo ou planejamento de pesquisa.

Generalizada porque faz a busca de todos os artigos publicados sobre o assunto a ser revisado. Parece paradoxal ser específica e generalizada ao mesmo tempo, mas não é. A generalização se refere a busca dos estudos que deve ser abrangente, considerando todas as bases eletrônicas de informações científicas e outras fontes que tenham afinidade com o assunto da metanálise.

Crítica porque leva em consideração a qualidade metodológica dos estudos incluídos para o novo resultado a ser criado pela metanálise. Todos os estudos incluídos contribuem para o resultado final da metanálise, mas com pesos diferentes. Estudos com melhor qualidade metodológica têm pesos maiores. Veja “*Weight*” na cor azul na Figura 2

acima: o estudo da Duncan and Earhart (2012) contribuiu com 10,5%, enquanto que o estudo de Michels et al. (2018) contribuiu com 4%. Como o “Weight” foi obtido e o seu significado será abordado também mais a frente neste livro.

FAZENDO A METANÁLISE

Para facilitar a compreensão de como se faz uma metanálise, serão utilizados quatro exemplos de metanálises elaboradas por alunos de mestrado e/ou doutorado sob minha orientação e dois exemplos fictícios:

Exemplo 1 – Metanálise de 13 estudos para verificar a eficácia da dança sobre o controle postural de pessoas com doença de Parkinson;

Exemplo 2 – Metanálise de 05 estudos para verificar a eficácia de exercícios físicos crônicos sobre a capacidade funcional de pessoas com cirrose hepática;

Exemplo 3 – Metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos;

Exemplo 4 – Metanálise de 17 estudos para verificar o exercício físico como fator de proteção contra lesões musculoesqueléticas;

Exemplo 5 - Metanálise de 5 estudos para verificar a obesidade como fator de risco para hipertensão arterial (exemplo fictício);

Exemplo 6 - Metanálise de 5 estudos para verificar a associação do fator “não lavagem das mãos pelos médicos” e infecção (exemplo fictício).

Antes de começar a fazer uma metanálise é interessante conhecer o aplicativo EndNote disponível em: <https://access.clarivate.com/login?app=endnote> e a plataforma Registro Internacional Prospectivo de Revisões Sistemáticas (PROSPERO) disponível em <https://www.crd.york.ac.uk/prospero/>. O aplicativo EndNote será útil para detectar os estudos duplicados e para a citação e referências dos estudos. A plataforma PROSPERO mostrará se há alguma metanálise, idêntica à sua, em andamento e será a plataforma onde você deverá registrar a metanálise que quer fazer. Os seguintes tutoriais poderão ajuda-los quanto a plataforma PROSPERO e aplicativo EndNote: <http://www.gpces.com.br/arquivos/133670721615421087.7682.pdf>

<http://www.gpces.com.br/arquivos/418177581615421007.3003.pdf>

Como todo e qualquer estudo, a metanálise requer um planejamento rigoroso, detalhado e cuidadoso. Este planejamento, que irá lhe poupar tempo e facilitar a execução das tarefas necessárias, compreende oito passos, nesta ordem:

- 1) Justificativa para elaboração da metanálise
- 2) Definição do objetivo
- 3) Estabelecimento dos critérios de inclusão dos estudos
- 4) Criação da estratégia de busca e seleção dos estudos
- 5) Avaliação da qualidade metodológica e vieses dos estudos incluídos
- 6) Extração dos dados dos estudos incluídos
- 7) Metanálise dos dados extraídos e
- 8) Avaliação do nível de evidência da metanálise

Justificativa para elaboração da metanálise

Se você tiver a pretensão de publicar a metanálise em um periódico, há que ter uma boa justificativa para a sua elaboração. Faz-se uma metanálise quando há estudos suficientes, diferentes e/ou contraditórios, mas não há metanálise publicada sobre o assunto que se quer pesquisar.

Porém, há outras situações que requerem a elaboração da metanálise: quando há metanálise, mas há estudos suficientes para alterar o resultado dessa metanálise ou quando a metanálise existente apresentou limitações que prejudicaram a precisão e confiança nos resultados obtidos.

Definição do objetivo

Sem objetivo não se chega a lugar algum. É fundamental definir com clareza e concisão o objetivo da metanálise, porque o objetivo tem relações com o título da metanálise e com os dois passos seguintes: O estabelecimento dos critérios de inclusão e a criação da estratégia de busca dos estudos. O objetivo deve conter o tratamento, o desfecho e as principais características dos participantes.

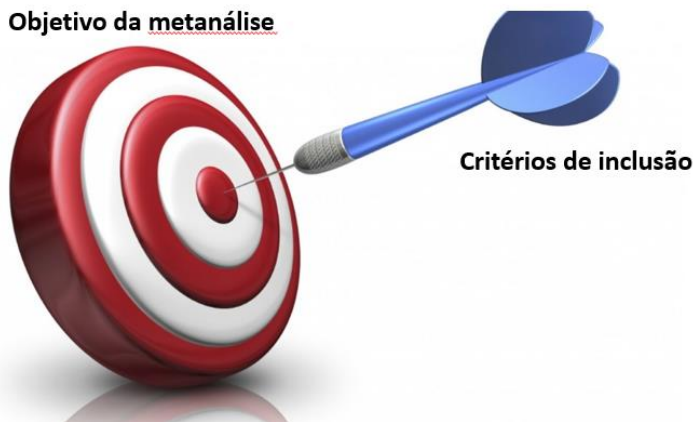
Utilizando o Exemplo 1 - Metanálise de 13 estudos para verificar a eficácia da dança sobre o controle postural de pessoas com doença de Parkinson, por exemplo, o objetivo desta metanálise foi verificar o efeito da dança sobre o controle postural de pessoas com doença de Parkinson, onde a dança é o tratamento, o controle postural é o desfecho e ter a doença de Parkinson é a principal característica dos participantes dos estudos.

Estabelecimento dos critérios de inclusão dos estudos

Os critérios de inclusão delimitam as características que devem ter os estudos para poderem ser incluídos na metanálise. Há diferentes siglas mnemônicas para auxiliar no estabelecimento dos critérios de inclusão. Esqueça todas elas e observe o objetivo da metanálise que foi definido. Este objetivo irá orientar os critérios de inclusão dos estudos (Figura 3). Por exemplo, considerando o objetivo enunciado acima, os critérios de inclusão seriam estudos que aplicaram a dança como tratamento e mediram o controle postural como desfecho em pessoas com doença de Parkinson.

Poderiam ser acrescentados, se necessários, mais dois critérios: o grupo de comparação com o tratamento dança e o tipo de estudo. Não esqueça que não se deve misturar diferentes tipos de estudo na mesma metanálise. Outro cuidado que se deve ter é não estabelecer data de publicação e idioma dos estudos como critérios de inclusão, porque criam um viés no resultado da metanálise muito difícil de ser explicado.

Figura 3 – Os critérios de inclusão dos estudos devem ser coerentes com o objetivo da metanálise



Criação da estratégia de busca e seleção dos estudos

A estratégia de busca dos estudos é a combinação engenhosa de descritores, termos de busca, que se faz para recuperar os estudos nas bases eletrônicas de informações científicas e outras fontes. Ela requer criatividade e paciência.

Para se criar uma boa estratégia de busca é necessário conhecer as bases eletrônicas de informações científicas. Na área das ciências da saúde há, pelo menos, 11 bases que devem ser consultadas (Figura 4). As bases eletrônicas de informações científicas têm inteligência artificial que recuperam os artigos por palavras-título e/ou palavras-chave informadas no idioma inglês. Estas palavras são denominadas descritores.

Para criar a estratégia de busca, é necessário definir quais são os descritores que serão utilizados. O objetivo ou os critérios de inclusão dos estudos da metanálise irão ajudar a encontrar estes descritores. Por exemplo, considerando os critérios de inclusão listados acima, os descritores poderiam ser “dança” e “controle postural”. Para que a busca seja abrangente, os sinônimos devem ser considerados. Há duas plataformas que ajudarão com os descritores e seus sinônimos: Descritores em Ciências da Saúde (DeCS) disponível em: <https://decs.bvsalud.org/en/> e *Medical Subject Headings* (MeSH) disponível em: <https://www.ncbi.nlm.nih.gov/mesh/>.

Figura 4 – Bases eletrônicas de informações científicas na área de ciências da saúde



Algumas bases admitem a frase de busca, que é uma combinação dos descritores e seus sinônimos, utilizando os operadores booleanos AND entre os descritores e OR entre os sinônimos. Muito cuidado com o uso de frase de busca, porque podem gerar erros facilmente. Uma sugestão de estratégia de busca simples e eficaz é a seguinte: Insira o primeiro descritor e observe a quantidade de registros recuperados; acrescente o segundo descritor utilizando o AND e observe a quantidade de registros recuperados, que deverá ser menor que a quantidade anterior; acrescente um sinônimo do primeiro descritor utilizando o OR e observe a quantidade de registros recuperados, que deverá ser maior ou igual a quantidade anterior; acrescente um sinônimo do segundo descritor utilizando o OR e observe a quantidade de registros recuperados, que deverá ser maior ou igual a quantidade anterior; repita isto ora para o primeiro ora para o segundo descritor até utilizar todos os sinônimos (Figura 5). Evite usar o idioma e a data de publicação dos estudos como filtros, porque eles criam uma tendenciosidade difícil de ser explicada depois.

Sugiro a leitura dos seguintes tutoriais gratuitos na internet, que ensinam como navegar nas 11 bases eletrônicas de informações científicas listadas na Figura 4:

<http://www.gpces.com.br/arquivos/412209971565728992.6664.pdf>
(Instruções iniciais)

<http://www.gpces.com.br/arquivos/943079921565729115.5075.pdf>

(Base LILACS)

<http://www.gpces.com.br/arquivos/84896171565729193.7943.pdf>

(Base SciELO)

<http://www.gpces.com.br/arquivos/392513071565730517.0659.pdf>

(Base Scopus)

<http://www.gpces.com.br/arquivos/846014961565729302.3299.pdf>

(Base Science Direct)

<http://www.gpces.com.br/arquivos/905902041565729368.3225.pdf>

(Base Web of Science)

<http://www.gpces.com.br/arquivos/387999861565729418.3892.pdf>

(Base PEDro)

<http://www.gpces.com.br/arquivos/770238731565729510.0364.pdf>

(base Cochrane)

<http://www.gpces.com.br/arquivos/479037731565729583.8048.pdf>

(Base MEDLINE)

<http://www.gpces.com.br/arquivos/836832691565730380.1179.pdf>

(Base CINAHL)

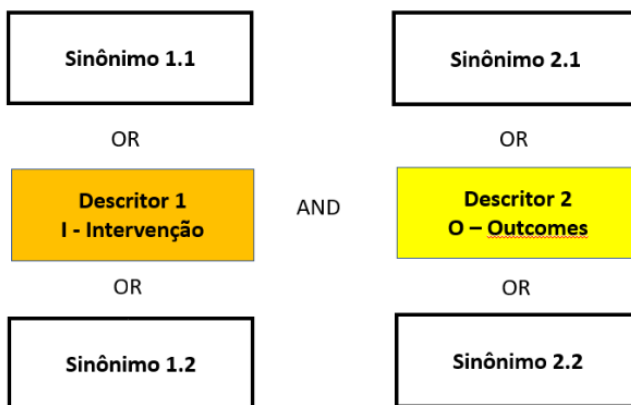
<http://www.gpces.com.br/arquivos/39889301565730405.276.pdf>

(Base Sport Discus)

Figura 5 – Uma sugestão de estratégia de busca dos estudos simples e eficaz

Menos é mais...
Simples é melhor...

Quantidade de estudos



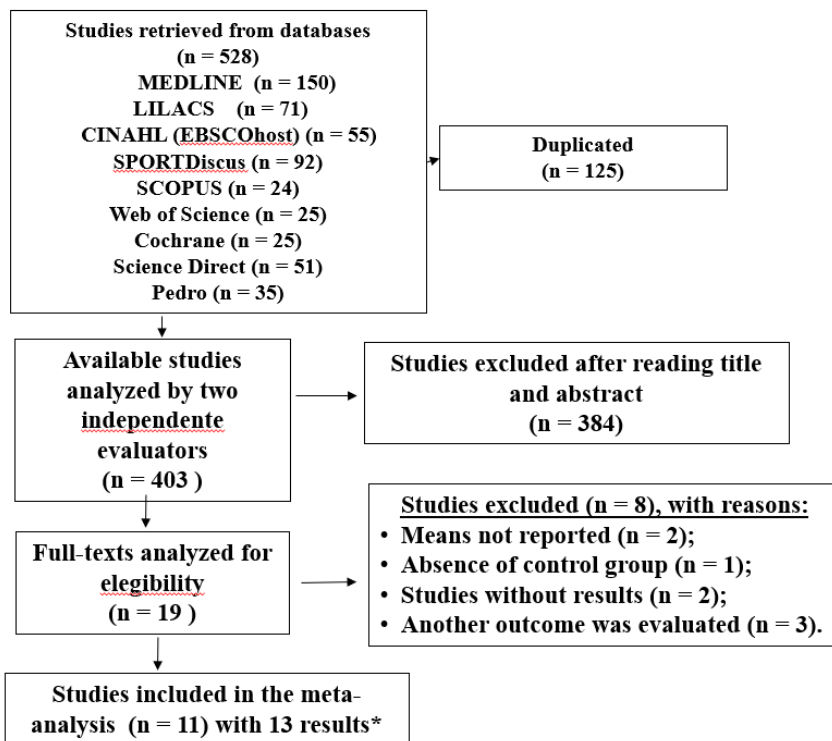
A seleção dos estudos deve ser feita por dois avaliadores independentes. Um terceiro avaliador deverá ser acionado no caso de divergências. A seleção dos estudos duplicados pode ser feita pelo

aplicativo *EndNote*, caso você esteja utilizando-o. Do contrário, será feita manualmente e isto tomará muito tempo e gastará sua paciência. Excluídos os duplicados, serão lidos os títulos e resumos dos estudos remanescentes, para ver se atendem integralmente aos critérios de inclusão estabelecidos. Excluídos os que não atenderam aos critérios de inclusão estabelecidos, serão lidos os textos completos dos estudos remanescentes, para ver se atendem integralmente aos critérios de inclusão estabelecidos (isso mesmo, mais uma vez).

A execução dos dois passos anteriores gera o primeiro resultado da metanálise que é o diagrama de fluxo dos estudos, que pode ser feito manualmente ou utilizando a plataforma *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA 2020 Statement), disponível em <http://www.prisma-statement.org/>

O diagrama de fluxo dos estudos é uma contabilidade da entrada e saída dos estudos. A entrada será a quantidade de estudos recuperados nas bases eletrônicas de informações científicas e outras fontes e a saída é a quantidade de estudos excluídos por duplicação, por não atender os critérios de inclusão estabelecidos e outras razões, que devem ser descritas. A Figura 6 mostra um exemplo de diagrama de fluxo dos estudos.

Figura 6 – Diagrama de fluxo dos estudos



Fonte: Almeida H. (2020)

Avaliação da qualidade metodológica e vieses dos estudos incluídos

A avaliação da qualidade metodológica e vieses dos estudos deve ser feita por dois avaliadores independentes. Um terceiro avaliador deverá ser acionado no caso de divergências. Normalmente, utiliza-se ferramentas ou escalas de avaliação já validadas e publicadas, que variam de acordo com o tipo de estudo incluído na metanálise. Conhecer os possíveis vieses dos diferentes tipos de estudo empregados nas ciências da saúde e como se pode controlar os vieses é fundamental, porque as escalas foram construídas para avaliar a presença ou ausência de vieses. A escolha da ferramenta ou escala fica a critério dos autores. No caso do Exemplo 1, foram escolhidas a Escala PEDro e a Ferramenta da Cochrane. Sugiro a leitura dos seguintes tutoriais gratuitos na internet, que ensinam como utilizar a Escala PEDro, Jadad e a RoB 2 - Ferramenta da Colaboração Cochrane:

<http://www.gpces.com.br/arquivos/132853411558411427.5384.pdf>
(Escala PEDro)
<http://www.gpces.com.br/arquivos/915507671558410792.4702.pdf>
(Escala Jadad)
<http://www.gpces.com.br/arquivos/802745161617660757.6202.pdf>
(RoB 2 - Ferramenta da Colaboração Cochrane)

A execução deste passo gera o segundo e terceiro resultados da metanálise que é o quadro ou tabela com o resultado desta avaliação (Figura 7 e 8).

Figura 7 – Avaliação da qualidade metodológica dos estudos, utilizando a Escala PEDro, disponível em <https://pedro.org.au/portuguese/resources/pedro-scale/>, que contém 11 critérios. Cada critério de 2 a 11, pontuou 1 ponto quando atendido, totalizando até 10 pontos. Yes= sim; No = não. O critério 1 não pontua.

Study	Score	1	2	3	4	5	6	7	8	9	10	11
Hackney et al., 2007	5 / 10	yes	yes	no	yes	no	no	yes	no	no	yes	yes
Hackney and Earhart, 2009	5 / 10	yes	yes	no	yes	no	no	yes	no	no	yes	yes
Duncan and Earhart, 2012	7 / 10	yes	yes	yes	yes	no	no	yes	no	yes	yes	yes
Volpe et al., 2013	6 / 10	yes	yes	yes	yes	no	no	yes	no	no	yes	yes
Romenets et al., 2015	6 / 10	yes	yes	no	yes	no	no	no	yes	yes	yes	yes
Hashimoto et al., 2015	6 / 10	yes	yes	yes	yes	no	no	yes	no	no	yes	yes
Lee NY et al., 2015	4 / 10	no	yes	no	yes	no	no	no	no	no	yes	yes
Kunkel et al., 2017	6 / 10	yes	yes	no	yes	no	no	yes	yes	no	yes	yes
Lee HJ et al., 2018	8 / 10	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes
Michels et al., 2018	7 / 10	yes	yes	yes	no	no	no	yes	yes	yes	yes	yes
Sollas et al., 2019	6 / 10	yes	yes	no	yes	no	no	yes	yes	no	yes	yes

Fonte: Almeida H. (2020)

Figura 8 – Avaliação do risco de viés dos estudos, utilizando a Ferramenta Cochrane, disponível em <https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2>, que contém 7 domínios, podendo cada domínio ser avaliado com um dos três níveis: baixo, incerto ou alto risco. O risco de viés do estudo foi aquele de maior nível entre os 7 domínios. Low risk = baixo risco; Uncertain risk = risco incerto; High risk = alto risco.

Studies	1	2	3	4	5	6	7	Risk
(Hackney et al., 2007)	Low risk	Low risk	Low risk	Uncertain risk	Low risk	Uncertain risk	Low risk	Uncertain risk
(Hackney and Earhart, 2009)	Uncertain risk	Uncertain risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk
(Duncan and Earhart, 2012)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk	Uncertain risk
(Volpe et al., 2013)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk	Uncertain risk
(Rios Romenets et al., 2015)	High risk	Uncertain risk	Low risk	Low risk	Low risk	Uncertain risk	Uncertain risk	High risk
(Hashimoto et al., 2015)	Low risk	Uncertain risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk
(Lee NY et al., 2015)	Uncertain risk	Uncertain risk	Low risk	Uncertain risk	Uncertain risk	Low risk	Uncertain risk	Uncertain risk
(Kunkel et al., 2017)	Uncertain risk	Uncertain risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk
(Lee HJ et al., 2018)	Low risk	Uncertain risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk
(Michels et al., 2018)	Low risk	Uncertain risk	Low risk	Uncertain risk	Low risk	Low risk	Low risk	Uncertain risk
(Sollas et al., 2019)	Low risk	Uncertain risk	Low risk	Low risk	Low risk	Low risk	Low risk	Uncertain risk

Fonte: Almeida H. (2020)

Extração dos dados dos estudos incluídos

Para evitar possíveis erros na leitura dos dados, no registro dos dados e na interpretação de quadros, tabelas ou figuras, é obrigatório que este passo seja executado por dois avaliadores e um terceiro avaliador para dirimir as discordâncias. As principais características dos participantes dos estudos, os parâmetros do tratamento aplicado e os resultados do desfecho devem ser extraídos, para tornar possível a formulação de diretrizes. Os critérios de inclusão dos estudos podem orientar a decisão de quais dados devem ser extraídos dos estudos. A execução deste passo gera o quarto resultado da metanálise que é o quadro com os dados extraídos (Figura 9).

Figura 9 – Exemplo de um quadro com os dados extraídos dos estudos. Foram suprimidos os dados dos outros 08 estudos. *Authors* = autores; *Sample size, age (years) and sex* = tamanho da amostra, idade (anos) e sexo; *Staging Hohen and Yarh* = Estágio da doença de Parkinson, segundo o teste Hohen e Yarh; *Intervention* = intervenção; *Duration and Weekly frequency of intervention* = duração e frequência semanal da intervenção; *Postural control (Berg Balance and MiniBest)* = controle postural medido pelos testes Berg Balance e MiniBest.

Público-alvo			Diretrizes		Metanálise	
Authors	Study Status	Sample size, age (years) and sex	Staging Hohen and Yarh *	Intervention	Duration and weekly frequency of intervention	Postural Control (Berg Balance and MiniBest)
(Hackney et al., 2007)	Old	Dance Group: n=9 Age: 72.6 ± 6.6 Gender (M/F): 6/3 Control Group: n=10 Age 69.6 ± 6.64 Gender (M/F): 6/4	Between stages 2 and 3	Tango X Supervised exercises	13 weeks 21 sessions of 1h each 2 x per week for Tango and the same for the exercises	Postural balance measured with Berg Balance Scale (BBS) Always at the same time of day, with subjects taking medication Scores: Dance: 50.6 ± 3 Control: 47.1 ± 2.8
(Hackney and Earhart, 2009a)	Old	Dance Group: n=17 Age: 68.2 ± 5.23 Gender (M/F): 14/3 Control Group: n=17 Age: 66.5 ± 11.54 Gender (M/F): 12/5	Between stages 1 and 3	Tango X Unsupervised control	13 weeks 20 sessions of 1h each 2 x per week	Postural balance measured with Berg Balance Scale (BBS) Always at the same time of day, with subjects taking medication Scores: Dance: 52.1 ± 4.9 Control: 47.0 ± 10.30
(Hackney and Earhart, 2009b)	Old	Dance Group: n=14 Age: 66.8 ± 9.89 Gender (M/F): 11/3 Control Group: n=17 Age: 66.5 ± 11.54 Gender (M/F): 12/5	Between stages 1 and 3	Waltz X Unsupervised control	13 weeks 20 sessions of 1h each 2 x per week	Postural balance measured with Berg Balance Scale (BBS) Always at the same time of day, with subjects taking medication Scores: Dance: 52 ± 3 Control: 47 ± 10.3

Fonte: Almeida H. (2020)

Metanálise dos dados extraídos

Aqui chegamos ao cerne deste livro. Dependendo do tipo de desfecho e do tipo de estudo, há diferentes formas de metanalisar (técnicas estatísticas) os dados extraídos dos estudos incluídos na metanálise (estudo, artigo científico). Sendo assim, vamos organizar esta parte “Metanálise dos dados extraídos” pelas medidas de efeito média, diferença de médias (DM), diferença de médias padronizada (DMP), risco relativo (RR), diferença de risco (DR) e *odds ratio* (OR), com um exemplo prático para cada medida de efeito.

Antes de vermos os exemplos práticos da técnica estatística de metanálise para cada medida efeito, o entendimento de tipo de desfecho e medidas de efeito, modelo de análise e heterogeneidade, pesos dos estudos e resultado da metanálise e viés de publicação é fundamental para compreensão destes exemplos.

Tipo de desfecho e medidas de efeito

Cuidado para não confundir tipo de desfecho com medida de efeito. O tipo de desfecho se refere a qualquer desfecho, como por exemplo, pressão arterial (mmHg), batimentos cardíacos (bpm), distância percorrida (m), quantidade de lesões ocorridas, proporção de pessoas doentes e quantidade de pessoas que tem plano de saúde, que podem ser contínuos ou dicotômicos. Os três primeiros desfechos são considerados contínuos, porque têm uma unidade de medida, os três últimos são considerados dicotômicos, porque são quantidades ou proporções adimensionais.

O tipo de desfecho tem uma relação direta com as medidas de efeito e com o tipo de estudo. O tipo de desfecho contínuo tem como medidas de efeito a média, diferença de médias (DM) e diferença de médias padronizada (DMP), enquanto o desfecho dicotômico tem como medidas de efeito o risco relativo (RR), a diferença de risco (DR) e a *odds ratio* (OR).

A média é uma medida de efeito obtida com tipo de estudo transversal, enquanto a diferença de médias (DM) e diferença de médias padronizada (DMP) são obtidas com tipo de estudo experimental ou de comparação de grupos. O risco relativo (RR) e a diferença de riscos (DR) são medidas de efeito obtidas com tipo de estudo coorte, no qual um grupo é exposto prospectivamente ao tratamento ou fator de risco (GE) e o outro grupo não (GNE), enquanto o *odds ratio* (OR) é obtida com tipo de estudo caso-controle, no qual um grupo de casos foi exposto retrospectivamente ao tratamento ou fator de risco (GE) e outro grupo controle não (GNE).

Para o tipo de desfecho contínuo, os dados extraídos dos estudos a serem incluídos na metanálise são a média, desvio-padrão e quantidade de participantes do (s) grupo(s), enquanto que para o tipo de desfecho dicotômico, os dados extraídos dos estudos a serem incluídos na metanálise são a quantidade de eventos e a quantidade de participantes dos grupos. Eventos são o próprio desfecho dicotômico, podendo ser a quantidade de lesões ocorridas, proporção de pessoas doentes, quantidade de pessoas que tem plano de saúde.

A média representa o resultado do grupo de participantes da pesquisa, obtida pela soma dos resultados individuais dos participantes da pesquisa dividida pela quantidade de participantes do grupo. O grau

de variabilidade dos resultados individuais do grupo de participantes da pesquisa em relação à média do grupo representa o desvio-padrão. Quanto menos diferentes da média forem estes resultados individuais, menor será o desvio-padrão. Quanto maior for a quantidade de participantes do grupo, menor será o desvio-padrão. A diferença de médias (DM) é a aritmética simples de subtração entre duas médias. A diferença de médias padronizada (DMP) é a razão entre a diferença de médias e o desvio-padrão combinado dos dois grupos. Isto será visto mais a frente.

O risco relativo (RR) é razão entre a incidência absoluta do evento no GE e a incidência absoluta do evento no GNE. A diferença de risco (DR) é, obviamente, a diferença entre estes dois riscos absolutos do GE e GNE. A *odds ratio* (OR), por ser uma medida de efeito típica do estudo retrospectivo caso-controle, não pode obter as incidências do evento, mas pode obter a razão entre a chance do evento no GE e a chance do evento no GNE. Risco relativo (RR), diferença de riscos (DR) e *odds ratio* (OR) não têm desvios-padrão por razões óbvias, mas têm erro padrão (EP).

Erro padrão (EP) representa o grau de variabilidade das médias, diferenças de médias (DM), diferenças de médias padronizadas (DMP), riscos relativos (RR), diferenças de riscos (DR) e *odds ratios* (OR). Assim como há variabilidade dos resultados individuais do grupo de participantes da pesquisa, há também variabilidade das médias, diferenças de médias (DM), diferenças de médias padronizadas (DMP), riscos relativos (RR), diferenças de riscos (DR) e *odds ratios* (OR), caso diversos grupos de participantes da pesquisa fossem selecionados da mesma população. Esta variabilidade das médias, diferenças de médias (DM), diferenças de médias padronizadas (DMP), riscos relativos (RR), diferenças de riscos (DR) e *odds ratios* (OR), conhecida também como erro amostral, pode ser estimada a partir do desvio-padrão (DP) e quantidade de participantes do(s) grupo(s) no caso dos desfechos contínuos ou quantidade de eventos e quantidade de participantes dos GE e GNE no caso de desfechos dicotômicos. Quanto menor o desvio-padrão do(s) grupo(s) e/ou quanto maior a quantidade de participantes do(s) grupo(s), menor será o erro padrão para os desfechos contínuos. Para os desfechos dicotômicos, quanto maior a quantidade de eventos e de participantes nos grupos, menor será o erro padrão. Isto vale tanto para a medida de efeito do estudo quanto para a medida de efeito da metanálise.

Intervalo de confiança 95% representa os limites prováveis das médias, diferenças de médias (DM), diferenças de médias padronizadas (DMP), riscos relativos (RR), diferenças de riscos (DR) e *odds ratios* (OR), que poderiam ocorrer para uma certeza de 95%. Estes limites são obtidos a partir da diferença entre a medida de efeito e o erro padrão desta medida de efeito: o limite inferior é igual a medida de efeito menos 1,96 vezes o erro padrão desta medida de efeito; o limite superior é igual a medida de efeito mais 1,96 vezes o erro padrão desta medida de efeito. 1,96 é o escore Z que garante 95% das medidas de efeito entre estes limites. Importante saber que há intervalos para quaisquer outros níveis de confiança, como 60%, 75%, 95%, 99%. A metanálise, bem como a ciência como um todo, utiliza o intervalo de confiança 95%. Outro ponto importante é que há necessidade que o desfecho considerado no estudo tenha uma distribuição normal dos resultados.

Há diferentes formas de explicar uma distribuição normal. Da mais complexa, que é a fórmula matemática, a mais simples que é esta aqui: um desfecho apresentará distribuição normal quando a maioria das medidas de efeito estão próximas da média destas medidas, tanto acima como abaixo dela, poucas medidas de efeito estão afastadas da média destas medidas, tanto acima como abaixo dela, e a minoria das medidas de efeito estão muito afastadas da média destas medidas, tanto acima como abaixo dela. Como as medidas de efeito risco relativo (RR) e *odds ratio* (OR) não apresentam distribuição normal, a função logarítmica (ln) é utilizada para obtenção de uma distribuição aproximadamente normal para o cálculo do intervalo de confiança 95% dos riscos relativos (RR) e *odds ratios* (OR). Em seguida, a função exponencial (Exp) converte o intervalo de confiança 95% dos riscos relativos (RR) e *odds ratios* (OR) para os valores originais.

As metanálises de desfechos contínuos são mais difíceis de serem realizadas, porque muitas vezes os estudos ao invés de informar a média e o desvio-padrão dos grupos, informam a diferença de médias e/ou erro padrão. Não que seja errado informar a diferença de médias e o erro padrão, mas isto prejudica a realização de metanálises. Caso o estudo informe a média e o erro padrão da média, pode-se calcular o desvio-padrão, pela raiz quadrada da quantidade de participantes do grupo multiplicada pelo erro padrão da média (desvio-padrão = erro padrão da média * raiz [quantidade de participantes do grupo]).

Modelo de análise e heterogeneidade

Há dois modelos para metanalisar os dados extraídos dos estudos: de efeito fixo e de efeitos randômicos, também chamado de efeitos aleatórios. No modelo de efeito fixo assume-se que o resultado é comum em todos os estudos, que os estudos são semelhantes e que as pequenas diferenças entre os resultados dos estudos são provenientes do erro amostral, ou seja, das diferenças de características dos participantes. Por outro lado, no modelo de efeitos randômicos o resultado varia entre os estudos incluídos, provavelmente devido a diferenças entre as características dos participantes e/ou dos métodos empregados em cada estudo. Essa variabilidade entre os resultados dos estudos incluídos, chamada de heterogeneidade, deve ser incorporada na atribuição dos pesos aos estudos, no resultado da metanálise e explicada suas prováveis causas.

Há duas formas de verificar a existência da heterogeneidade: observando e quantificando. Considerando que os prováveis resultados dos estudos e da metanálise são aqueles contidos em seus respectivos intervalos de confiança 95%, no modelo de efeito fixo observa-se que o resultado da metanálise, linha vertical vermelha, está contido nos resultados de todos os estudos, enquanto que no modelo de efeitos randômicos, o resultado da metanálise não está contido em três estudos (Figura 10).

A quantificação da heterogeneidade da medida de efeito entre os estudos depende do método estatístico utilizado e pode ser feita pelo teste de Cochran (1954). No método da variância inversa (IV), a heterogeneidade é obtida pelo somatório dos produtos dos pesos dos estudos pelos quadrados das diferenças entre as medidas de efeito dos estudos e medida de efeito da metanálise. No método de Mantel-Haenszel (MH), a heterogeneidade é obtida de maneira semelhante, mas os pesos dos estudos passam a ser o inverso do quadrado dos erros padrão dos estudos e para os desfechos risco relativo (RR) e *odds ratio* (OR) são utilizadas as funções logarítmicas e exponenciais para as medidas de efeito dos estudos e medida de efeito da metanálise.

Quando a quantidade de estudos é pequena, o teste de Cochran pode não detectar heterogeneidade mesmo existindo grande variabilidade entre os resultados dos estudos, ou ainda, quando a quantidade de estudos é muito grande pode detectar alta heterogeneidade mesmo existindo pouca variabilidade entre os

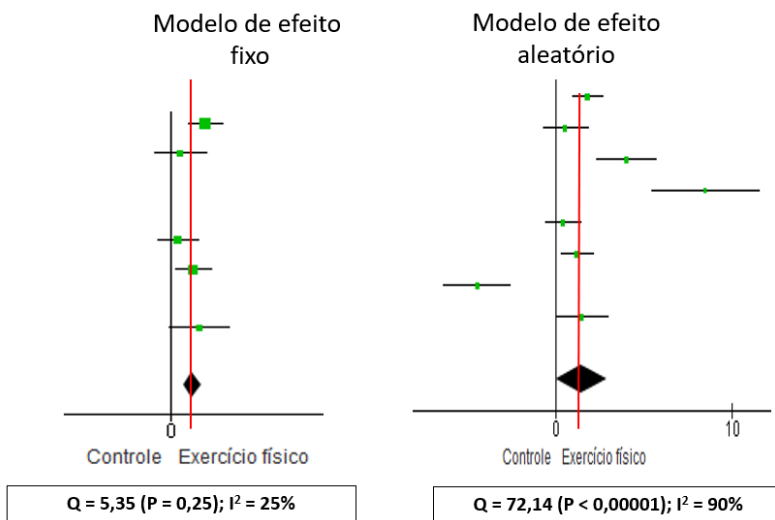
resultados dos estudos. Por isso, é recomendável considerar também o teste de Higgins (2002) que quantifica, a partir do resultado do teste de Cochran, a inconsistência (I^2) total relativa à heterogeneidade e não ao erro amostral.

Valores de I^2 próximos de 0% indicam ausência de heterogeneidade, próximos de 25% indicam baixa heterogeneidade, próximos de 50% indicam moderada heterogeneidade e próximos de 75% indicam alta heterogeneidade.

Há uma discussão científica interessante sobre qual teste utilizar para quantificar a heterogeneidade. O Handbook da Cochrane aconselhou a não escolher o modelo de análise, se fixo ou de efeitos randômicos, a partir do teste de heterogeneidade. Higgins et al., (2003) disseram que o I^2 é preferível para testar a heterogeneidade. Ora, com todo respeito a Higgins et al., (2003) há um conflito de interesses nessa preferência, além do que, o I^2 se baseia no teste de Cochran. Sedgwick (2015) sugere utilizar o nível de significância de 10% ao invés de 5%, como forma de atenuar a pouca validade do teste de Cochran.

Bem, a minha sugestão sobre essa discussão é que você reflita sobre ela e que considere nesta reflexão como eu trabalho com a heterogeneidade na metanálise: unindo a observação gráfica, teste de Cochran 5% e I^2 . Na Figura 10 você poderá ver tudo isso. Na metanálise da esquerda, a heterogeneidade foi igual a 5,35, a probabilidade de ter sido por acaso foi de 0,25, alta esta probabilidade, então foi por acaso e a I^2 foi igual a 25%, indicando baixa heterogeneidade. Neste caso, o modelo de análise de efeito fixo é o indicado. Na metanálise da direita, a heterogeneidade foi igual a 72,14, a probabilidade de ter sido por acaso foi de 0,00001, muito baixa esta probabilidade, então não foi por acaso e a I^2 foi igual a 90%, indicando alta heterogeneidade. Neste caso, o modelo de análise de efeitos randômicos é o indicado, para incorporar a heterogeneidade na atribuição dos pesos aos estudos e, conseqüentemente, ao resultado da metanálise.

Figura 10 - No modelo de efeito fixo o resultado da metanálise, linha vertical vermelha, está contido nos resultados de todos os estudos, enquanto que no modelo de efeitos randômicos, o resultado da metanálise não está contido em três estudos



Por analogia, podemos comparar a heterogeneidade da metanálise à confiabilidade da medida entre dois ou mais avaliadores. Se dois ou mais avaliadores, utilizaram o mesmo material para medir o mesmo desfecho no mesmo grupo de participantes do estudo, é de se esperar que as médias e variâncias dos desfechos entre esses avaliadores sejam as mesmas, mas muitas vezes não são. Os estudos de confiabilidade da medida têm este desenho para medir o erro de uma medida entre avaliadores. Algo semelhante ocorre com as metanálises: se os estudos incluídos na metanálise tiveram o mesmo objetivo, é de se esperar que os resultados desses estudos sejam os mesmos, mas nem sempre são. Apesar do objetivo ter sido o mesmo entre os estudos, as diferenças entre as amostras e entre os métodos dos estudos geram uma heterogeneidade que pode ser por acaso ou verdadeira, baixa ou alta. Quando verdadeira e/ou alta, a heterogeneidade se transforma em um erro, que deve ser inserido na metanálise para que haja correção da medida de efeito e do intervalo de confiança 95% da metanálise.

Caso a heterogeneidade da medida de efeito entre os estudos seja verdadeira e/ou alta é conveniente que seja explicada. Há três

diferentes maneiras para identificar a causa da heterogeneidade: exclusão dos estudos *outliers*, análise de subgrupo e metaregressão.

A exclusão de *outliers* consiste em retirar da análise estudos destoantes, responsáveis pela heterogeneidade observada. A retirada desses estudos idealmente deve ser guiada por uma justificativa de cunho clínico ou metodológico como, por exemplo, ter avaliado uma população com características diferentes dos demais estudos ou ter utilizado diferentes critérios de avaliação para o desfecho de interesse. Jamais a exclusão de *outliers* irá ocorrer se o planejamento da metanálise foi detalhado e cuidadoso estabelecendo critérios de inclusão dos estudos adequados e coerentes com o objetivo da metanálise.

A análise de subgrupos consiste em dividir a amostra em dois ou mais grupos e avaliar separadamente cada um desses subgrupos. A separação desses grupos pode ser feita a partir de variáveis clínicas (por exemplo, gravidade da doença) ou metodológica (por exemplo, qualidade dos estudos). A análise de subgrupos tenta explicar a heterogeneidade a partir de variáveis categóricas.

A análise de meta-regressão é semelhante à análise de regressão linear. Diferente da análise de subgrupos, se propõe a explicar a heterogeneidade a partir de variáveis contínuas como, por exemplo, a dose de um medicamento ou tempo de seguimento do estudo.

Muito importante considerar a recomendação de Borenstein et al. (2017): I^2 nada tem a ver com o tamanho da medida do efeito. I^2 descreve o erro entre os estudos (variância observada), excluindo o erro amostral. O erro entre os estudos decorrentes das diferenças nos MÉTODOS empregados, apesar dos objetivos dos estudos terem sido os mesmos.

Se você quiser considerar o tamanho da medida do efeito, calcule o tamanho do efeito (eles recomendam o *d de Cohen*) menor e maior entre os estudos. Eu prefiro interpretar o tamanho da medida do efeito pelo IC95% da metanálise.

Pesos dos estudos e resultado da metanálise

Uma forma simplória de obter um resultado da metanálise seria calcular a média aritmética dos resultados dos estudos incluídos. Porém, esses estudos apresentam resultados com diferentes precisão e confiança. Seria um erro considerar igualmente o resultado de cada

estudo para o resultado da metanálise. A técnica estatística metanálise apresentou uma solução magistral para isto: atribuir pesos aos estudos de acordo com suas qualidades metodológicas. O estudo de melhor qualidade metodológica deveria pesar mais para o resultado da metanálise.

O peso de cada estudo depende do modelo de análise, efeito fixo ou efeitos randômicos, e da medida de efeito. Para cada combinação de modelo análise com a medida de efeito, a forma de medir o peso dos estudos e, conseqüentemente, o resultado da metanálise é diferente.

Como saber a qualidade metodológica dos estudos a partir das médias, desvios-padrão, quantidade de eventos e quantidade de participantes dos grupos? Pela variância dos estudos. Estudos com maiores quantidades de participantes, maiores proporções de eventos no grupo exposto (GE), menores proporções de eventos no grupo não exposto (GNE), adequados critérios de inclusão dos participantes, formação aleatória dos grupos, grupo controle, cegamento e pouca perda amostral são estudos de boa qualidade metodológica e, por isso, apresentam em seus resultados desvios-padrão, erros padrão, e intervalos de confiança 95% menores, conseqüentemente variâncias menores.

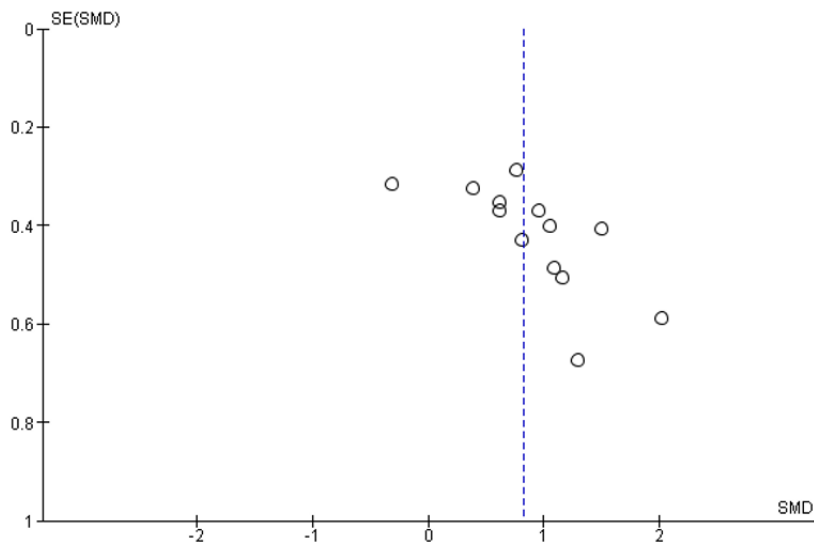
Por isso, o resultado da metanálise atribui pesos proporcionais ao intervalo de confiança 95% de cada estudo. Quanto menor o intervalo de confiança 95%, maior o peso do estudo. Na Figura 21 a quantidade de participantes no estudo Roman et al., (2016) foi bem maior que a do estudo de Chen et al., (2020) e os desvios-padrão do estudo Roman et al., (2016) foram bem menores que do estudo de Chen et al., (2020). Conseqüentemente, a amplitude do intervalo de confiança 95% do estudo de Roman et al., (2016) foi de 5,78 (0,89 – (-4,89)), enquanto que o de Chen et al., (2020) foi de 15,58 (9,79 – (-5,79)), quase três vezes maior que o de Roman et al., (2016). Os pesos atribuídos aos estudos de Roman et al., (2016) e Chen et al., (2020) foram de 47,6% e 6,5%, respectivamente. O estudo de Roman et al., (2016) contribuiu sete vezes mais que o estudo de Chen et al., (2020) para o resultado da metanálise.

Viés de publicação

A Figura 11 abaixo representa o *funnel plot* dos estudos incluídos na metanálise do controle postural de pessoas com doença de Parkinson que realizaram danças regulares versus controle (Figura 22). O *funnel plot*

é um gráfico que avalia subjetivamente a suspeição de viés de publicação, ou seja, a não inclusão na metanálise de algum estudo, principalmente aqueles que não tenham sido publicados porque a quantidade de participantes fora pequena.

Figura 11 – Viés de publicação da metanálise sobre do controle postural de pessoas com doença de Parkinson que realizaram danças regulares versus controle



Egger test: bias = 8,77 (95% CI = 5,84 to 11,69); P < 0,0001
Objetivo = erro

Fonte: Almeida H. (2020)

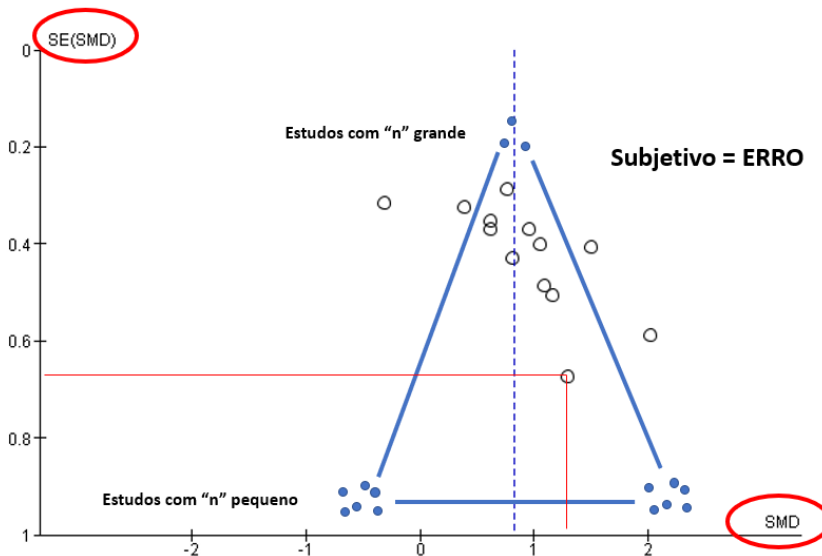
A Figura 12 é mesma Figura 11, mas com a inclusão de estudos fictícios para melhorar a explicação sobre o *funnel plot*. Como dito, o *funnel plot* é um gráfico cujo o eixo Y invertido contém os erros padrão da medida de efeito e o eixo X contém as medidas de efeito considerados nos estudos. A linha vertical pontilhada representa a medida de efeito resultante da metanálise. Cada pequeno círculo representa o estudo incluído, que é plotado no gráfico de acordo com sua medida de efeito e erro padrão desta medida de efeito.

O *funnel plot* se baseia em dois pressupostos: (1) A quantidade de estudos com pequena quantidade de participantes (“n” pequeno) é bem maior do que estudos com grande quantidade de participantes (“n”

grande); e (2) Estudos com pequena quantidade de participantes (“n” pequeno) são mais imprecisos. Conseqüentemente, estudos com “n” pequeno apresentam erros padrão da medida de efeito maiores, portanto, estariam plotados na parte inferior do eixo Y, equidistantes do resultado da metanálise por serem imprecisos. Já estudos com “n” grande apresentam erros padrão da medida de efeito menores, portanto, estariam plotados na parte superior do eixo Y, próximos do resultado da metanálise por serem precisos (triângulo azul da Figura 12 criado pelo autor = *funnel plot*).

Quando o gráfico apresenta este triângulo, não há viés de publicação. Devido a esta subjetividade que pode gerar um erro maior na avaliação do viés de publicação, outros testes mais objetivos, como o de Egger para tipo de desfecho contínuo e o de Begg para tipo de desfecho dicotômico, podem ser utilizados. No exemplo dado, o teste de Egger mostrou que a probabilidade do viés de publicação ser por acaso foi menor que 0,0001, confirmando a suspeição de viés de publicação detectado pelo *funnel plot*. Quando houver divergência de resultados entre o *funnel plot* e o teste de Egger ou Begg, opte pelos dois últimos.

Figura 12 – Criação do autor para explicar o *funnel plot*



Egger test: bias = 8,77 (95% CI = 5,84 to 11,69); P < 0,0001
Objetivo = erro

Média

O JAMOVI, disponível gratuitamente em <https://www.jamovi.org/download.html> é um bom aplicativo para metanalisar médias. Para saber como utilizar o JAMOVI o link <https://www.gpces.com.br/arquivos/970736341663865744.9471.pdf>.

O Exemplo 3 - Metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos será utilizado para facilitar a compreensão. Os dados extraídos de cada estudo incluído na metanálise foram média, desvio-padrão e quantidade de participantes do grupo (Tabela 1).

Tabela 1 – Dados extraídos dos estudos

Estudo	Média	DP	n
Clarkson et al., (1985)	48,9	3,6	13
Chmelar et al., (1989)	47	2,1	10
Guidetti et al., (2007)	46	2	12
Guidetti et al., (2008) a	38,1	1,9	13
Guidetti et al., (2008) b	46,2	2,1	12

Fonte: Matos MI. (2021)

DP = desvio-padrão do estudo; n = quantidade de participantes do estudo

No aplicativo JAMOVI, selecione o tipo de desfecho contínuo, o método estatístico de máxima verossimilhança (é o mesmo que variância inversa) e modelo de efeito fixo.

A Figura 13 é o *forest plot* da metanálise. Foram 5 estudos. O menor consumo de oxigênio foi de 38,10 ml/kg/min (estudo do Guidetti et al., (2008a), enquanto que o maior consumo de oxigênio foi de 48,90 ml/kg/min (estudo de Clarkson et al., (1985)). Os pesos dos estudos variaram de 12,31% a 23,32%. O resultado da metanálise foi de 44,77 [43,46 a 46,07] ml/kg/min. O mesmo resultado está contido na Figura 14 (Estimativa + IC95%).

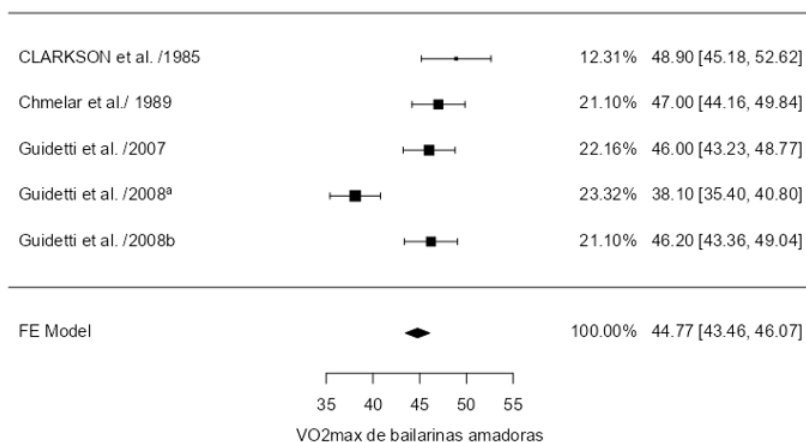
A média do consumo de oxigênio do estudo de Clarkson et al., (1985) foi igual a 48,90 ml/kg/min, resultado extraído do próprio estudo. O erro padrão da média ($EP_{\text{média estudo}}$) foi igual a raiz do desvio-padrão do estudo (DP_{estudo}). $EP_{\text{média estudo}} = \text{raiz}(3,6) = 1,8973$. O intervalo de confiança 95% desse estudo ($IC_{95\% \text{ estudo}}$) foi igual a média de consumo de oxigênio ($Média_{\text{estudo}} \pm 1,96 * = 48,90 \pm 1,96 * EP_{\text{média estudo}} = 48,90 \pm$

$1,96 * 1,8973 = [45,18 \text{ a } 52,62]$. Calcule o intervalo de confiança 95% dos outros 4 estudos.

O peso desse estudo ($\text{Peso}_{\text{estudo}}$) foi igual a $1 / \text{DP}_{\text{estudo}} = 1 / 3,6 = 0,2778$. Calcule o peso dos outros 4 estudos. Some o peso dos 5 estudos. Se a soma dos pesos foi igual 2,2565, você acertou os pesos dos outros 4 estudos. O peso percentual desse estudo ($\text{Peso}\%_{\text{estudo}}$) foi igual ao $\text{Peso}_{\text{estudo}} / \text{soma dos pesos} = 0,2778 / 2,2565 = 12,31\%$. Calcule o peso% dos outros 4 estudos.

A média da metanálise ($\text{Média}_{\text{meta}}$) foi igual à soma ($\text{Peso}\%_{\text{estudo}} * \text{Média}_{\text{estudo}}$) = $0,1231*48,90 + 0,2110*47,00 + 0,2216*46,00 + 0,2332*38,10 + 0,2110*46,20 = 44,77 \text{ ml/kg/min}$. O erro padrão da média da metanálise ($\text{EP}_{\text{média meta}}$) foi igual ao inverso da raiz da soma dos pesos dos estudos = $1 / \text{raiz}(\text{soma dos } \text{Peso}_{\text{Sestudos}}) = 1 / \text{raiz}(2,2565) = 0,6657$. O intervalo de confiança 95% da metanálise ($\text{IC } 95\%_{\text{meta}}$) foi igual a $\text{Média}_{\text{meta}} \pm 1,96 * \text{EP}_{\text{média meta}} = 44,77 \pm 1,96 * 0,6657 = 44,77 \pm 1,3048 = [43,46 \text{ a } 46,07] \text{ ml/kg/min}$.

Figura 13 – *Forest plot* da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinas; da esquerda para a direita: estudos incluídos na metanálise, resultado gráfico dos estudos (retângulo preto = peso do estudo, traço horizontal = intervalo de confiança 95% do estudo), peso percentual do estudo, resultado numérico dos estudos (média [intervalo de confiança 95%]; última linha: *FE Model* = modelo de efeito fixo, resultado gráfico da metanálise, soma dos pesos percentuais dos estudos, resultado da metanálise (média [intervalo de confiança 95%])



Fonte: Matos MI. (2021)

A Figura 14 faz parte ainda do relatório do JAMOV. Na parte superior replicou o resultado da metanálise (Estimativa = 44,8), o erro padrão da média da metanálise (EPM = 0,666), o intervalo de confiança 95% da metanálise (IC95% = 43,463 a 46,072), já mostrados no forest plot (Figura 13). Além disso, mostrou uma informação importante na parte inferior que foi a heterogeneidade do consumo médio de oxigênio entre os 5 estudos igual à Soma $(\text{Peso}_{\text{estudo}} * (\text{Média}_{\text{estudo}} - \text{Média}_{\text{meta}})^2) = 0,2778*(48,9 - 44,77)^2 + 0,4762*(47 - 44,77)^2 + 0,5*(46 - 44,77)^2 + 0,5263*(38,1 - 44,77)^2 + 0,4762*(46,2 - 44,77)^2 = Q = 32,251$, que representou uma inconsistência $I^2 = ((Q - (\text{quantidade de estudos} - 1)) / Q) * 100 = ((32,251 - (5-1)) / 32,251) * 100 = 87,6\%$ e a probabilidade de ter sido por acaso $< 0,001$.

Figura 14 – Resultado da metanálise (Estimativa), o erro padrão da média (EPM) da metanálise (EPM), o intervalo de confiança 95% da metanálise (IC95%) e a heterogeneidade do consumo médio de oxigênio entre os 5 estudos (Q, I² e P)

Tamanho do efeito e variância amostral

Modelo de efeito fixo (k = 5)

	Estimativa	EPM	Z	P	Inferior IC95%	Superior IC95%
Intercepto	44.8	0.666	67.2	< .001	43.463	46.072
.



Heterogeneidade

Tau	Tau ²	I ²	H ²	R ²	df	Q	p
0.000	0 (SE= NA)	87.6%	8.063	.	4.000	32.251	< .001

Fonte: Matos MI. (2021)

Figura 15 – Viés de publicação da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos; *Test Name* = nome do teste (no caso 3 diferentes testes, entre eles o teste de Egger; *value* = valor do teste; *P* = probabilidade do acaso

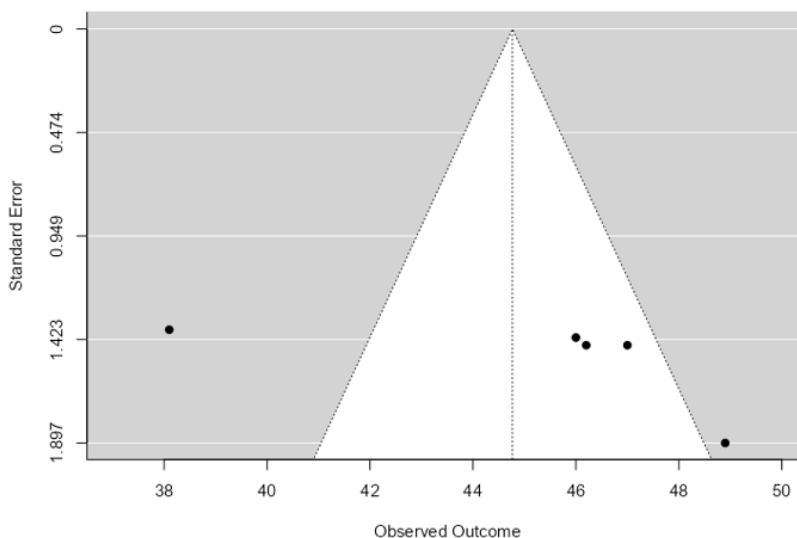
Viés de publicação

<u>Test Name</u>	<u>value</u>	<u>p</u>
<u>Fail-Safe N</u>	8340.000	< .001
<u>Kendalls Tau</u>	0.949	0.023
<u>Egger's Regression</u>	3.108	0.002

Note. Fail-safe N Calculation Using the Rosenthal Approach

Fonte: Matos MI. (2021)

Figura 16 – *Funnel plot* da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos



Fonte: Matos MI. (2021)

Continuando o relatório JAMOV, as Figuras 15 e 16 se referem ao viés de publicação. Comparando as Figuras 15 e 16, o teste de Egger, mais objetivo e com menor chance de erro, detectou que a probabilidade da suspeição de viés de publicação ter sido por acaso foi de 0,002. Muito pequena a probabilidade de ter sido por acaso, então há viés de publicação, contradizendo o resultado do *funnel plot*. Entre os dois, opte sempre pelo teste mais objetivo.

Observe que na Figura 14 a heterogeneidade foi de $Q = 32,251$, $I^2 = 87,6\%$ e $P < 0,001$ e que a metanálise foi no modelo de efeito fixo (*FE Model* – Figura 13). Então, o erro do consumo médio de oxigênio entre os estudos (heterogeneidade) não foi incorporado ao resultado da metanálise. Toda a análise da Figura 13 até a 16 está errada, porque o modelo de efeito fixo deve ser substituído pelo modelo de efeitos randômicos.

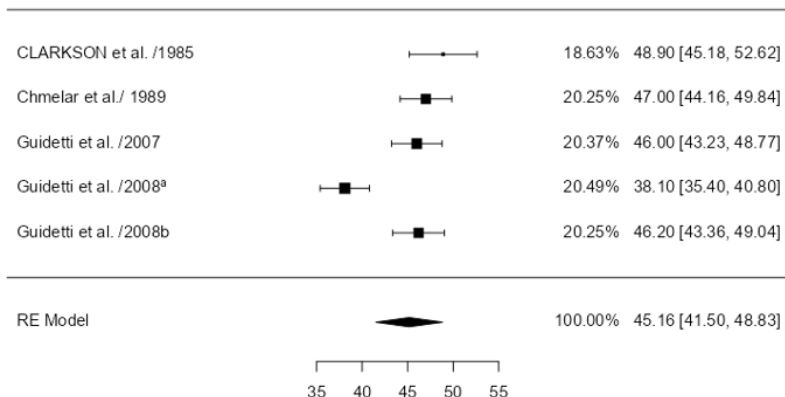
Ao incorporar a heterogeneidade do consumo médio de oxigênio entre os estudos, substituindo o modelo de efeito fixo pelo de efeitos randômicos, são corrigidos os pesos dos estudos e, conseqüentemente, o resultado da metanálise que era de 44,77 [43,46 a 46,07] passou para 45,16 [41,50 a 48,83] ml/kg/min (Figuras 17 e 18).

Ocorreu alteração também no resultado do teste de Egger que era de $P = 0,002$ (Figura 15) e passou para $P = 0,211$ (Figura 19). No modelo de análise de efeito fixo, que não havia incorporado a heterogeneidade de $Q = 32,251$ e $I^2 = 87,6\%$ ao resultado da metanálise, o teste de Egger acusou suspeição de viés de publicação (Figura 15). Quando a metanálise foi corrigida, incorporando a heterogeneidade pelo uso do modelo de efeitos randômicos (*RE Model*, Figura 17), o teste de Egger não acusou mais a suspeição de viés de publicação (Figura 19).

Qual a razão para isto ter ocorrido, se as plotagens dos estudos, os pequenos círculos pretos que representam a média e erro padrão da média de cada estudo, não alteraram? A suspeição de viés de publicação é avaliada considerando o resultado da metanálise, que foi alterado com a troca do modelo de análise.

A interpretação do resultado da metanálise de 45,16 [41,50 a 48,83] ml/kg/min (Figura 17) é que bailarinas amadoras têm um consumo médio de oxigênio igual a 45,16 ml/kg/min e que 95% delas têm consumo de oxigênio entre 41,50 e 48,83 ml/kg/min.

Figura 17 – *Forest plot* da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos



Fonte: Matos MI. (2021)

Figura 18 – Resultado da metanálise (Estimativa), o erro padrão da média da metanálise (EPM), o intervalo de confiança 95% da metanálise (IC95%) e a heterogeneidade do consumo médio de oxigênio entre os 5 estudos

Tamanho do efeito e variância amostral



Modelo de efeitos randômicos (k = 5)

	Estimativa	EPM	Z	P	Inferior IC95%	Superior IC95%
<u>Intercept</u>	45.2	1.87	24.2	< .001	41.503	48.826
.

Note. Tau² Estimator: Restricted Maximum-Likelihood



Heterogeneidade

Tau	Tau ²	I ²	H ²	R ²	df	Q	p
3.890	15.131 (SE= 12.3364)	87.12%	7.763	.	4.000	32.251	< .001

Fonte: Matos MI. (2021)

Figura 19 – Viés de publicação da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos

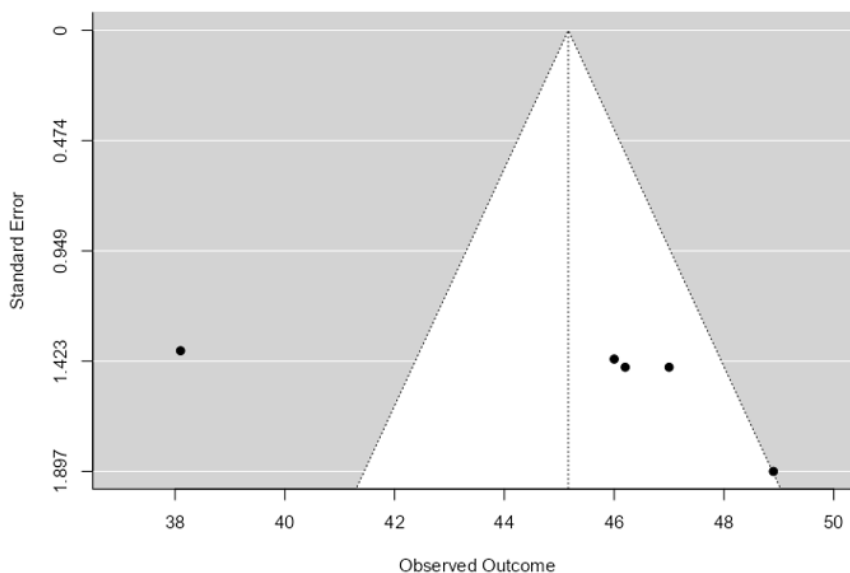
Viés de publicação

Test Name	value	p
Fail-Safe N	8340.000	< .001
Kendalls Tau	0.949	0.023
Egger's Regression	1.250	0.211

Note. Fail-safe N Calculation Using the Rosenthal Approach

Fonte: Matos MI. (2021)

Figura 20 - Funnel plot da metanálise de 05 estudos para verificar o consumo médio de oxigênio de bailarinos



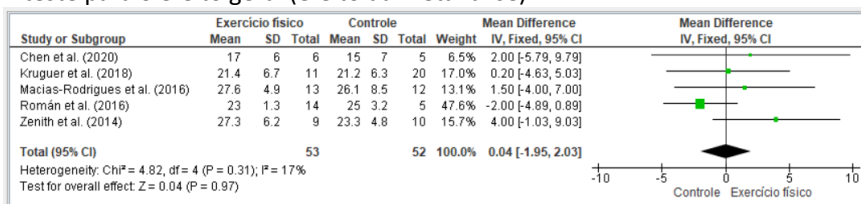
Fonte: Matos MI. (2021)

Diferença de médias (DM)

Agora, vamos considerar o Exemplo 2 - Metanálise de 05 estudos para verificar a eficácia de exercícios físicos crônicos sobre a capacidade funcional de pessoas com cirrose hepática. O aplicativo Review Manager da Colaboração Cochrane, disponível gratuitamente em <https://training.cochrane.org/online-learning/core-software-cochranereviews/revman/revman-5-download> é muito bom para fazer metanálise comparando duas médias entre dois grupos de pesquisas ou comparando duas médias intragrupo de pesquisa. Para saber como utilizar o Review Manager, acesse os links <https://www.gpces.com.br/arquivos/985974921566484955.9322.pdf> e <https://www.gpces.com.br/arquivos/563592671548367669.0211.pdf>.

Inicialmente, considere o tipo de desfecho que foi extraído dos estudos: contínuo ou dicotômico? No Exemplo 2, o desfecho foi a capacidade funcional, mensurada pelo consumo de oxigênio por ergoespirometria, em ml/kg/min. Este é um desfecho contínuo.

Figura 21 – Metanálise da capacidade funcional de pessoas com cirrose hepática após realizarem exercício físico crônico versus controle. *Study or Subgroup* = estudos incluídos na metanálise; *Mean* = média; *SD* = desvio-padrão; *Total* = quantidade de participantes; *Weight* = peso % do estudo; *Mean Difference* = diferença de médias; *IV* = método estatístico da variância inversa; *Fixed* = modelo de análise de efeito fixo; *95% CI* = intervalo de confiança 95%; *Heterogeneity* = heterogeneidade; Chi^2 = qui-quadrado = Q; *df* = graus de liberdade; *P* = probabilidade do acaso; I^2 = inconsistência; *Test for overall effect* = teste para o efeito geral (efeito da metanálise)



Fonte: Rocha FCC. (2021)

Sendo a capacidade funcional um desfecho contínuo, o método estatístico selecionado foi o da variância inversa (*IV* na Figura 21). Este método, conhecido também como método da máxima verossimilhança, atribui um peso (*Weight* na Figura 21) a cada estudo incluído, equivalente ao inverso da variância desse estudo e o intervalo de confiança da

metanálise é estimado pelo inverso da soma desses pesos atribuídos aos estudos.

Após a seleção do método estatístico da variância inversa, dois modelos de análise se apresentaram para seleção: modelo de efeito fixo ou modelo de efeitos randômicos? Como não sabemos ainda se há ou não heterogeneidade entre os estudos, o modelo de efeito fixo deve ser selecionado. Após a seleção do modelo de análise de efeito fixo, duas medidas de efeito se apresentaram para seleção: diferença de médias ou diferença de médias padronizada? A diferença de médias deve ser selecionada quando o desfecho foi medido na mesma escala em todos os estudos, enquanto que a diferença de médias padronizada quando o desfecho foi medido em escalas diferentes entre os estudos.

A diferença de médias é obtida por aritmética simples, a média de um grupo menos a média do outro grupo. O intervalo de confiança 95% desta diferença de médias é obtido pela diferença de médias $\pm 1,96$ erro padrão da diferença, sendo o erro padrão da diferença obtido a partir das quantidades de participantes e dos desvios-padrão dos dois grupos.

Já a diferença de médias padronizada é obtida pela divisão da diferença de médias pelo desvio-padrão combinado, sendo este desvio-padrão combinado obtido a partir das quantidades de participantes e dos desvios-padrão dos dois grupos. O intervalo de confiança 95% desta diferença de médias padronizada é obtido pela diferença de médias padronizada $\pm 1,96$ erro padrão da diferença padronizada, sendo o erro padrão da diferença padronizada obtido a partir das quantidades de participantes e da própria diferença de médias padronizada.

Na Figura 21 – Metanálise da capacidade funcional de pessoas com cirrose hepática, o método estatístico foi a variância inversa (IV), modelo de análise de efeito fixo (*Fixed*) e a medida de efeito foi a diferença de média (*Mean Difference*). Tomando como exemplo o primeiro estudo, de Chen et al., (2020), a média (*Mean*) da capacidade funcional do grupo exercício físico foi igual a 17, enquanto que do grupo controle foi igual 15. A diferença de média (DM_{estudo}) foi igual a $17 - 15 = 2$. No grupo exercício físico, o desvio-padrão (*SD*) foi de 6 e a quantidade de participantes (*Total*) foi de 6, enquanto que no grupo controle de 7 e 5, respectivamente. O erro padrão da diferença (EP_{DM}) foi igual a raiz da soma das razões dos quadrados dos desvios-padrão dos dois grupos pela quantidade de participantes de cada grupo. Erro padrão da diferença (EP_{DM}) igual a raiz $(6^2 / 6 + 7^2 / 5) = 3,974921$. Então, o intervalo de confiança 95% ($IC_{95\%_{\text{estudo}}}$) foi igual a $DM_{\text{estudo}} \pm 1,96 *$

$EP_{DM} = 2 \pm 1,96 * 3,974921 = [-5,79 \text{ a } 9,79]$. Repita isto para os outros 4 estudos incluídos.

O peso do estudo ($Peso_{estudo}$) de Chen et al., (2020) foi igual ao inverso de sua variância, ou seja, $1 / EP_{DM}^2 = Peso_{estudo} = 1 / 3,974921^2 = 0,063291$. Repita isto para os outros 4 estudos incluídos. Se você repetiu corretamente, a soma dos pesos dos 5 estudos foi igual a 0,968453. Dividindo o peso do estudo de Chen et al., (2020) que foi de 0,063291 pela soma dos pesos dos 5 estudos que foi de 0,968453, tem-se 6,5% (0,0651) que foi o peso percentual ($Peso\%_{estudo}$) de contribuição desse estudo para o resultado da metanálise. Calcule o percentual de contribuição dos outros 4 estudos.

A diferença de médias da metanálise (DM_{meta}) foi estimada pela soma do produto do peso percentual pela diferença de médias de cada estudo. A diferença de médias da metanálise foi igual $0,065 * 2 + 0,17 * 0,2 + 0,131 * 1,5 + 0,476 * -2 + 0,157 * 4 = 0,037541 = 0,04$ por arredondamento. O erro padrão da diferença de médias da metanálise ($EP_{DM\ meta}$) foi igual ao inverso da raiz da soma dos pesos dos estudos. $EP_{DM\ meta} = 1 / Raiz (Soma\ dos\ Pesos_{estudos}) = 1 / raiz (0,968453) = 1,02$. O intervalo de confiança 95% da metanálise ($IC\ 95\%_{meta}$) foi igual a diferença de médias da metanálise $\pm 1,96 * EP_{DM\ meta} = 0,04 \pm 1,96 * 1,02 = [-1,95 \text{ a } 2,03]$.

A interpretação do resultado da metanálise de 0,04 [-1,95 a 2,03]; $P = 0,97$ (Figura 21) foi de que a execução do exercício físico crônico não melhorou a capacidade funcional das pessoas com cirrose hepática. São três as razões para esta interpretação:

(1) O intervalo de confiança 95% [-1,95 a 2,03] contém o valor zero, ou seja, a capacidade funcional do grupo que fez exercício físico crônico pode ser a mesma do grupo controle, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise cruzou a linha de nulidade da capacidade funcional, linha vertical que passa pelo valor zero; e

(3) A probabilidade das diferenças de capacidade funcional entre os grupos ser por acaso foi alta, de 0,97.

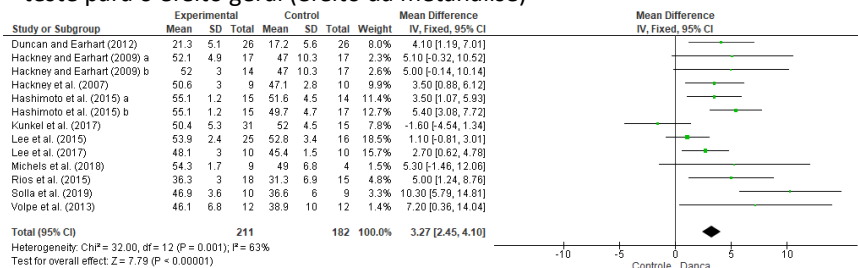
Na Figura 21, a heterogeneidade da capacidade funcional entre os estudos incluídos na metanálise foi igual à $Q = Chi^2 = Soma (Peso_{estudo} * (Efeito_{estudo} - Efeito_{metanálise})^2) = 0,063291 * (2 - 0,04)^2 + 0,164869 * (0,2 - 0,04)^2 + 0,127101 * (1,5 - 0,04)^2 + 0,461103 * (-2 - 0,04)^2 + 0,152089 * (4 -$

$0,04)^2 = 4,82$. A inconsistência I^2 foi igual a $((Q - (\text{quantidade de estudos} - 1)) / Q) * 100\% = ((4,82 - (5-1)) / 4,82) * 100\% = 17\%$, indicando baixa heterogeneidade. A probabilidade de ter sido por acaso foi de 0,31, alta esta probabilidade, então foi por acaso. Neste caso, o modelo de análise de efeito fixo foi selecionado adequadamente (*Fixed*) e a metanálise termina aí.

Diferença de médias padronizada (DMP)

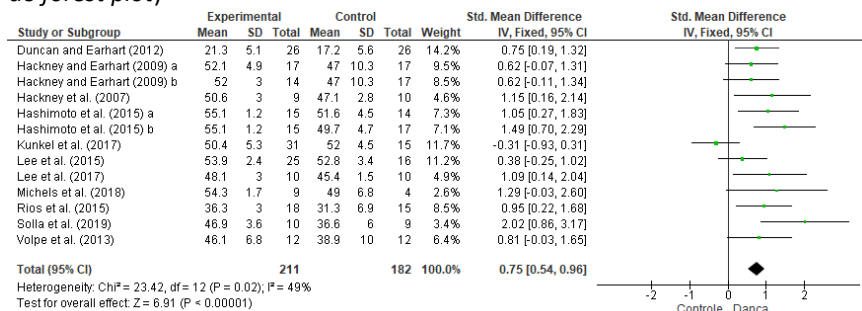
Vamos ver como fica uma metanálise ainda com tipo de desfecho contínuo, mas que apresentou heterogeneidade entre as diferenças de médias dos estudos e o desfecho foi medido por diferentes escalas entre os estudos: Exemplo 1 – Metanálise de 13 estudos para verificar a eficácia da dança sobre o controle postural de pessoas com doença de Parkinson (Figura 22 A, B e C).

A Figura 22-A é de uma metanálise do controle postural de pessoas com doença de Parkinson que realizaram danças regulares versus controle; *Study or Subgroup* = estudos incluídos na metanálise; *Mean* = média; *SD* = desvio-padrão; *Total* = quantidade de participantes; *Weight* = peso % do estudo; *Mean Difference* = diferença de médias; *IV* = método estatístico da variância inversa; *Fixed* = modelo de análise de efeito fixo; *95% CI* = intervalo de confiança 95%; *Heterogeneity* = heterogeneidade; *Chi²* = qui-quadrado = *Q*; *df* = graus de liberdade; *P* = probabilidade do acaso; I^2 = inconsistência; *Test for overall effect* = teste para o efeito geral (efeito da metanálise)



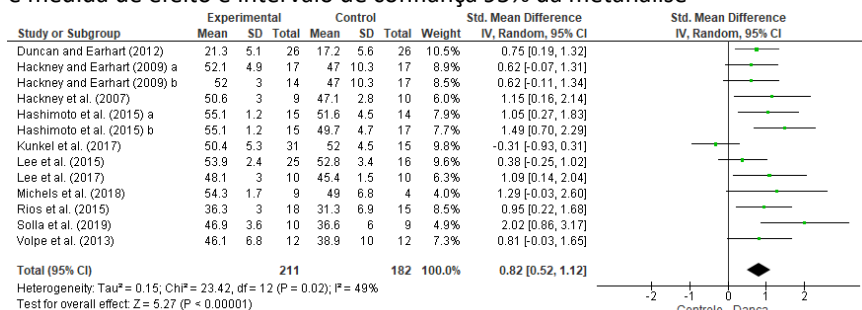
Fonte: Almeida H. (2020)

A Figura 22-B é o mesmo *forest plot* da Figura 22-A, com uma ressalva: a diferença de médias foi padronizada (*Std. Mean Difference*), para ajustar o controle postural que foi medido com escalas diferentes, *MiniBest* com 32 pontos e *Balance Berg* com 56 pontos (informação extraída dos estudos e não do *forest plot*)



Fonte: Almeida H. (2020)

A Figura 22-C é o mesmo *forest plot* da Figura 22-B, com uma ressalva: o modelo de análise foi o de efeitos randômicos (*Random*), para incorporar a heterogeneidade do controle postural entre os 13 estudos ao peso dos estudos e medida de efeito e intervalo de confiança 95% da metanálise



Fonte: Almeida H. (2020)

Na Figura 22-A foi feita a metanálise do controle postural utilizando o método da variância inversa (*IV*), modelo de efeito fixo (*Fixed*) e medida de efeito diferença de média (*Mean Difference*). Os resultados dos estudos e da metanálise foram obtidos seguindo o processo de metanalisar já descrito na Figura 21. Entretanto, observe duas informações importantes: (1) Há uma heterogeneidade moderada de $Chi^2 = Q = 32$; $I^2 = 63\%$ ($P = 0,001$); (2) O controle postural foi medido com escalas diferentes, *MiniBest* com 32 pontos e *Balance Berg* com 56 pontos (informação extraída dos estudos e não do *forest plot*). O modelo

de efeito fixo (*Fixed*) não incorporou a heterogeneidade aos pesos dos estudos e ao resultado da metanálise e a diferença de médias (*Mean Difference*) não corrigiu a diferença de escalas do *MiniBest* e *Balance Berg* na medição do controle postural.

Observe que o resultado em diferença de médias (*Mean Difference*) na Figura 22-A apontou um ganho de controle postural de 3,27 pontos para quem dançou. Porém, 3,27 pontos em qual escala? Na escala do *MiniBest* com 32 pontos, o ganho de 3,27 pontos representaria 10%, enquanto que na escala do *Balance Berg* com 56 pontos representaria 6%. Por esta razão, a medida de efeito obrigatória para corrigir a diferença de escalas na medição do controle postural foi a diferença de médias padronizada (*Std. Mean Difference* na Figura 22-B). Isto alterou as medidas de efeito dos estudos, os intervalos de confiança 95% dos estudos, os pesos dos estudos, a medida de efeito da metanálise e o intervalo da confiança 95% da metanálise e a heterogeneidade.

Tomando como exemplo o estudo de Duncan and Earhart (2012), a medida de efeito desse estudo foi obtida pela diferença de médias padronizada (DMP_{estudo}) pelo desvio padrão combinado ($DP_{\text{combinado}}$), com o ajustamento para o erro de pequenas amostras de $1 - 3/(4 * \text{total de participantes do estudo} - 9)$. O $DP_{\text{combinado}}$ foi obtido pela Raiz($((n_E - 1) * DP_E^2 + (n_C - 1) * DP_C^2 / (n_E + n_C - 2))$), sendo n_E e n_C as quantidades de participantes dos grupos experimental e controle, respectivamente. Então, $DP_{\text{combinado}} = \text{raiz}(((26-1)*5,1^2 + (26-1)*5,6^2) / (26+26-2)) = 5,3558$. A $DMP_{\text{estudo}} = ((21,3 - 17,2) / 5,3558) * (1 - 3 / (4 * 52 - 9)) = 0,7539$. O erro padrão da diferença de média padronizada (EP_{DMP}) foi igual a Raiz($((n_E + n_C) / (n_E * n_C) + DMP^2 / (2 * (n_E + n_C - 3,14)))$), sendo n_E e n_C as quantidades de participantes dos grupos experimental e controle, respectivamente. Então, $EP_{DMP} = \text{raiz}((26+26) / (26*26) + 0,7539^2 / (2*(26+26-3,14))) = 0,2876$. O intervalo de confiança 95% ($IC95\%_{\text{estudo}}$) foi igual $DMP_{\text{estudo}} \pm 1,96 * EP_{DMP} = 0,7539 \pm 1,96 * 0,2876 = [0,19 \text{ a } 1,32]$. Repita isto para os outros 12 estudos.

O peso do estudo ($\text{Peso}_{\text{estudo}}$) de Duncan and Earhart (2012) foi igual ao inverso da variância, ou seja, $1 / EP_{DMP}^2 = 1 / 0,2876^2 = 12,0859$. Repita isto para os outros 12 estudos. Se você repetiu corretamente, a soma dos pesos dos 13 estudos foi igual a 85,5576. Dividindo o peso do estudo de Duncan and Earhart (2012) que foi de 12,0859 pela soma dos pesos dos 13 estudos que foi de 85,5576, tem-se 0,1412 (14,1%) que foi o peso percentual ($\text{Peso}\%_{\text{estudo}}$) de contribuição desse estudo para o resultado da metanálise. Calcule o percentual de contribuição dos outros 12 estudos.

A diferença de médias padronizada da metanálise (DMP_{meta}) foi estimada pela soma do produto do peso percentual pela diferença de médias padronizada de cada estudo. A diferença de médias padronizada da metanálise foi igual a $0,142*0,75 + 0,095*0,62 + 0,085*0,62 + 0,046*1,15 + 0,073*1,05 + 0,071*1,49 + 0,117*(-0,31) + 0,112*0,38 + 0,049*1,09 + 0,026*1,29 + 0,085*0,95 + 0,034*2,02 + 0,064*0,81 = 0,7504$. O erro padrão da diferença de médias padronizada da metanálise ($EP_{DMP_{meta}}$) = $1 / \text{Raiz}(\text{Soma dos Pesos}_{estudos}) = 1 / \text{raiz}(85,5576) = 0,1081$. O intervalo de confiança 95% da metanálise ($IC_{95\%_{meta}}$) foi igual a diferença de médias padronizada da metanálise $\pm 1,96 * EP_{DMP_{meta}} = 0,7539 \pm 1,96 * 0,1081 = [0,54 \text{ a } 0,96]$.

Para incorporar a heterogeneidade aos pesos dos estudos e ao resultado da metanálise foi preciso permutar o modelo de análise de efeito fixo para efeitos randômicos (*Random* na Figura 22-C). Com este modelo foram corrigidos os pesos dos estudos, a medida de efeito da metanálise e o intervalo de confiança 95% da metanálise de 0,75 [0,54 a 0,96]; $P < 0,00001$ (Figura 22-B) para 0,82 [0,52 a 1,12]; $P < 0,00001$ (Figura 22-C).

A interpretação do resultado da metanálise de 0,82 [0,52 a 1,12]; $P < 0,00001$ (Figura 22-C) foi que dançar melhorou o controle postural das pessoas com doença de Parkinson. São três as razões para esta interpretação:

(1) O intervalo de confiança 95% [0,52 a 1,12] não contém o valor zero, ou seja, o controle postural do grupo que dançou foi diferente e maior do que a do grupo controle, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise não cruzou a linha de nulidade do controle postural, linha vertical que passa pelo valor zero; e

(3) A probabilidade das diferenças de controle postural entre os grupos ter sido por acaso foi baixa, de 0,00001.

O Quadro 1 mostra as fórmulas para calcular os resultados dos estudos (medidas de efeitos dos estudos, intervalos de confiança 95% dos estudos e pesos dos estudos), resultado da metanálise (medida de efeito da metanálise e intervalo de confiança 95% da metanálise), heterogeneidade e inconsistência para as medidas de efeitos média, diferença de médias e diferença de médias padronizada, no modelo de efeito fixo e método estatístico da variância inversa. Por haver diferentes

maneiras de inserção da heterogeneidade dos resultados entre os estudos e devido à complexidade de suas fórmulas, o cálculo manual não foi mostrado ao se substituir o modelo de efeito fixo pelo de efeitos randômicos. Porém, mais importante que as fórmulas omitidas, é saber que a heterogeneidade dos resultados entre os estudos ao ser inserida na metanálise, não altera as medidas de efeitos e intervalos de confiança 95% (IC95%) dos estudos, mas há aumento ou diminuição dos pesos dos estudos e, conseqüentemente, alteração na medida de efeito da metanálise e aumento no IC95% da metanálise.

Quadro 1- Fórmulas de metanálises de desfecho contínuo que utilizam o método estatístico da variância inversa, modelo de efeito fixo e medida de efeito diferença de média ou diferença de média padronizada, do pacote estatístico *RevMan*

	Modelo de Efeito Fixo		
	Média	Diferença de médias (DM)	Diferença de médias padronizada (DMP)
MEfeito _{estudo} =	Média _{estudo}	DM _{estudo} = Média _E - Média _C	DMP _{estudo} = (Média _E - Média _C) / DP _{combinado} * (1 - 3 / (4*(n _E +n _C - 9))) DP _{combinado} = Raiz(((n _E -1) * DP _E ² + (n _C -1) * DP _C ² / (n _E +n _C -2)))
EP _{estudo} =	EP _{média estudo} = Raiz (DP _{estudo})	EP _{DM} = Raiz (DP _E ² /n _E + DP _C ² /n _C)	EP _{DMP} = Raiz((n _E +n _C) / (n _E *n _C) + DMP ² / (2*(n _E +n _C -3,14)))
IC 95% _{estudo} =	Média _{estudo} ± 1,96 * EP _{média estudo}	DM _{estudo} ± 1,96 * EP _{DM}	DMP _{estudo} ± 1,96 * EP _{DMP}
Peso _{estudo} =	1 / DP _{estudo}	1/ EP _{DM} ²	1/ EP _{DMP} ²
Peso% _{estudo} =	Peso _{estudo} / Soma Pesos _{estudos}	Peso _{estudo} / Soma Pesos _{estudos}	Peso _{estudo} / Soma Pesos _{estudos}

MEfeito_{meta} =	Média _{meta} = Soma (Peso% _{estudo} * Média _{estudo})	DM _{meta} = Soma (Peso% _{estudo} * DM _{estudo})	DMP _{meta} = Soma (Peso% _{estudo} * DMP _{estudo})
EP_{meta} =	EP _{média meta} = 1 / Raiz (Soma dos Pesos _{estudos})	EP _{DM meta} = 1 / Raiz (Soma dos Pesos _{estudos})	EP _{DMP meta} = 1 / Raiz (Soma dos Pesos _{estudos})
IC 95%_{meta} =	Média _{meta} ± 1,96 * EP _{média meta}	DM _{meta} ± 1,96 * EP _{DM meta}	DMP _{meta} ± 1,96 * EP _{DMP meta}
Heterogeneidade (Q) =	Soma (Peso _{estudo} * (MEfeito _{estudo} - MEfeito _{meta}) ²)		
Inconsistência (I²) =	((Q - (quantidade de estudos - 1)) / Q) * 100		

MEfeito_{estudo} = medida de efeito do estudo; IC 95%_{estudo} = intervalo de confiança 95% do estudo; Peso_{estudo} = peso do estudo; Peso%_{estudo} = peso percentual do estudo; MEfeito_{meta} = medida de efeito da metanálise; IC 95%_{meta} = intervalo de confiança 95% da metanálise; Média_{estudo} = média do estudo; EP_{média estudo} = erro padrão da média do estudo; DP_{estudo} = desvio-padrão do estudo; Soma Pesos_{estudo} = soma dos pesos dos estudos; Média_{meta} = média da metanálise; EP_{média meta} = erro padrão da média da metanálise; DM_{estudo} = diferença de médias do estudo; Média_E = média do grupo experimental; Média_C = média do grupo controle; EP_{DM} = erro padrão da diferença de médias; DP_E = desvio-padrão do grupo experimental; DP_C = desvio-padrão do grupo controle; n_E = quantidade de participantes do grupo experimental; n_C = quantidade de participantes do grupo controle; DM_{meta} = diferença de médias da metanálise; EP_{DM meta} = erro padrão da diferença de médias da metanálise; DMP_{estudo} = diferença de médias padronizada do estudo; DP_{combinado} = desvio-padrão combinado; EP_{DMP} = erro padrão da diferença médias padronizada; DMP_{meta} = diferença de médias padronizada da metanálise; EP_{DMP meta} = erro padrão da diferença de médias padronizada da metanálise

Risco relativo (RR)

Até agora, vimos as medidas de efeito contínuo média, diferença de médias e diferença de médias padronizada. Se o desfecho for dicotômico, as medidas de efeito serão o risco relativo (RR) ou diferença de risco (DR) ou *odds ratio* (OR) ou *Peto odds ratio*. Desfecho dicotômico é aquele que expressa uma quantidade ou proporção sem unidade de medida. Já foram citados exemplos anteriores, como a quantidade de lesões ocorridas, proporção de pessoas doentes, quantidade de pessoas

que tem plano de saúde. Como novos exemplos, a quantidade de infectados, a quantidade de vacinados, a quantidade de pessoas com efeitos colaterais e a quantidade de óbitos. Os desfechos dicotômicos são chamados de eventos, que podem estar presentes ou não, que podem ocorrer ou não, quando um grupo é exposto a um tratamento ou fator de risco (GE) e outro grupo não exposto (GNE).

O método estatístico a ser adotado poderá ser o da variância inversa (IV), de Mantel-Haenszel (M-H) ou de Peto. Entre os três métodos estatísticos, o de Mantel-Haenszel (M-H) é preferível para desfechos dicotômicos, porque é válido também para amostras pequenas e/ou pequenas taxas de eventos nos grupos exposto (GE) e não exposto (GNE). Quando a taxa de eventos em um dos grupos for nula, usa-se o de Peto.

O modelo de análise será de efeito fixo ou efeito randômico, dependendo da heterogeneidade encontrada. Para facilitar a compreensão, vamos continuar no *RevMan* e utilizar o Exemplo 4 - Metanálise de 17 estudos para verificar o exercício físico como fator de proteção contra lesões musculoesqueléticas (Figura 23).

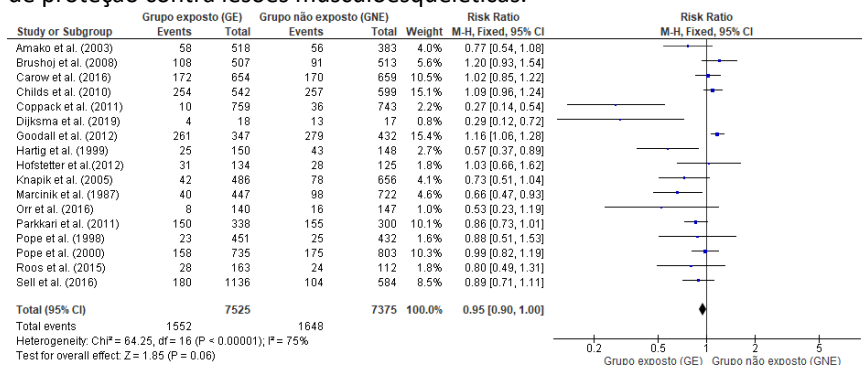
Na Figura 23 são 17 estudos, cada estudo com dois grupos: grupo exposto ao programa de exercícios físicos (GE) e grupo não exposto (GNE). De cada grupo foi extraída a quantidade de lesões ocorridas (*Events*) e a quantidade de participantes do grupo (*Total*). Considere o primeiro estudo, de Amako et al., (2003), no qual o risco absoluto de lesão do GE foi de $Events / Total = 58/518 = 0,11$ enquanto o risco absoluto de lesão do GNE foi de $Events / Total = 56/383 = 0,15$. O risco relativo do estudo (RR_{estudo}) foi a divisão dos dois riscos absolutos $= 0,11/0,15 = 0,77$. O erro padrão do RR do estudo (EP_{RR}) foi igual a raiz $(1/Events_{GE} - 1/Total_{GE} + 1/Events_{GNE} - 1/Total_{GNE}) = \text{raiz}(1/58 - 1/518 + 1/56 - 1/383) = 0,174806$. O intervalo de confiança 95% do estudo ($IC_{95\%_{estudo}}$) foi obtido pela $\text{Exp}[\ln(RR_{estudo}) \pm 1,96 * EP_{RR}] = \text{Exp}[\ln(0,77) \pm 1,96 * 0,174806] = [0,54 \text{ a } 1,08]$. A função logarítmica (\ln) foi utilizada para obter uma distribuição aproximadamente normal para o cálculo do intervalo de confiança 95% dos RR. Em seguida, a função exponencial (Exp) converteu o intervalo de confiança 95% dos RR para os valores originais.

O peso do estudo ($Peso_{estudo}$) foi obtido por $Events_{GNE} * Total_{GE} / (Total_{GE} + Total_{GNE}) = 56 * 518 / (518 + 383) = 32,19534$. O peso percentual do estudo ($Peso\%_{estudo}$) foi obtido pelo peso do estudo / soma dos pesos dos estudos $= 32,19534 / 809,37389 = 0,04$ (4%). Calcule o RR, EP_{RR} , o IC

95%_{estudo}, $\text{Peso}_{\text{estudo}}$ e $\text{Peso}\%_{\text{estudo}}$ dos outros 16 estudos. Estará certo se encontrou a soma dos pesos dos estudos igual a 809,37389.

O RR da metanálise (RR_{meta}) foi obtido pela soma ($\text{Peso}\%_{\text{estudo}} * \text{RR}_{\text{estudo}}$) = $0,04*0,77 + 0,056*1,20 + 0,105*1,02 + 0,151*1,09 + 0,022*0,27 + 0,008*0,29 + 0,154*1,16 + 0,027*0,57 + 0,018*1,03 + 0,041*0,73 + 0,046*0,66 + 0,01*0,53 + 0,101*0,86 + 0,016*0,88 + 0,103*0,99 + 0,018*0,80 + 0,085*0,89 = 0,95$. O intervalo de confiança 95% da metanálise ($\text{IC95}\%_{\text{meta}}$) foi obtido pela $\text{Exp}[\ln(\text{RR}_{\text{meta}}) \pm 1,96 * \text{EP}_{\text{RR meta}}] = \text{Exp}[\ln(0,95) \pm 1,96*0,02551] = [0,90 \text{ a } 1,00]$. O cálculo do $\text{EP}_{\text{RR meta}} = 0,02551$ não será mostrado, porque apesar de ser obtido a partir dos dados $\text{Events}_{\text{GE}}$, $\text{Events}_{\text{GNE}}$, Total_{GE} e $\text{Total}_{\text{GNE}}$, os quatro dados juntos aparecem 15 vezes na fórmula do $\text{EP}_{\text{RRmeta}}$, tornando sua extensão não viável de mostrar.

Figura 23 - Metanálise de 17 estudos para verificar o exercício físico como fator de proteção contra lesões musculoesqueléticas.



Fonte: Bunn PS. (2020)

Observe na Figura 23 que a heterogeneidade foi de $Q = 64,25$; $I^2 = 75\%$; $P < 0,00001$, ou seja, uma heterogeneidade alta e verdadeira entre os resultados de lesões dos estudos. Então, o modelo de análise de efeito fixo (*Fixed* na Figura 23) deve ser substituído pelo de efeitos randômicos, para que o resultado da metanálise que foi de 0,95 [0,90 a 1,00] seja reajustado em função desta heterogeneidade. Isto foi feito na Figura 24, o modelo de efeitos randômicos (*Random*) incorporou a heterogeneidade entre os estudos aos pesos dos estudos e resultado da metanálise, alterando de 0,95 [0,90 a 1,00] para 0,86 [0,76 a 0,98]. A metanálise compensou a inconsistência de 75% aumentando o intervalo de confiança que era de $[0,90 \text{ a } 1,00] = 0,10$ para $[0,76 \text{ a } 0,98] = 0,22$.

A interpretação do resultado da metanálise de 0,86 [0,76 a 0,98]; $P = 0,02$ (Figura 24) foi quem fez exercício físico teve um risco menor de ser acometido por lesões musculoesqueléticas, ou seja, o exercício físico foi um fator protetor contra a ocorrência de lesões musculoesqueléticas. São três as razões para esta interpretação:

(1) O intervalo de confiança 95% [0,76 a 0,98] não contém o valor 1, ou seja, as incidências absolutas prováveis de lesões musculoesqueléticas dos grupos exercício físico e controle foram diferentes, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise não cruzou a linha de nulidade da lesão musculoesquelética, linha vertical que passa pelo valor 1; e

(3) A probabilidade do risco relativo de 0,86 ter sido por acaso foi baixa, de 0,02.

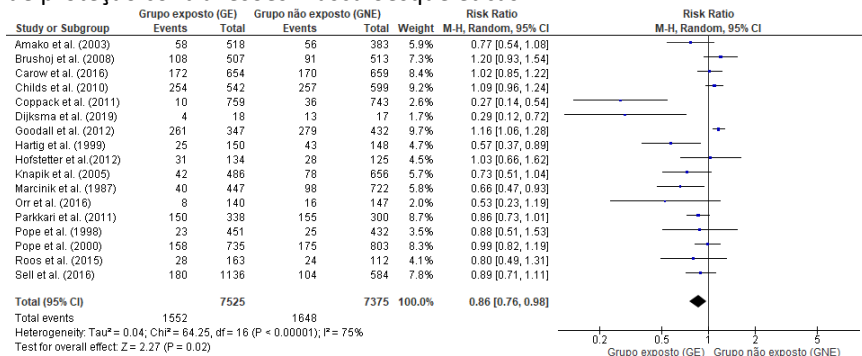
A seleção errada do modelo de análise pode levar a conclusões erradas. Se fosse mantido o resultado da metanálise de 0,95 [0,90 a 1,00]; $P = 0,06$, aquele advindo do modelo de efeito fixo (Figura 23), modelo inadequado porque não incorporou a alta e verdadeira heterogeneidade de $Q = 64,25$; $I^2 = 75\%$; $P < 0,00001$ entre os estudos, a interpretação teria sido o risco de ser acometido por lesões musculoesqueléticas foi o mesmo para quem fez exercício físico e para quem não o fez, ou seja, o exercício físico não foi um fator protetor contra a ocorrência de lesões musculoesqueléticas. Seriam três as razões para esta interpretação:

(1) O intervalo de confiança 95% [0,90 a 1,00] contém o valor 1, ou seja, as incidências absolutas prováveis de lesões musculoesqueléticas dos grupos exercício físico e controle foram iguais, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise cruzou a linha de nulidade da lesão musculoesquelética, linha vertical que passa pelo valor 1; e

(3) A probabilidade do risco relativo de 0,95 ter sido por acaso foi alta, de 0,06.

Figura 24 - Metanálise de 17 estudos para verificar o exercício físico como fator de proteção contra lesões musculoesqueléticas.



Fonte: Bunn PS. (2020)

Diferença de riscos (DR)

Para a medida de efeito diferença de riscos utilizaremos o Exemplo 5 - Metanálise de 5 estudos para verificar a obesidade como fator de risco para hipertensão arterial (fictício). Na Figura 25 são 5 estudos, cada estudo com dois grupos: grupo exposto ao fator obesidade (GE) e grupo não exposto (GNE). De cada grupo foi extraída a quantidade de diagnósticos de hipertensão arterial (*Events*) e a quantidade de participantes do grupo (*Total*). Considere o primeiro estudo, de A (ano), no qual a diferença de risco do estudo (DR_{estudo}) foi obtida pela diferença entre o risco absoluto do GE e do GNE, ou seja, $Events_{GE}/Total_{GE} - Events_{GNE}/Total_{GNE} = 4/10 - 1/8 = 0,28$. O erro padrão da DR_{estudo} (EP_{DR}) foi igual a Raiz $(Events_{GE} * (Total_{GE} - Events_{GE}) / Total_{GE}^3 + Events_{GNE} * (Total_{GNE} - Events_{GNE}) / Total_{GNE}^3) = Raiz (4*(10-4) / 10^3 + 1*(8-1) / 8^3) = 0,194092$. O intervalo de confiança 95% do estudo ($IC_{95\%estudo}$) foi igual a $DR_{estudo} \pm 1,96 * EP_{DR} = 0,28 \pm 1,96 * 0,194092 = [-0,11 \text{ a } 0,66]$.

O peso do estudo ($Peso_{estudo}$) de A (ano) foi igual a $Total_{GE} * Total_{GNE} / (Total_{GE} + Total_{GNE}) = 10*8 / (10+8) = 4,444444$. O peso percentual do estudo ($Peso\%_{estudo}$) foi igual a $Peso_{estudo} / Soma\ Pesos_{estudo} = 4,444444 / 40,09995 = 11,1\%$. Calcule a DR, EP_{DR} , $IC_{95\%estudo}$, $Peso_{estudo}$ e $Peso\%_{estudo}$ dos outros 4 estudos.

A diferença de riscos da metanálise (DR_{meta}) foi obtida pela Soma $(Peso\%_{estudo} * DR_{estudo}) = 0,111*0,28 + 0,299*0,25 + 0,236*0,06 + 0,150*0,13 + 0,204*0,24 = 0,19$. O intervalo de confiança 95% da metanálise ($IC_{95\%meta}$) foi igual $DR_{meta} \pm 1,96 * EP_{DR\ meta} = 0,19 \pm 1,96 *$

0,071428571 = [0,05 a 0,33]. O cálculo do $EP_{DR\ meta} = 0,071428571$ não será mostrado, porque apesar de ser obtido a partir dos dados $Events_{GE}$, $Events_{GNE}$, $Total_{GE}$ e $Total_{GNE}$, os quatro dados juntos aparecem 20 vezes na fórmula do $EP_{DR\ meta}$, tornando sua extensão não viável de mostrar.

A heterogeneidade $Q = Soma (1/EP_{DR}^2 * (DR_{estudo} - DR_{meta})^2) = 1/0,194092^2 * (0,28 - 0,19)^2 + 1/0,119896^2 * (0,25 - 0,19)^2 + 1/0,161918^2 * (0,06 - 0,19)^2 + 1/0,175541^2 * (0,13 - 0,19)^2 + 1/0,156154^2 * (0,24 - 0,19)^2 = 1,36$. A inconsistência $I^2 = ((Q - (quantidade\ de\ estudos - 1)) / Q) * 100 = ((1,36 - (5-1)) / 1,36) * 100 = -193,248 = 0\%$ (sempre que I^2 for negativo seu resultado será zero). Como a heterogeneidade foi nula ($I^2 = 0\%$) e devido ao acaso ($P = 0,85$) não foi necessário substituir o modelo de efeito fixo pelo de efeitos randômicos.

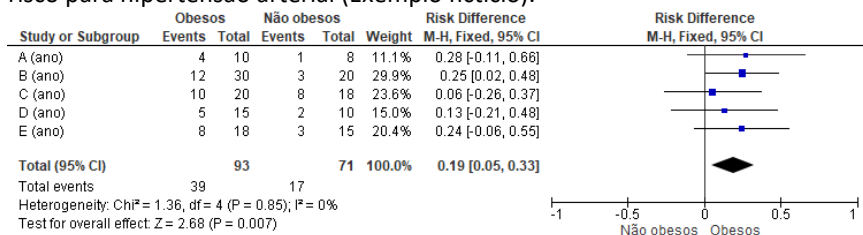
A interpretação do resultado da metanálise de 0,19 [0,05 a 0,33]; $P = 0,007$ (Figura 25) é que obesos tem 0,19 mais incidência de hipertensão arterial do que não obesos, ou seja, a obesidade é um fator de risco para hipertensão arterial. São três as razões para esta interpretação:

(1) O intervalo de confiança 95% [0,05 a 0,33] não contém o valor zero, ou seja, as incidências prováveis de hipertensão arterial dos GE e GNE não foram iguais, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise não cruzou a linha de nulidade da hipertensão arterial, linha vertical que passa pelo valor 0; e

(3) A probabilidade da diferença de risco de 0,19 ter sido por acaso foi baixa, de 0,007.

Figura 25 - Metanálise de 5 estudos para verificar a obesidade como fator de risco para hipertensão arterial (Exemplo fictício).



Odds ratio (OR)

Para a medida de efeito *odds ratio* utilizaremos o Exemplo 6 - Metanálise de 5 estudos para verificar a associação do fator “lavagem das mãos pelos médicos” e infecção (exemplo fictício). Na Figura 26 são 5 estudos, cada estudo com dois grupos: grupo exposto ao atendimento médico sem lavagem de mãos (GE) e grupo não exposto (GNE). De cada grupo foi extraída a quantidade de infecções ocorridas (*Events*) e a quantidade de participantes do grupo (*Total*).

Considere o primeiro estudo, de A (ano), no qual a *odds ratio* do estudo (OR_{estudo}) foi igual a $Events_{GE} * (Total_{GNE} - Events_{GNE}) / (Total_{GE} - Events_{GE}) * Events_{GNE} = 12 * (25 - 5) / (20 - 12) * 5 = 6,00$. O erro padrão da OR_{estudo} (EP_{OR}) foi igual a raiz $(1/Events_{GE} + 1/(Total_{GE} - Events_{GE}) + 1/Events_{GNE} + 1/(Total_{GNE} - Events_{GNE})) = Raiz (1/12 + 1/(20-12) + 1/5 + 1/(25-5)) = 0,677003$. O intervalo de confiança 95% do estudo ($IC_{95\% \text{estudo}}$) foi igual a $Exp[\ln(OR_{\text{estudo}}) \pm 1,96 * EP_{OR}] = Exp[\ln(6,00) \pm 1,96 * 0,677003] = [1,59 \text{ a } 22,62]$.

O peso do estudo ($Peso_{\text{estudo}}$) foi igual a $(Total_{GE} - Events_{GE}) * Events_{GNE} / (Total_{GE} + Total_{GNE}) = (20 - 12) * 5 / (20 + 25) = 0,888889$. O peso percentual do estudo ($Peso\%_{\text{estudo}}$) foi igual ao $Peso_{\text{estudo}} / \text{Soma Pesos}_{\text{estudo}} = 0,888889 / 4,838889 = 18,4\%$. Calcule a OR, EP_{OR} , $IC_{95\% \text{estudo}}$, $Peso_{\text{estudo}}$ e $Peso\%_{\text{estudo}}$ dos outros 4 estudos.

A *odds ratio* da metanálise (OR_{meta}) foi obtida pela Soma $(Peso\%_{\text{estudo}} * OR_{\text{estudo}}) = 0,184 * 6,00 + 0,103 * 3,89 + 0,413 * 0,50 + 0,207 * 6,00 + 0,093 * 5,44 = 3,46$. O intervalo de confiança 95% da metanálise ($IC_{95\% \text{meta}}$) foi igual a $Exp[\ln(OR_{\text{meta}}) \pm 1,96 * EP_{OR \text{ meta}}] = Exp[\ln(3,46) \pm 1,96 * 0,337924] = [1,78 \text{ a } 6,71]$. O cálculo do $EP_{OR \text{ meta}} = 0,337924$ não será mostrado, porque apesar de ser obtido a partir dos dados $Events_{GE}$, $Events_{GNE}$, $Total_{GE}$ e $Total_{GNE}$, os quatro dados juntos aparecem 43 vezes na fórmula do $EP_{OR \text{ meta}}$, tornando sua extensão não viável de mostrar.

A heterogeneidade $Q = \text{Soma } (1/EP_{OR}^2 * (\ln(Efeito_{\text{estudo}}) - \ln(Efeito_{\text{meta}}))^2) = 1/0,677003^2 * (\ln(6,00) - \ln(3,46))^2 + 1/1,004751^2 * (\ln(3,89) - \ln(3,46))^2 + 1/0,83666^2 * (\ln(0,50) - \ln(3,46))^2 + 1/0,645497^2 * (\ln(6,00) - \ln(3,46))^2 + 1/0,9759^2 * (\ln(5,44) - \ln(3,46))^2 = 6,96$. A inconsistência $I^2 = ((Q - (\text{quantidade de estudos} - 1)) / Q) * 100 = ((6,96 - (5 - 1)) / 6,96) * 100 = 43\%$. Como a heterogeneidade foi moderada ($I^2 =$

43%) e devido ao acaso ($P = 0,14$) não foi necessário substituir o modelo de efeito fixo pelo de efeitos randômicos.

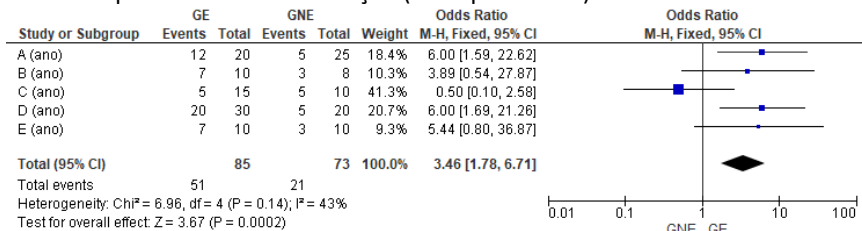
A interpretação do resultado da metanálise de 3,46 [1,78 a 6,71]; $P = 0,0002$ (Figura 26) foi que a chance de infecção dos pacientes atendidos por médicos que não lavam as mãos é 3,46 vezes maior que aqueles cujos os médicos lavam as mãos. São três as razões para esta interpretação:

(1) O intervalo de confiança 95% [1,78 a 6,71] não contém o valor 1,00, ou seja, as chances prováveis de infecções dos GE e GNE foram diferentes, com a probabilidade de 95% disto ocorrer;

(2) O losango (diamante) que representa o resultado gráfico da metanálise não cruzou a linha de nulidade da infecção, linha vertical que passa pelo valor 1; e

(3) A probabilidade da *odds ratio* de 3,46 ter sido por acaso foi baixa, de 0,0002.

Figura 26 - Metanálise de 5 estudos para verificar a associação do fator lavagem das mãos pelos médicos” e infecção (exemplo fictício)



O Quadro 2 mostra as fórmulas para calcular os resultados dos estudos (medidas de efeitos dos estudos, intervalos de confiança 95% dos estudos e pesos dos estudos), resultado da metanálise (medida de efeito da metanálise e intervalo de confiança 95% da metanálise), heterogeneidade e inconsistência para as medidas de efeitos risco relativo (RR), diferença de riscos (DR) e odds ratio (OR), no modelo de efeito fixo e método estatístico de Mantel-Haenszel (M-H). Por haver diferentes maneiras de inserção da heterogeneidade dos resultados entre os estudos e devido à complexidade de suas fórmulas, o cálculo manual não foi mostrado ao se substituir o modelo de efeito fixo pelo de efeitos randômicos. Porém, mais importante que as fórmulas omitidas, é saber que a heterogeneidade dos resultados entre os estudos ao ser inserida na metanálise, não altera as medidas de efeitos e intervalos de

confiança 95% (IC95%) dos estudos, mas há aumento ou diminuição dos pesos dos estudos e, conseqüentemente, alteração na medida de efeito da metanálise e aumento no IC95% da metanálise.

Quadro 2- Fórmulas de metanálises de desfecho dicotômico que utilizam o método estatístico Mantel-Haenszel (M-H), modelo de efeito fixo e medida de efeito risco relativo, diferença de risco ou *odds ratio* do pacote estatístico RevMan

	Modelo de Efeito Fixo		
	Risco relativo (RR)	Diferença de riscos (DR)	Odds ratio (OR)
MEfeito_{estudo} =	$RR_{\text{estudo}} = \frac{(Events_{GE}/Total_{GE})}{(Events_{GNE}/Total_{GNE})}$	$DR_{\text{estudo}} = \frac{Events_{GE}/Total_{GE} - Events_{GNE}/Total_{GNE}}$	$OR_{\text{estudo}} = \frac{(Events_{GE} * (Total_{GNE} - Events_{GNE}))}{((Total_{GE} - Events_{GE}) * Events_{GNE})}$
EP_{estudo} =	$EP_{RR} = \frac{Raiz(1/Events_{GE} - 1/Total_{GE} + 1/Events_{GNE} - 1/Total_{GNE})}{Raiz(1/Events_{GE} - 1/Total_{GE} + 1/Events_{GNE} - 1/Total_{GNE})}$	$EP_{DR} = \frac{Raiz(Events_{GE} * (Total_{GNE} - Events_{GNE}) / Total_{GE}^3 + Events_{GNE} * (Total_{GNE} - Events_{GNE}) / Total_{GNE}^3)}{Raiz(Events_{GE} * (Total_{GNE} - Events_{GNE}) / Total_{GE}^3 + Events_{GNE} * (Total_{GNE} - Events_{GNE}) / Total_{GNE}^3)}$	$EP_{OR} = \frac{Raiz(1/Events_{GE} + 1/(Total_{GE} - Events_{GE}) + 1/Events_{GNE} + 1/(Total_{GNE} - Events_{GNE}))}{Raiz(1/Events_{GE} + 1/(Total_{GE} - Events_{GE}) + 1/Events_{GNE} + 1/(Total_{GNE} - Events_{GNE}))}$
IC 95%_{estudo} =	$Exp(\ln(RR_{\text{estudo}}) \pm 1,96 * EP_{RR})$	$DR_{\text{estudo}} \pm 1,96 * EP_{DR}$	$Exp(\ln(OR_{\text{estudo}}) \pm 1,96 * EP_{OR})$
Peso_{estudo} =	$\frac{Events_{GNE} * Total_{GE}}{(Total_{GE} + Total_{GNE})}$	$\frac{Total_{GE} * Total_{GNE}}{(Total_{GE} + Total_{GNE})}$	$\frac{(Total_{GE} - Events_{GE}) * Events_{GNE}}{(Total_{GE} + Total_{GNE})}$
Peso%_{estudo} =	$Peso_{\text{estudo}} / \text{Soma Pesos}_{\text{estudo}}$	$Peso_{\text{estudo}} / \text{Soma Pesos}_{\text{estudo}}$	$Peso_{\text{estudo}} / \text{Soma Pesos}_{\text{estudo}}$
MEfeito_{meta} =	$RR_{\text{meta}} = \text{Soma } (Peso\%_{\text{estudo}} * RR_{\text{estudo}})$	$DR_{\text{meta}} = \text{Soma } (Peso\%_{\text{estudo}} * DR_{\text{estudo}})$	$OR_{\text{meta}} = \text{Soma } (Peso\%_{\text{estudo}} * OR_{\text{estudo}})$
EP_{meta} =	$EP_{RR \text{ meta}} = \text{vide (1) abaixo}$	$EP_{DR \text{ meta}} = \text{vide (1) abaixo}$	$EP_{OR \text{ meta}} = \text{vide (1) abaixo}$

IC95%_{meta} =	$\text{Exp}(\ln(\text{RR}_{\text{meta}}) \pm 1,96 * \text{EP}_{\text{RR meta}})$	$\text{DR}_{\text{meta}} \pm 1,96 * \text{EP}_{\text{DR meta}}$	$\text{Exp}(\ln(\text{OR}_{\text{meta}}) \pm 1,96 * \text{EP}_{\text{OR meta}})$
Heterogeneidade e (Q) =	$\text{Soma } (1/\text{EP}_{\text{RR}}^2 * (\ln(\text{RR}_{\text{estudo}}) - \ln(\text{RR}_{\text{meta}}))^2)$	$\text{Soma } (1/\text{EP}_{\text{DR}}^2 * (\text{DR}_{\text{estudo}} - \text{DR}_{\text{meta}})^2)$	$\text{Soma } (1/\text{EP}_{\text{OR}}^2 * (\ln(\text{OR}_{\text{estudo}}) - \ln(\text{OR}_{\text{meta}}))^2)$
Inconsistência (I²) =	$((Q - (\text{quantidade de estudos} - 1)) / Q) * 100$		

MEfeito_{estudo} = medida de efeito do estudo; IC 95%_{estudo} = intervalo de confiança 95% do estudo; Peso_{estudo} = peso do estudo; Peso%_{estudo} = peso percentual do estudo; MEfeito_{meta} = medida de efeito da metanálise; IC 95%_{meta} = intervalo de confiança 95% da metanálise; RR_{estudo} = risco relativo do estudo; Events_{GE} = quantidade de eventos no grupo exposto ao tratamento ou fator de risco; Events_{GNE} = quantidade de eventos no grupo não exposto ao tratamento ou fator de risco; Total_{GE} = total de participantes do grupo exposto ao tratamento ou fator de risco; Total_{GNE} = total de participantes do grupo não exposto ao tratamento ou fator de risco; EP_{RR} = erro padrão do RR do estudo; ln = função logarítmica utilizada para obtenção de uma distribuição aproximadamente normal para o cálculo do intervalo de confiança 95% dos RR ou OR; Exp = função exponencial para converter o intervalo de confiança 95% dos RR ou OR para os valores originais; Soma Peso_{Sestudo} = soma dos pesos dos estudos; RR_{meta} = risco relativo da metanálise; EP_{RRmeta} = erro padrão do RR da metanálise; DR_{estudo} = diferença de riscos do estudo; EP_{DR} = erro padrão da diferença de riscos do estudo; DR_{meta} = diferença de riscos da metanálise; EP_{DR meta} = erro padrão da diferença de riscos da metanálise; (1) obtido a partir de 4 dados somente: Events_{GE}, Events_{GNE}, Total_{GE} e Total_{GNE}, mas os quatro dados juntos aparecem 43 vezes na fórmula do EP_{RRmeta}, 20 vezes na fórmula do EP_{DR meta} e 15 vezes na fórmula do EP_{OR meta} que devido a sua extensão não vale a pena mostrar.

Comparando os Quadros 1 e 2, pode-se distinguir dois resultados: (a) dos estudos, composto da medida de efeito, erro padrão desta medida de efeito, intervalo de confiança 95%, peso e peso %; e (b) da metanálise, composto de medida de efeito, erro padrão desta medida de efeito e intervalo de confiança 95%. Os cálculos de cada um destes resultados podem ser gerais, que se aplicam igualmente a todas as medidas de efeito ou específicos que se aplicam diferentemente a cada uma das medidas de efeito. Por exemplo, são cálculos gerais:

(1) os intervalos de confiança 95%, tanto dos estudos como da metanálise, obtidos a partir da medida de efeito $\pm 1,96 * \text{erro padrão}$ da medida de efeito, com a atenção para as medidas de efeito risco relativo (RR) e *odds ratio* (OR), que utilizam a função logarítmica para obter uma

distribuição dos riscos relativo (RR) e *odds ratio* (OR) aproximada da normal e a exponencial para converter aos valores iniciais de riscos relativo (RR) e *odds ratio* (OR);

(2) o peso % do estudo, obtido a partir da razão do peso do estudo pela soma dos pesos dos estudos;

(3) a medida de efeito da metanálise, obtida pela soma do produto dos pesos % dos estudos pelas medidas de efeitos dos estudos;

(4) a heterogeneidade, obtida a partir da soma do produto dos pesos dos estudos pelo quadrado das diferenças entre as medidas de efeito dos estudos e medida de efeito da metanálise, com atenção para as medidas de efeito dicotômicas nas quais o peso passa a ser o inverso do quadrado do erro padrão da medida de efeito e para o risco relativo (RR) e *odds ratio* (OR), nos quais são utilizadas a função logarítmica para obter uma distribuição das medidas de efeitos dos estudos e medida de efeito da metanálise aproximada da normal e a exponencial para converter aos valores iniciais das medidas de efeitos dos estudos e medida de efeito da metanálise;

(5) a inconsistência, obtida a partir da heterogeneidade e quantidade de estudos.

Como cálculos específicos, podem ser exemplificados:

(1) as medidas de efeito dos estudos

(2) os erros padrão dos estudos e da metanálise

(3) os pesos dos estudos

No caso de desfechos contínuos, observe como a variabilidade dos resultados e a quantidade de participantes dos estudos são muito importantes. Quanto menor os desvios-padrão e maior as quantidades de participantes nos estudos, menor serão os erros padrão dos estudos, conseqüentemente maior serão os pesos dos estudos, menor será o erro padrão da metanálise e menor será o IC95% da metanálise. No caso de desfechos dicotômicos, observe que quanto maior a quantidade de eventos no GE e não-eventos no GNE e maior a quantidade de participantes nos GE e GNE, menores serão os erros padrão dos estudos, conseqüentemente maiores serão os pesos dos estudos, menor será o erro padrão da metanálise e menor será o IC95% da metanálise.

Cronologicamente, podemos definir os seguintes passos para a metanálise dos dados extraídos dos estudos:

(1) definição do tipo de desfecho: se contínuo ou dicotômico. Sendo contínuo, as medidas de efeito poderão ser média, diferença de médias (DM) ou diferença de médias padronizada (DMP). Sendo dicotômico, as medidas de efeito poderão ser risco relativo (RR), diferença de riscos (DR) ou *odds ratio* (OR);

(2) seleção do método estatístico: inverso da variância (IV), quando os desfechos forem contínuos ou Mantel-Haenszel (M-H), quando os desfechos forem dicotômicos;

(3) seleção da medida de efeito;

(4) caso esta medida de efeito tenha sido a diferença de médias e a escala ou unidade da medida tenha sido diferente entre os estudos, selecionar a diferença de médias padronizada;

(5) seleção do modelo de análise de efeito fixo;

(6) na presença de heterogeneidade alta e significativa, o modelo de análise de efeito fixo deve ser substituído pelo de efeitos randômicos, que incorporará o efeito desta heterogeneidade no resultado da metanálise, aumentando ou diminuindo o peso de cada estudo, alterando a medida de efeito da metanálise e aumentando o intervalo de confiança 95% da metanálise.

A linha de nulidade que torna sem efeito o resultado da metanálise pode ser zero ou 1, dependendo da medida de efeito considerada. Quando a medida de efeito for uma diferença de efeitos entre os grupos, a linha de nulidade será igual a zero. Enquadram-se neste caso a diferença de médias (DM), diferença de médias padronizada (DMP) e diferença de riscos (DR) (Figuras 21, 22-C e 25, respectivamente). Há uma razão para isto, a linha de nulidade igual a zero significa que os efeitos foram iguais entre os grupos, com 95% de probabilidade de isto ocorrer. Quando a medida de efeito for uma razão entre os efeitos dos grupos, a linha de nulidade será igual a 1. Enquadram-se neste caso o risco relativo (RR) e o *odds ratio* (OR) (Figuras 24 e 26, respectivamente). A razão é a mesma, a linha de nulidade igual a 1 significa que os efeitos foram iguais entre os grupos, com 95% de probabilidade de isto ocorrer.

Avaliação do nível de evidência da metanálise

Grading of Recommendations, Assessment, Development and Evaluations (GRADE) é um sistema desenvolvido com objetivo de criar uma avaliação universal, transparente e sensível para graduar a

qualidade das evidências e a força das recomendações de uma metanálise. O nível de evidência representa a confiança na informação utilizada em apoio a uma determinada recomendação, que deve ser avaliado para cada desfecho analisado.

Na GRADE, a qualidade da evidência é classificada em quatro níveis: alto, moderado, baixo, muito baixo. Esses níveis representam a confiança que possuímos na estimativa dos resultados apresentados na metanálise. A classificação inicial da qualidade da evidência é definida a partir do delineamento dos estudos. Estudos experimentais são delineamentos de estudo mais adequados para questões relacionadas à intervenção, eles começam com qualidade da evidência alta. Para estudos observacionais, a qualidade da evidência inicia como baixa.

Os fatores responsáveis pela redução no nível de evidência dos estudos experimentais são limitações metodológicas (risco de viés), inconsistência, evidência indireta, imprecisão e viés de publicação. Os fatores responsáveis pelo aumento no nível de evidência dos estudos observacionais são elevada magnitude de efeito, controle dos fatores de confusões residuais que aumentam a confiança na estimativa e gradiente dose-resposta.

A criação da GRADE para avaliar o nível de evidência das metanálises foi uma iniciativa fundamental e muito importante para a ciência baseada em evidências, mas há dois pontos que deveriam ser melhor esclarecidos: (1) a subjetividade para a classificação de cada um dos fatores pela redução ou aumento do nível de evidência; (2) a inconsistência da metanálise como fator de redução do nível de evidência; e (3) a suspeição de viés de publicação como fator de redução do nível de evidência.

Para cada um dos cinco fatores que reduzem o nível de evidência a GRADE prevê três classificações: não sério, sério e muito sério ou não grave, grave ou muito grave. Não sério ou não grave não altera o nível de evidência, sério ou grave reduz um nível de evidência, muito sério ou muito grave reduz dois níveis de evidência. A GRADE ao deixar a critério do autor classificar em não sério, sério e muito sério ou não grave, grave ou muito grave as limitações metodológicas (risco de viés), inconsistência, evidência indireta, imprecisão e viés de publicação cria uma subjetividade, que faz com que a mesma metanálise receba diferentes classificações do seu nível de evidência, quando avaliada por diferentes pesquisadores que utilizaram o mesmo instrumento de

classificação, no caso a GRADE. A classificação GRADE ficaria mais objetiva adotando a sugestão do Quadro 3.

Quadro 3 – Classificação GRADE com mais objetividade

Fatores	Classificação	Objetividade
Risco de viés	Menos um nível	Quando a quantidade de estudos com risco alto for $\geq 25\%$
	Menos dois níveis	Quando a quantidade de estudos com risco alto for $\geq 50\%$
Inconsistência	Menos um nível	Quando $\geq 50\%$
	Menos dois níveis	Quando $\geq 75\%$
Evidência indireta	Menos dois níveis	Quando qualquer um dos critérios de inclusão dos estudos estiverem incoerentes com o objetivo da metanálise
Imprecisão	Menos um nível	Quando a quantidade de participantes da metanálise ≤ 138
	Menos dois níveis	Quando a quantidade de participantes da metanálise ≤ 98
Viés de publicação	Menos um nível	Sempre que houver suspeição
Grande magnitude de efeito*	Mais dois níveis	Quando $RR \geq 5$ ou $\leq 0,2$
	Mais um nível	Quando o $RR \geq 2$ ou $\leq 0,5$
Gradiente dose-resposta	Mais um nível	Quando o coeficiente de associação entre a dose-resposta for superior a $\pm 7,0$
Variáveis de confusão	Mais um nível	Quando o autor deixar claro que controlou os principais fatores de confusão que poderiam interferir no resultado final.

* Mesma classificação sugerida pela GRADE

A GRADE considera a inconsistência da metanálise como fator de redução no nível de evidência, mas quando há inconsistência e o pesquisador seleciona o modelo de efeitos randômicos, esta inconsistência é incorporada na análise dos dados, havendo uma alteração nos pesos dos estudos, com consequente alteração na medida de efeito da metanálise e aumento no intervalo de confiança 95% da metanálise. Nesta situação, a inconsistência deixaria de ser um fator de redução? Acho que sim. Sugiro quando isto ocorrer que a inconsistência

seja informada, mas atribuída a ela a classificação não séria ou não grave, porque foi incorporada ao resultado da metanálise.

A GRADE considera também o viés de publicação como fator de redução no nível de evidência da metanálise. Porém, os próprios criadores da GRADE reconhecem a dificuldade para avaliação do viés de publicação. Tanto é que eles orientaram para se usar o eufemismo de “há suspeição” ou “não foi detectado” para a presença ou ausência de viés de publicação, respectivamente.

A ausência de estudos no rol dos estudos incluídos na metanálise é o que caracteriza o viés de publicação. Vamos supor duas situações que podem ocorrer, quando se faz uma revisão sistemática para futura metanálise dos dados extraídos dos estudos: (1) uma pequena quantidade de estudos incluídos; e (2) uma grande quantidade de estudos incluídos. Na primeira suposição ficará claro a “suspeição de viés de publicação”, porque outros estudos futuros que serão publicados poderão alterar o resultado da metanálise obtido com poucos estudos. Entretanto, na segunda suposição, apesar de muitos estudos incluídos na metanálise, ainda poderá ter estudos que não foram incluídos na metanálise porque não foram publicados. A razão para não publicação foi certamente a presença de vieses no estudo detectada pelo próprio autor do estudo, que, por isso, não submeteu o estudo para publicação no periódico ou foi detectada pelo editor do periódico que rejeitou a publicação. Ora, se a busca dos estudos para inclusão na metanálise foi abrangente, ou seja, na maioria das bases eletrônicas relacionadas ao assunto que está sendo estudado, que redundou em grande quantidade de estudos e neste rol de estudos há tanto estudos que mostraram resultados favoráveis e desfavoráveis, então aqueles estudos que não foram publicados, por qualquer que seja o motivo, não devem ser considerados para caracterizar a “suspeição de viés de publicação”.

Importante ressaltar que a classificação GRADE nada tem a ver com o planejamento rigoroso, detalhado e cuidadoso da equipe de pesquisadores que elaborou a metanálise. Dos cinco fatores que reduzem o nível de evidência e dos três fatores que aumentam o nível de evidência, somente dois tem relação com este planejamento: evidência indireta e viés de publicação. Estes dois fatores deixam de existir quando o planejamento prevê critérios de inclusão coerentes com o objetivo da metanálise e busca abrangente dos estudos.

Em outras palavras, a classificação GRADE depende dos estudos incluídos e não do planejamento detalhado, cuidadoso e rigoroso adotado pelo grupo de pesquisadores. Com isso, pode-se ter uma metanálise com nível de evidência muito baixo, apesar do planejamento muito detalhado, muito cuidadoso, muito rigoroso por parte da equipe de pesquisadores. Por isso, criei o seguinte jargão a ser inserido na DISCUSSÃO da metanálise: “Apesar do esforço deste grupo de pesquisadores para alcançar um alto rigor científico no planejamento e execução desta metanálise, o nível de evidência da metanálise foi (muito) baixo devido ... (ver GRADE). Sugere-se a realização de novos estudos que contemplem ... ou que evitem as limitações apresentadas pelos estudos incluídos.”

O nível de evidência da metanálise do Exemplo 1 – Metanálise de 13 estudos para verificar a eficácia da dança sobre o controle postural de pessoas com doença de Parkinson não foi avaliado, porque a revista para qual ela foi submetida não exigia este procedimento. Por este motivo este resultado não será apresentado aqui.

SIGNIFICÂNCIA ESTATÍSTICA E/OU SIGNIFICÂNCIA CLÍNICA?

Sobre o assunto significância estatística e significância clínica, Rob Herbert (2019) em seu artigo intitulado *“Research Note: Significance testing and hypothesis testing: meaningless, misleading and mostly unnecessary”* faz uma dura crítica ao teste de significância estatística e teste de hipótese. O título do seu artigo já seria suficiente para isso. Entretanto, ao ler o texto completo do seu artigo e, principalmente, as alternativas que ele apresentou para se evitar o “sem sentido, enganoso e principalmente desnecessário teste de significância estatística e teste de hipótese”, conclui-se que o teste de significância estatística tem sentido, não é enganoso e principalmente é necessário, pelas razões que apresentarei abaixo para cada alternativa por ele apresentada:

a. Alternativa apresentada por Rob Herbert (2019): *“Define study objectives in terms of estimation of effects”*.

O teste de significância estatística não foi criado para estimar o tamanho de efeito. O teste de significância estatística nos diz simplesmente a probabilidade do efeito obtido, que poderá ser grande ou pequeno, ter sido por acaso. Quando esta probabilidade é pequena ($P \leq 5\%$), não significa que temos a probabilidade de 95% de atribuir o efeito como causa da intervenção. A estatística foi concebida a partir da hipótese que todos os efeitos obtidos em uma pesquisa são 100% por acaso. Por isso, quando atribuímos o efeito como causa da intervenção, porque $P \leq 5\%$, na realidade estamos utilizando a lógica de como a probabilidade de ser por acaso é muito pequena, então o efeito foi devido a intervenção. Claro que apesar desta probabilidade do acaso ter sido pequena ela pode acontecer. Por isso mesmo, esta mesma significância estatística da hipótese nula é também o erro que poderá ser cometido pelo pesquisador ao rejeitar esta hipótese nula. Não podemos esquecer que, aqui no Brasil, a probabilidade de uma pessoa acertar na megasena é de 1/50.000.000 (1 / 50 milhões), probabilidade bem menor do que 1 / 20 (5%) utilizada no teste de significância estatística, mas assim mesmo há sempre alguém acertando a megasena (ganhando o prêmio).

Sempre digo aos meus alunos de mestrado e doutorado que um resultado científico não representa a verdade e sim o quanto estamos errados. O que difere o bom pesquisador dos demais é que a probabilidade de ele ter cometido um erro na pesquisa é muito pequena, menor que 5%.

b. Alternativa apresentada por Rob Herbert (2019): “*Do not report statistical significance*”.

Não apresentar a significância estatística é correr o risco de concluir a eficácia de uma intervenção pelo grande tamanho de efeito gerado, sem considerar que o acaso possa ser a causa do grande efeito gerado. Por exemplo, vamos imaginar que após uma determinada intervenção aplicada ao grupo B, e não ao grupo A, cada grupo com três participantes, apresentaram os seguintes resultados para um determinado desfecho: grupo A = 4, 5 e 6 e grupo B = 10, 50 e 90. A média do grupo A foi igual a 5, enquanto a do grupo B foi igual a 50. A diferença de médias $B - A = 50 - 5 = 45$. Isto representa um tamanho de efeito 9 vezes maior a favor do grupo B. Pela alternativa proposta por Rob Herbert (2019) acima, a conclusão seria que a intervenção aplicada foi eficaz para o desfecho considerado. Esta conclusão baseada somente no grande tamanho do efeito, sem considerar a significância estatística estaria errada. Por quê? O desvio-padrão (DP) do grupo A foi igual a 1, enquanto o do grupo B foi igual a 40. O intervalo de confiança 95% (IC95%) para a diferença de médias A - B (tamanho do efeito) de 45 obtido, foi igual a [0 a 90]. Isto significa que a probabilidade do tamanho do efeito de 45 é a mesma probabilidade do tamanho de efeito 0 e igual a 95%. O teste t de *Student* independente, que testou a significância estatística da diferença de médias dos grupos A e B apresentou resultado de $P = 0,12$ mostrando que o tamanho do efeito obtido foi por acaso, ratificando o resultado apresentado pelo intervalo de confiança 95%. Por que o teste de significância estatística atribuiu o grande efeito de 45 ao acaso? Para entendermos isto, precisamos observar os efeitos individuais dos três participantes do grupo B = 10, 50 e 90. Ora, como pode a mesma intervenção produzir efeitos tão diferentes? Se estes efeitos fossem provenientes da mesma intervenção não poderiam variar tanto assim. Para a estatística, todas as diferenças de médias (efeitos) são decorrentes do acaso (erro intrínseco ao se extrair duas amostras da mesma população). Quando estes efeitos se distribuem simetricamente em relação a sua média, sendo a maioria destes efeitos próximos a média e

a minoria destes efeitos afastados da média, 95% destes efeitos estão contidos entre a diferença de médias $\pm 1,96$ vezes a variabilidade desta diferença de médias, que pode ser estimada a partir do DP e da quantidade de participantes dos dois grupos considerados. Quanto maior o DP e menor a quantidade de participantes dos grupos maior será a variabilidade da diferença de médias (efeito) e maior será a probabilidade deste efeito estar contido no intervalo de confiança 95% e ser por acaso. Observe a quantidade de participantes e o DP dos grupos A e B e tire sua conclusão.

c. Alternativa apresentada por Rob Herbert (2019): “*Focus on the estimated size of the effect*”.

Para discutir esta terceira alternativa, vou utilizar o resultado padronizado de uma metanálise que foi 1,99 [-0,54 a 4,52]. Poderíamos atribuir o efeito padronizado de 1,99 como sendo mais do que duas vezes o valor de grande tamanho de efeito de 0,8 considerado por Cohen (1988). Este valor representa uma grande relevância clínica. Porém, não podemos dizer isso, porque a probabilidade de ser 1,99 é a mesma de ser 0 (zero) e igual a 95%. Foi isto que o IC95% [-0,54 a 4,52] nos mostrou. O resultado do IC95% foi ratificado pelo teste de hipótese nula ou significância estatística que apresentou $P = 0,12$ para o resultado da metanálise, ou seja, a probabilidade deste efeito padronizado de 1,99 ter sido por acaso foi grande, então foi por acaso. Não interessa para a ciência tamanhos de efeitos grandes ou muito grandes, como foi o caso desta metanálise, mas que tenham sido por acaso. Assim como não vai interessar para o paciente grandes efeitos que podem ou não ocorrer. Da mesma forma, não interessa para ciência efeitos pequenos, mas certos de que vão ocorrer (significância estatística ou $P \leq 5\%$). Assim como não vai interessar para o paciente gastar dinheiro, tempo, esforço para um efeito certo de ocorrer, mas irrelevante clinicamente. Sendo assim, me vejo obrigado a replicar Rob Herbert (2019):

It is important to understand that null hypothesis statistical testing in general, and p-values specifically, tell us nothing about the size of effects. A small and statistically significant p-value could arise when the data provide clear evidence of a large and clinically important effect. However, the same p-value could also arise when the data do not provide clear evidence of an important effect. Even more alarming, the same p-value could arise when the data provide clear evidence that the effect is small and unimportant. Statistically non-significant findings are

also uninformative about sizes of effects. A large, statistically non-significant p-value could arise when the data provide clear evidence of little or no effect, but the same p-value could arise when the data provide evidence that there could be a large and clinically important effect. So p-values or null hypothesis significance testing cannot be relied on for information about the size of effects.

Na realidade o artigo “*Significance testing and hypothesis testing: meaningless, misleading and mostly unnecessary*” de Rob Herbert (2019) trata de dois assuntos que a estatística faz uma distinção clara entre um e o outro, mas que algumas pessoas teimam em achar que é a mesma coisa: significância estatística e significância clínica, que eu gosto mais de chamar relevância clínica, a fim de estabelecer uma diferença entre elas já na denominação. Uma é diferente da outra, elas são independentes, mas se complementam. A significância estatística representa a probabilidade do efeito ter ocorrido por acaso, enquanto que a relevância clínica (significância clínica) nada tem a ver com a probabilidade, mas com o tamanho do efeito obtido. E elas se complementam porque não interessa uma relevância clínica por acaso, bem como não interessa uma certeza de uma irrelevância clínica. Por isso que o problema está no “e/ou”. A estatística nunca nos disse para utilizar uma ou outra na avaliação dos resultados de uma pesquisa, nem tampouco que uma pode substituir a outra. Pelo contrário, como elas se complementam, utilize as duas na mesma avaliação dos resultados de uma pesquisa. Reforço mais uma vez: não interessa para a ciência nem para o paciente um tamanho de feito grande, mas que ocorreu por acaso, nem tampouco um tamanho de feito pequeno, que ocorreu devido a intervenção.

Objetivamente, como a ciência classifica a relevância clínica (significância clínica)? Primeiro, temos que distinguir as medidas de efeito. No caso da diferença de médias de dois grupos, o Grupo de Revisões da Cochrane estabeleceu como relevância clínica para dor a redução de 30% (3 de 10 pontos), quando esta dor tiver sido mensurada pela Escala Visual Analógica ou Escala Numérica de Dor e de 8 e 12% (2 e 3 de 25 pontos, respectivamente), quando esta dor tiver sido mensurada pelo Questionário de Incapacidade de Roland Morris. Não sei dizer o porquê não foi estabelecido 7 e 8 pontos para o Questionário de Incapacidade de Roland Morris. Para o desfecho dor, no caso de diferença de médias, pelo menos se tem um parâmetro para tomada de

decisão. Entretanto, para os outros desfechos, a relevância clínica será decidida pela experiência do pesquisador.

Por outro lado, a ciência tem um parâmetro para tomada de decisão sobre a relevância clínica, quando a medida de efeito for a diferença de médias padronizada. Cohen (1988) propôs 0,2 para pequena relevância clínica; 0,5 para moderada relevância clínica; e 0,8 para grande relevância clínica.

Para a medida de efeito risco relativo (RR), o sistema GRADE considera como grande relevância clínica $0,5 > RR > 2,0$ e como muito grande relevância clínica $0,2 > RR > 5,0$. Como a diferença de risco (DR) é obtida pela diferença entre os riscos absolutos dos dois grupos, pode-se deduzir que haverá grande relevância clínica para $-1 > DR > 1$ e muito grande relevância clínica para $-8 > DR > 4$. Não há por enquanto uma referência para *odds ratio*.

REDIGINDO A METANÁLISE

Todo artigo científico, entre eles a metanálise, é composto das seguintes seções: TÍTULO, RESUMO, ABSTRACT, INTRODUÇÃO, MÉTODO, RESULTADOS, DISCUSSÃO, CONCLUSÃO, CONFLITO DE INTERESSE e REFERÊNCIAS.

TÍTULO

Título é o resumo do RESUMO. Deve ser escrito em até duas linhas e tem relação direta com o objetivo da metanálise. O título de um estudo tem que ser bem elaborado por dois motivos: (1) porque as pessoas buscam os estudos nas bases eletrônicas de informações científicas por palavras-título; (2) porque a seleção inicial dos artigos é feita pela leitura dos títulos. Para escrever um bom título adote a seguinte equação: TÍTULO = I + D + P: Metanálise. “I” foi a intervenção ou tratamento aplicado; “D” foi o desfecho mensurado; e “P” a característica principal dos participantes do estudo, se for o caso.

RESUMO

Resumo é outra parte do texto que merece atenção, pelos mesmos dois motivos do título. Quando não for por palavra-título, as pessoas buscam os estudos nas bases eletrônicas de informações científicas por palavras-resumo. A seleção seguinte dos artigos é feita pela leitura dos resumos. O resumo deve conter o objetivo do estudo, uma descrição sumária do método empregado, os resultados obtidos e a conclusão da metanálise. Não há discussão no resumo. Na linha seguinte, três palavras-chaves que não estejam no título.

INTRODUÇÃO

Nesta parte do estudo, você deve responder a seguinte pergunta: “O que foi estudado?” Para isso contextualize as variáveis da metanálise,

mostre a relevância e a justificativa para realiza-la e finalize a introdução definindo o objetivo da metanálise. Mostre clareza e concisão no texto da INTRODUÇÃO.

MÉTODO

O método é um planejamento racional, lógico, coerente e minucioso das ações executadas para atingir o objetivo da metanálise. Nesta parte da metanálise, você deve responder a seguinte pergunta: “Como foi estudado?” Para isso descreva com detalhes todos os procedimentos realizados, subdividindo o método nas suas partes: Protocolo de registro e local de acesso para a metanálise; Estabelecimento dos critérios de inclusão dos estudos; Criação da estratégia de busca e seleção dos estudos; Avaliação da qualidade metodológica e vieses dos estudos incluídos; Extração dos dados dos estudos incluídos; Metanálise dos dados extraídos e Avaliação do nível de evidência da metanálise.

Na subparte “Protocolo de registro e local de acesso para a metanálise” registre o número do protocolo informado pela plataforma PROSPERO e o link da internet de acesso à metanálise.

Na subparte “Estabelecimento dos critérios de inclusão dos estudos” registre os critérios estabelecidos para os estudos serem incluídos na metanálise, por exemplo, característica(s) dos participantes, tipo de intervenção, tipo de grupo de comparação, desfecho avaliado e tipo de estudo.

Na subparte “Criação da estratégia de busca e seleção dos estudos” registre a data da busca e da atualização da busca, quem fez a busca, as bases eletrônicas e outras fontes consultadas, os descritores e sinônimos utilizados na busca, se fez uso das bases Descritores em Ciências da Saúde (DeCS) e *Medical Subject Headings* (MeSH) para selecionar os descritores, se fez uso dos operadores booleanos, se fez uso de filtros e como foi feita a seleção dos estudos.

Na subparte “Avaliação da qualidade metodológica e vieses dos estudos incluídos” registre a escala utilizada na avaliação, como funcionou essa escala e quem fez a avaliação.

Na subparte “Extração dos dados dos estudos incluídos” registre quais foram os dados extraídos dos estudos incluídos na metanálise.

Na subparte “Metanálise dos dados extraídos” registre o pacote estatístico, o tipo de desfecho, o método estatístico, o modelo de análise e a medida de efeito selecionados que foram utilizados para metanalisar os dados.

Na subparte “Avaliação do nível de evidência da metanálise” registre a escala utilizada na avaliação, como funcionou essa escala e quem fez a avaliação.

É fundamental que cada ação desta do MÉTODO seja realizada por três pesquisadores independentemente: dois executam as ações e o terceiro decide sobre as divergências encontradas na comparação dos resultados dos dois primeiros. Se não for possível para todas as ações, pelo menos para a ação de extração dos dados dos estudos deve ser compulsório três pesquisadores.

RESULTADOS

Nesta parte do estudo, você deve responder a seguinte pergunta: “Quais os resultados encontrados?” Para isso apresente os resultados na ordem cronológica dos procedimentos executados no método. Os resultados podem ser apresentados em texto, tabelas, quadros ou figuras. Evite a redundância de resultados, ou seja, a apresentação de resultados em texto e tabelas ou quadros ou figuras, simultaneamente. As tabelas e quadros devem ser nominados na parte superior, enquanto as figuras na parte inferior.

É comum se ver um texto descrevendo criteriosamente as tabelas, quadros ou figuras seguintes. É como se alguém estivesse ao seu lado lendo as tabelas, quadros ou figuras. Isto é desnecessário, principalmente agora que os periódicos limitam o tamanho do artigo a ser publicado. Se as tabelas, quadros ou figuras contiverem todas as informações necessárias, não esquecer legendas para as siglas, e forem corretamente identificadas, não há necessidade desse texto descritivo antecedendo-a.

DISCUSSÃO

Discussão é a parte do estudo que responde a seguinte pergunta: “qual a interpretação dos resultados encontrados?” É a parte do estudo na qual você terá a oportunidade de CONVENCER o leitor de que o

RESULTADO da sua metanálise está CORRETO. Você ficaria convencido de algo, a partir da opinião de uma pessoa? Claro que não. Por isso, A DISCUSSÃO se apoia, se fundamenta, se baseia, se prova, se comprova em informações científicas publicadas (outros artigos). Um juiz não absolve um réu a partir da opinião do próprio réu ou de seu advogado, mas sim a partir das provas legítimas apresentadas. Então, você tem que apresentar provas legítimas ao leitor (artigos científicos) para convencê-lo de que o RESULTADO de sua metanálise está CORRETO. Discussão requer leitura e conhecimento do estado da arte.

CONCLUSÃO

É uma resposta ao objetivo, apoiada na racionalidade, lógica e coerência do método, considerando os resultados encontrados. Delimite a conclusão ao objetivo do estudo, método adotado para chegar ao objetivo e os resultados encontrados. Não extrapole a conclusão para aquilo que não foi pesquisado.

CONFLITO DE INTERESSE

É uma questão de ética na pesquisa. Deixe claro os apoios recebidos para a realização da metanálise, principalmente o financeiro e informe se a metanálise está atendendo algum interesse desses apoiadores ou não.

REFERÊNCIAS

Liste todos os artigos, livros e outras fontes citadas no texto da metanálise.

LENDO E INTERPRETANDO A METANÁLISE

Metanálise, como qualquer outro estudo científico, pode ter sido bem ou mal conduzida pelo autor. Antes de se encantar com o diamante, sugere-se que sejam verificadas a racionalidade e coerência interna nas ações executadas:

1) É razoável a justificativa para a elaboração da metanálise?

Uma metanálise é justificada quando não há uma metanálise anterior ou, havendo metanálise anterior, há também novos estudos que não puderam ser considerados na metanálise anterior ou ainda quando a metanálise anterior possui severas limitações que justifiquem a realização de nova metanálise, com um planejamento mais cuidadoso, detalhado e rigoroso.

2) A metanálise é específica, ou seja, os “critérios de inclusão dos estudos” delimitam adequadamente o assunto, convergindo para o objetivo da metanálise?

Compare o objetivo com os critérios de inclusão dos estudos na metanálise. Os critérios de inclusão devem espelhar o objetivo da metanálise e devem delimitar claramente o tratamento, a comparação, o desfecho, a característica principal dos participantes e o tipo de estudo. Somente assim será evitado a evidência indireta, ou seja, inclusões de estudos que nada têm a ver com o objetivo da metanálise, que criarão resultados não confiáveis.

3) A metanálise é abrangente, ou seja, considera as principais bases eletrônicas de informações científicas e têm descritores adequados com os “critérios de inclusão dos estudos” para a recuperação dos estudos?

Verifique se a estratégia de busca da metanálise alcançou todas as principais bases eletrônicas de informações científicas. A busca abrangente inibe a suspeição de viés de publicação. Observe se os

descritores estão coerentes com o objetivo da metanálise e se os sinônimos foram considerados.

4) A extração dos dados dos estudos incluídos foi realizada, no mínimo, por três pessoas?

Esta é a ação mais sensível da metanálise, porque um equívoco na extração dos dados do estudo ou um erro de digitação destes dados pode comprometer seriamente o resultado da metanálise. A presença de dois pesquisadores para extrair os dados e de um terceiro para decidir sobre as discordâncias diminuirá a chance do erro.

5) Havendo inconsistência verdadeira entre os estudos incluídos, ela foi incorporada na “análise dos dados” e explicada?

Observe no *forest plot* a quantidade de inconsistência (I^2) e a probabilidade dela ter sido por acaso (o valor-P). Trace uma linha vertical passando pelo centro do losango da metanálise. Se esta linha vertical não perfurar os intervalos de confiança 95% de todos os estudos incluídos e a probabilidade da inconsistência ter sido por acaso for menor ou igual a 5%, obrigatoriamente o modelo de análise deverá ter sido o de efeitos randômicos. Caso contrário, o resultado da metanálise não pode ser confiável.

6) Havendo unidades de medida ou escalas diferentes para o mesmo desfecho entre os estudos incluídos, a medida de efeito foi padronizada?

Este cuidado é aplicável ao desfecho contínuo. Observe no Quadro com os dados extraídos dos estudos como foi mensurado o desfecho entre os estudos incluídos. Se tiver escalas ou unidades de medida do desfecho diferentes entre os estudos, a medida de efeito obrigatoriamente deverá ter sido a diferença de médias padronizada.

7) A quantidade de estudos incluídos é satisfatória?

Metanálises com poucos estudos inflacionam os pesos dos estudos incluídos, que por sua vez inflacionam a heterogeneidade. O aumento da heterogeneidade pode aumentar o intervalo de confiança 95% da metanálise. Estudos com poucos participantes podem aumentar os intervalos de confiança 95% dos estudos e diminuir os pesos dos estudos, conseqüentemente vão diminuir a medida de efeito da metanálise e

aumentar o intervalo de confiança 95% da metanálise, gerando um resultado pouco confiável.

8) Não ocorreu viés de publicação?

Observe o *funnel plot*. Caso haja o triângulo isósceles, não há suspeição de viés de publicação. Porém, a avaliação com o *funnel plot* é muito subjetiva e pouco confiável quando a quantidade de estudos é inferior a dez. Em qualquer das situações veja se o(s) autor(es) utilizaram também o teste de Egger para desfechos contínuos ou de Begg para desfechos dicotômicos. Estes testes são mais objetivos que o *funnel plot* e devem ser considerados. Ambos os testes informam a probabilidade do viés de publicação existente ter sido por acaso. Se esta probabilidade for maior que 5%, atribui-se o viés de publicação ao acaso, ou seja, não há suspeição de viés de publicação.

9) O resultado da metanálise foi significativo, ou seja, teve significância estatística?

Observe o intervalo de confiança 95% da metanálise no *forest plot*. Se a medida de efeito for o risco relativo (RR) ou *odds ratio* (OR), o intervalo de confiança 95% da metanálise não deve conter o valor 1,00, ou seja, as incidências ou chances prováveis do desfecho nos GE e GNE foram diferentes, com a probabilidade de 95% disto ocorrer. Se a medida de efeito for a média ou diferença de médias (DM) ou diferença de médias padronizada (DMP), o intervalo de confiança 95% da metanálise não deve conter o valor zero, ou seja, as médias ou diferença de médias prováveis do desfecho nos grupos foram diferentes, com a probabilidade de 95% disto ocorrer. Você deve confirmar isto, observando o losango (diamante) que representa o resultado gráfico da metanálise, que não deve cruzar a linha de nulidade, linha vertical que passa pelo valor 1 ou zero, respectivamente. Observe também a probabilidade do resultado da metanálise não ter ocorrido por acaso (o valor-P em *Test for overall effect*). Se o valor-P for menor ou igual a 5%, muito pouco provável ter sido por acaso, então o resultado é verdadeiro.

10) O resultado da metanálise teve relevância clínica, ou seja, significância clínica?

Considero esta questão tão importante, que dediquei uma boa parte deste livro para esta discussão. Por favor leia novamente SIGNIFICÂNCIA ESTATÍSTICA E/OU SIGNIFICÂNCIA CLÍNICA?

Caso as respostas a estas 10 questões tenham sido sim, agora sim você pode ficar encantado com o diamante encontrado.

REFERÊNCIAS

- ALMEIDA H. et al., Effect of Dance on Postural Control in People with Parkinson's Disease: A Meta-Analysis Review. **Journal of Aging and Physical Activity**, v. 29, ed. 1, p. 1-11. July 2020.
- BORENSTEIN, M; HEDGES, L; HIGGINS, J; ROTHSTEIN, H. **Introduction to Meta-Analysis**, Chichester: Simo Puntanen, 2009.
- BORENSTEIN, M; HIGGINS, J; HEDGES, L; ROTHSTEIN, H. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. **Res. Syn. Meth.**, v. 8, p. 5–18, 2017.
- BROOKE, BS; SCHWARTZ, TA; PAWLIK, TM. MOOSE Reporting Guidelines for Meta-analyses of Observational Studies. **JAMA surgery**, v. 156, p. 787-88, 2021.
- BUNN, Priscila dos Santos. **Padrão de movimento e o risco de lesões musculoesqueléticas**. 2020. 254 f. Tese (Doutorado em Ciências do Exercício e do Esporte) – Instituto de Educação Física e Desportos, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.
- COCHRANE HANDBOOK FOR SYSTEMATIC REVIEWS OF INTERVENTIONS VERSION 6.2*. Cochrane: HIGGINS, JPT et al., 2021. Available from www.training.cochrane.org/handbook. Updated February 2021.
- COHEN, J. **Statistical Power Analysis for the Behavioral Sciences**, 2ed ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 1988.
- EGGER, M; SMITH, GD; PHILLIPS AN. “Meta-Analysis: Principles and Procedures”. **BMJ: British Medical Journal**, vol. 315, pp. 1533-37, 1997.
- FUCHS, SCPC; PAIM, BS. Revisão sistemática de estudos observacionais com metanálise. **Revista HCPA**. Porto Alegre. Vol. 30, ed. 3, p. 294-301, 2010.
- FURLAN, AD; PENNICK, V; BOMBARDIER, C; VAN TULDER, M. 2009 updated method guidelines for systematic reviews in the cochrane back review group. **Spine**. v. 34, ed. 18, p. 1929-41. 2009.
- GLASS, GV. Primary, Secondary, and Meta-Analysis of Research. **Educational Researcher**, vol. 5, pp. 3-8, 1976.
- GORDON, HG. et al. GRADE guidelines 5: Rating the quality of evidence - publication bias. **Journal of Clinical Epidemiology**. v. 64, p. 1277-1282. 2011.

- GORDON, HG. et al. GRADE guidelines 6: Rating the quality of evidence – imprecision. **Journal of Clinical Epidemiology**. v. 64, p. 1283-1293. 2011.
- GORDON, HG. et al. GRADE guidelines 7: Rating the quality of evidence – inconsistency. **Journal of Clinical Epidemiology**. v. 64, p. 1294-1302. 2011.
- GORDON, HG. et al. GRADE guidelines 8: Rating the quality of evidence – indirectness. **Journal of Clinical Epidemiology**. v. 64, p. 1303-1310. 2011.
- GORDON, HG. et al. GRADE guidelines 9: Rating the quality of evidence. **Journal of Clinical Epidemiology**. v. 64, p. 1311-16. 2011.
- HERBERT, R. Research Note: Significance testing and hypothesis testing: meaningless, misleading and mostly unnecessary. **Journal of Physiotherapy**, v. 65, p. 178-81, 2019
- HIGGINS, JPT; THOMPSON, SG. Quantifying Heterogeneity in a Meta-Analysis. **Statistics in Medicine**, vol. 21, p. 1539–1558, 2002.
- HIGGINS, JPT; THOMPSON, SG; DEEKS, JJ; ALTMAN, DG. Measuring inconsistency in meta-analyses. **BMJ**, v. 327, p. 557-60, 2003.
- KULIK, JA; Kulik, CLC. The Concept of Meta-Analysis. **International Journal of Educational Research**, vol. 13, ed. 3, p. 227-340, 1989.
- MATOS, MI. **Power Ballet Original, Capacidade Cardiorrespiratória, Muscular e Performance no Ballet Clássico**. Tese em andamento- Universidade do Estado do Rio de Janeiro- UERJ, Rio de Janeiro, 2021.
- Meta-Analysis: Fixed Effect vs. Random Effects. BORENSTEIN, M; HEDGES, L; ROTHSTEIN, H. 2007. Available: <https://www.metaanalysis.com/downloads/Metaanalysis%20fixed%20effect%20vs%20random%20effects.pdf>. Acesso em 2020.
- NEYELOFF, JL; FUCHS, SC; MOREIRA, LB. Meta-analyses and Forest plots using a Microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. **BMC research notes**, v. 5, ed. 1, p. 1-6, 2012.
- ROCHA, FCC. **Efeitos do treinamento aeróbio na capacidade funcional de pacientes com cirrose hepática**. Tese em andamento. Universidade do Estado do Rio de Janeiro- UERJ, Rio de Janeiro, 2021.
- SCHÜNEMANN, HJ. et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. **Research synthesis methods**, v. 4, ed. 1, p. 49-62, 2013.
- SEDGWICK, P. Meta-analyses: what is heterogeneity? **British Medical Journal**, v. 350, p. 1435, 2015.

SOUSA, MRD; RIBEIRO ALP. Revisão Sistemática e Meta-análise de Estudos de Diagnóstico e Prognóstico: um Tutorial. **Arquivo Brasileiro de Cardiologia**, vol. 92, ed. 3, p. 229-238, 2009.

STATISTICAL ALGORITHMS IN REVIEW MANAGER 5: Deeks, JJ; HIGGINS, JPT, 2010. Disponível em: file:///C:/Program%20Files/Review%20Manager%205.4/doc/userguide/Statistical_algorithms.pdf

STROUP, DF. et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. **JAMA**, v. 283, ed. 15, p. 2008-2012, 2000.

WHITEHEAD, A. **Meta-Analysis of Controlled Clinical Trials**. Chichester: John Wiley & Sons, 2002. 352 p.

ZHANG, Y *et al.* GRADE Guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences – Inconsistency, Imprecision, and other Domains. **Journal of Clinical Epidemiology**, v. 11, p. 83-93, Jul 2018.



**Pois é dando que se recebe,
É perdoando que se é perdoado...**

(Trecho da Oração de São Francisco de Assis)



ISBN 978-85-7993-971-6

