

# Avaliação Crítica de uma Revisão Sistemática e Meta-Análise: Da Execução da Meta-Análise à Exploração da Heterogeneidade

## *Critical Appraisal of a Systematic Review and Meta-Analysis: Performing Meta-Analysis and Addressing Heterogeneity*

Bernardo Sousa-Pinto<sup>1,2</sup>, Luís Azevedo<sup>1,2,3\*</sup>

### Afiliação

<sup>1</sup> MEDCIDS – Departamento de Medicina da Comunidade, Informação e Decisão em Saúde, Faculdade de Medicina da Universidade do Porto, Universidade do Porto, Porto, Portugal.

<sup>2</sup> CINTESIS – Centro de Investigação em Tecnologias e Serviços de Saúde, Porto, Portugal.

<sup>3</sup> Centro Nacional de Observação em Dor, Faculdade de Medicina da Universidade do Porto, Porto, Portugal.

### Palavras-chave

Interpretação Estatística de Dados; Meta-Análise como Assunto; Projecto de Investigação

### Keywords

*Data Interpretation, Statistical; Meta-Analysis as Topic; Research Design*

## RESUMO

A meta-análise corresponde a um conjunto de metodologias estatísticas usadas na adequada agregação ou síntese quantitativa de medidas quantitativas ou estimativas provenientes de diferentes estudos primários. No decurso de uma revisão sistemática, revela-se adequado proceder a meta-análise quando se deseja efectuar uma síntese dos resultados quantitativos dos estudos primários incluídos, desde que estes sejam suficientemente homogéneos do ponto de vista metodológico e clínico. Pode ser feita meta-análise tendo por base qualquer medida quantitativa que sumarie os resultados dos estudos primários incluídos, pelo que a pertinência da medida escolhida se reveste da maior importância. Adicionalmente, deverá avaliar-se qual o modelo meta-analítico (de efeitos fixos ou de efeitos aleatórios) mais apropriado, atendendo às características específicas dos mesmos. Para além da medida meta-analítica, importa sempre avaliar a presença e magnitude da heterogeneidade (diferenças entre os estudos que estão para além do esperado). De uma forma geral, as medidas meta-analíticas podem ser vistas como as melhores respostas disponíveis para a questão de investigação em apreço, uma vez que resultam da adequada síntese da melhor evidência disponível, e desde que a selecção da mesma tenha sido feita de maneira abrangente e não enviesada. No entanto, quando estamos perante heterogeneidade substancial, as medidas meta-analíticas podem não ser essa melhor resposta e devem ser interpretadas com

particular cuidado, sendo fundamental nesses casos encetar esforços, aplicando métodos adequados, com o objectivo de identificar as possíveis causas da heterogeneidade. Nestas situações deve mesmo assumir-se que o objectivo primeiro da meta-análise passa a ser o da análise e adequada explicação da heterogeneidade, em vez da mera procura por medidas meta-analíticas de síntese que podem não espelhar, isoladamente, a melhor resposta à questão de investigação.

## ABSTRACT

A meta-analysis encompasses a set of statistical methods used to combine and synthesize a set of quantitative estimates or measures coming from multiple scientific primary studies. Within the context of a systematic review, it is adequate to perform meta-analysis when we aim to synthesize quantitatively the primary studies included, as long as they are sufficiently homogeneous from both a methodological and a clinical point of view. A meta-analysis may be based on any quantitative measure summarizing the results of the primary studies included; therefore, one of the most important first topics to be appraised in a meta-analysis is the adequate choice of the effect measures to be analyzed. Next, a particular attention should be placed on the choice and adequacy of meta-analytical model (fixed or random effects model) and the weighting methods used, taking into account the specific characteristics of the primary studies included. Beyond the pooled meta-analytical measure, it is of paramount relevance in a meta-analysis to assess the existence and extent of heterogeneity (between studies variability beyond what would be expected). Generally, a meta-analytical pooled measure, as long as its evidence base is sufficiently comprehensive and unbiased, may very

Autor Correspondente/Corresponding Author\*:

Luís Filipe Azevedo

Morada: IMEDCIDS - Departamento de Medicina da Comunidade, Informação e Decisão em Saúde, Faculdade de Medicina da Universidade do Porto, Rua Dr. Plácido da Costa, 4200-450, Porto, Portugal.

E-mail: lazevedo@med.up.pt

well be the best available answer to a given research question, since it encompasses and synthesizes the best available scientific evidence. However, when substantial heterogeneity is found, pooled meta-analytical measures should be interpreted with great caution and the primary effort should be targeted to the exploration of possible causes and moderators of heterogeneity. In such situations, the main objective of the meta-analysis becomes the analysis and adequate assessment of heterogeneity, and not only the mere search for meta-analytical pooled measures that may not be per se the best available answer to the research question.

## INTRODUÇÃO

Os artigos anteriores desta série versaram a avaliação crítica de uma revisão sistemática.<sup>1,2</sup> Como previamente discutido, as revisões sistemáticas têm um duplo objectivo de analisar e sintetizar toda a evidência científica produzida relativamente a uma determinada questão de investigação. Contudo, em algumas situações, mais que uma síntese apenas qualitativa da evidência científica, revela-se possível e desejável proceder a uma síntese quantitativa da mesma. Para tal, poderia pensar-se em simplesmente calcular uma média das medidas de efeito apresentadas pelos diferentes estudos primários - todavia, tal implicaria “atribuir o mesmo valor” a todos os estudos, não “premiando” os resultados de estudos que apresentam estimativas mais precisas das medidas de efeito. Ao invés, opta-se preferencialmente por recorrer a métodos de meta-análise. De facto, a meta-análise não constitui um tipo particular de estudos, mas antes diz respeito a um conjunto de metodologias estatísticas usadas para proceder à síntese quantitativa de resultados, medidas ou estimativas quantitativas provenientes de diferentes estudos primários.<sup>3</sup> Em termos gerais, comum a este conjunto de metodologias é a existência de uma ponderação dos diferentes estudos com base na “qualidade estatística” ou precisão das suas estimativas ou medidas. Não é obrigatório - e nem sempre é desejável - uma revisão sistemática acompanhar-se de meta-análise, assim como existem situações em que são executadas meta-análises fora do contexto de revisões sistemáticas. Nesse sentido, o artigo de perspectiva desta edição da Revista da Sociedade Portuguesa de Anestesiologia discute a avaliação crítica de meta-análises. Tal como para os dois artigos prévios desta série, a abordagem conceptual será acompanhada por uma análise crítica da revisão sistemática de Meng T *et al.* *Impact of spinal anaesthesia vs. general anaesthesia on peri-operative outcome in lumbar spine surgery: a systematic review and meta-analysis of randomised, controlled trials. Anaesthesia.* (2017;72:391-4016).<sup>4</sup>

## Adequação da execução de meta-análise

No decurso da avaliação crítica de uma meta-análise, importa desde logo perceber até que ponto a sua execução se revelou adequada. Se numa revisão sistemática é possível incluir estudos de diferentes desenhos (e.g., estudos experimentais e observacionais), é habitualmente recomendado que numa mesma meta-análise apenas constem estudos do mesmo tipo. Do mesmo modo, não podem ser misturadas medidas de efeito numa mesma meta-análise. De facto, é possível proceder a meta-análise tendo por base qualquer medida de efeito, definidas neste contexto de forma mais lata e incluindo toda e qualquer medida quantitativa que possa dar resposta à questão de investigação em apreço - desde médias a riscos relativos, passando por proporções, *odds ratio*, coeficientes de correlação... -, bastando para isso os autores conhecerem a medida de efeito em causa e o respectivo erro-padrão para cada um dos estudos primários incluídos na meta-análise (ou, ainda melhor, disporem dos dados brutos necessários para calculá-los!). Assim, compete aos autores identificarem a medida de efeito mais adequada em relação à qual procederão à meta-análise (e em relação à qual os resultados de todos os estudos primários deverão ser expressos).<sup>5</sup> Considerados estes aspectos, depreende-se que, embora isso fosse o ideal, nem todos os estudos primários incluídos numa revisão sistemática necessitam obrigatoriamente de constar da(s) meta-análise(s) que nela é(são) apresentada(s). Constata-se ainda que uma mesma revisão sistemática pode apresentar várias meta-análises, desde logo se avaliar diferentes *outcomes* ou se considerar separadamente determinados subgrupos - a título de exemplo, se uma revisão sistemática incluir estudos experimentais e estudos observacionais, poderá optar-se pela apresentação de uma meta-análise só com um dos tipos de estudos, ou pela apresentação de duas meta-análises separadas em função do tipo de estudos primários. A razão subjacente a esta recomendação está relacionada com o facto de sabermos, pelas características próprias destes estudos, que as medidas de efeito que cada um apresenta são de natureza diferente (desde logo, pela clara diferença relativamente à validade e risco de enviesamento das estimativas), justificando essa diferença a sua meta-análise separada. Note-se que, em alguns casos, não basta apenas aferir se a medida de efeito foi escolhida adequadamente. A título de exemplo, perante meta-análises de proporções, poderá ser boa prática recorrer a transformações logarítmicas das mesmas, nomeadamente caso se trate de proporções bastante pequenas ou elevadas.<sup>6</sup> No estudo de Meng *et al.*, os autores procederam a 11 meta-análises diferentes, cada uma avaliando um *outcome* distinto.<sup>4</sup> Sendo os estudos primários incluídos quanto à sua direccionalidade do tipo anterógrado, o risco relativo foi a medida de efeito utilizada para variáveis dicotómicas; para variáveis contínuas, foi utilizada a diferença de médias padronizada.

## Modelos de meta-análise e métodos de ponderação

Verificando-se a adequação de proceder à meta-análise, importa atentar na metodologia seguida para a sua execução. Muito simplificada, existem dois modelos de meta-análise, nomeadamente o modelo de efeitos fixos e o modelo de efeitos aleatórios. O modelo de efeitos fixos considera que todos os estudos primários provêm de uma mesma população, estando a estimar uma mesma realidade - nesse sentido, assumimos que as diferenças de resultados observadas entre os diferentes estudos são inteiramente explicadas pela variabilidade amostral.<sup>7</sup> Contudo, esta assumption nem sempre se verifica - muitas vezes, existem diferenças de tal magnitude entre os estudos primários que deixa de ser possível assumir que estes provêm de uma mesma população. Nessa situação, opta-se pelo modelo de efeitos aleatórios, o qual considera e incorpora na própria medida meta-analítica a possibilidade de existência de diferenças entre os estudos que vão para além da simples variabilidade amostral.<sup>7</sup> Ou seja, o modelo de efeitos aleatórios considera que os estudos primários provêm de populações diferentes e estão a medir realidades diferentes. Uma vez que, contrariamente ao modelo de efeitos fixos, o modelo de efeitos aleatórios necessita simultaneamente de acomodar a variabilidade dentro de cada estudo primário e a variabilidade entre estudos, a estimativa da medida meta-analítica é frequentemente menos precisa. Assim, na presença de uma meta-análise executada de acordo com o modelo de efeitos fixos, importa perceber se efectivamente esta opção se revelou adequada, ou se os autores deveriam antes ter optado por um modelo de efeitos aleatórios. Em bom rigor, existem formas estatísticas que ajudam a identificar qual o modelo mais adequado a utilizar (*vide infra*).

Para além do modelo adoptado, importa ter em conta o método de ponderação utilizado. Como referido anteriormente, a meta-análise caracteriza-se adequadamente pelo cálculo de uma medida de síntese que resulta da ponderação de cada um dos estudos primários incluídos em função da sua "qualidade", medida pela precisão das suas estimativas - existem diferentes métodos de ponderação, sendo o mais utilizado o de ponderação pelo inverso da variância. Tal como o nome indica, a ponderação atribuída a cada estudo primário (i.e., o "peso" que cada estudo tem na medida meta-analítica) é tanto maior quanto menor for a sua variância (neste caso entendida como o quadrado do erro-padrão - medida por excelência da precisão da estimativa); assim, é atribuída maior ponderação aos estudos com maior precisão.<sup>8</sup>

No estudo de Meng *et al*, verificou-se adopção de um modelo de efeitos fixos, excepto se fosse detectada heterogeneidade elevada (passando nesse caso a ser utilizado o modelo de efeitos aleatórios).<sup>4</sup> O método de ponderação não se encontra explícito na secção dos métodos, mas a análise das figuras permite perceber que foi usado o método de

Mantel-Haenszel, que é um método de ponderação também adequado e específico para variáveis de resultado (*outcome*) dicotómicas, muitas vezes utilizado nesses casos em alternativa à ponderação pelo inverso da variância.

## Avaliação e exploração da heterogeneidade

Da execução de uma meta-análise, resulta a obtenção de uma medida meta-analítica (ou seja, uma medida de efeito que faz a síntese, através de uma adequada ponderação, das medidas provenientes de vários estudos primários) bem como do respectivo intervalo de confiança. Todavia, este não é o único resultado no qual se deve atentar aquando da avaliação crítica de uma meta-análise - para além da medida meta-analítica, é fundamental aferir a heterogeneidade existente. A heterogeneidade diz respeito à existência de diferenças entre as medidas de efeito para além daquelas que seriam esperadas e meramente explicadas pelo acaso/variabilidade amostral.<sup>8</sup> A heterogeneidade pode ser quantificada através da estatística Q de Cochran ou da estatística de  $I^2$ . A estatística Q de Cochran permite testar a hipótese nula de não existência de heterogeneidade, permitindo a obtenção de um valor de  $p$  para esse teste que, se inferior ao limiar de significância (tipicamente, neste caso, 0,10 ou 0,05), indicia a presença de heterogeneidade significativa (rejeitando-se a hipótese nula). Por outro lado, a estatística de  $I^2$  é expressa sob a forma de uma percentagem, informando acerca da magnitude da heterogeneidade (proporção da variabilidade total que se deve às diferenças entre os estudos e que está para além da esperada) - não existindo valores de *cut-off* universalmente aceites, considera-se que um valor de  $I^2$  superior a 40%-50% configura heterogeneidade grave.<sup>8</sup>

De uma forma geral, as medidas meta-analíticas podem ser vistas como as melhores respostas disponíveis para a questão de investigação em apreço, uma vez que resultam da adequada síntese da melhor evidência disponível, desde que a selecção da mesma tenha sido feita de maneira abrangente e não enviesada. No entanto, quando estamos perante heterogeneidade grave, as medidas meta-analíticas podem não ser essa melhor resposta e devem ser interpretadas com particular cuidado, sendo fundamental nesses casos encetar esforços, aplicando métodos adequados, com o objectivo de identificar as possíveis causas da heterogeneidade. Nestas situações, deve mesmo assumir-se que o objectivo primeiro da meta-análise passa a ser o da análise e adequada explicação da heterogeneidade, em vez da mera procura por medidas meta-analíticas de síntese que podem não espelhar, isoladamente, a melhor resposta à questão de investigação.

Perante a detecção de heterogeneidade, é consensual a recomendação da opção pelo modelo de efeitos aleatórios (caso a heterogeneidade seja moderada a grave, poderá nem fazer sentido apresentar os resultados da medida meta-analítica, pois nesse caso esta resulta da agregação de

medidas aparentemente provenientes de realidades muito distintas). Adicionalmente, importa procurar identificar as causas da heterogeneidade - ou seja, é fundamental identificar as diferenças clínicas e/ou metodológicas entre os estudos que estão por trás e explicam a heterogeneidade detectada. Para tal, é possível recorrer a análises de sensibilidade, análises de subgrupos ou meta-regressão. A análise de sensibilidade implica a repetição de meta-análise tendo por base pressupostos alternativos dos da análise principal (e.g., avaliando apenas uma parcela de estudos primários)<sup>8</sup> - uma das formas mais comuns prende-se com a análise de sensibilidade *leave-one-out*, em que são repetidas meta-análises retirando-se cada estudo primário individualmente, permitindo esta análise detectar os estudos mais influentes na medida meta-analítica e na eventual heterogeneidade observada. A análise de subgrupos consiste na execução de meta-análises separadas para grupos de estudos tendo em conta variáveis categóricas que definem características dos estudos que podem potencialmente explicar a heterogeneidade observada (e.g., estudos europeus *versus* estudos norte-americanos; estudos realizados pela Indústria *versus* estudos da Academia).<sup>9</sup> Por fim, a meta-regressão consiste na construção de modelos estatísticos univariável ou multivariável específicos - análogos aos modelos de regressão habitualmente usados noutros contextos (e.g., regressão linear ou logística) - de modo a identificar variáveis moderadoras de heterogeneidade (i.e., variáveis que contribuam significativamente para explicar heterogeneidade).<sup>9</sup> Note-se que é desejável a existência de pelo menos 5 a 10 estudos primários incluídos para a execução de meta-regressão.

Em suma, no decurso da análise crítica de uma meta-análise, revela-se fundamental atentar na magnitude da heterogeneidade. Caso se verifique heterogeneidade substancial, os resultados deverão ser interpretados com cautela e, por vezes, uma única medida meta-analítica poderá até nem corresponder à resposta mais adequada à questão de investigação. Deverá ainda ser tido em conta se os autores utilizaram o modelo de efeitos aleatórios, discutiram adequadamente a existência de heterogeneidade e procuraram possíveis causas para a sua existência, nomeadamente através de análises de sensibilidade, análises de subgrupos e/ou construção de modelos de meta-regressão.

No estudo de Meng *et al*, várias meta-análises resultaram em heterogeneidade substancial.<sup>4</sup> Contudo, esta não foi adequadamente discutida nem sequer explorada. Se o número de estudos primários se revelou insuficiente para proceder a meta-regressão, teria sido possível proceder a análises de sensibilidade ou de subgrupos. Adicionalmente, deveria ter sido dada maior ênfase ao cuidado que deve ser colocado na interpretação das medidas meta-analíticas nas situações em que estava presente heterogeneidade grave.

## Representação gráfica dos resultados meta-analíticos

Em termos gráficos, o *forest plot* constitui a representação por excelência dos resultados de uma meta-análise. Neste gráfico, tipicamente, é possível encontrar representadas as estimativas e respectivos intervalos de confiança das medidas de efeito de cada um dos estudos primários incluídos e da medida meta-analítica, com indicação do tipo de modelo (de efeitos fixos ou aleatórios) e métodos de ponderação utilizados. Em função do anteriormente referido, será sempre importante que este gráfico contenha também estatísticas e medidas que permitam avaliar a heterogeneidade observada. A título de exemplo, a interpretação de um *forest plot* encontra-se explicada na Fig. 1.

## CONSIDERAÇÕES FINAIS

De uma forma geral, as medidas meta-analíticas podem ser vistas como as melhores respostas disponíveis para a questão de investigação em apreço, uma vez que resultam da adequada síntese da melhor evidência disponível. No entanto, isto só será verdade se todo o trabalho da revisão sistemática e meta-análise tiver sido feito de forma metodologicamente rigorosa e quando factores como a heterogeneidade observada assim o permitam. Por isso mesmo, na prática os resultados meta-analíticos deverão sempre ser interpretados criticamente, e nunca de forma imediata e acrítica como “números mágicos” correspondentes à resposta última à questão de investigação em apreço. Isto é particularmente importante hoje, e cada vez mais no futuro, uma vez que a divulgação e publicação de estudos de meta-análise de fraca qualidade e inadequadamente executados, apresentados e/ou interpretados (por razões mais ou menos dolosas!) é hoje uma realidade muito comum e nestas situações - que gostaríamos que fossem infrequentes - estes estudos podem constituir-se como perigosas armas de desinformação.<sup>10</sup>

Assim, e a título de sumário final, perante este tipo de estudos revela-se sempre fundamental analisar de forma sistemática a adequação da sua execução e apresentação, ter em conta se o modelo foi bem escolhido, atentar na heterogeneidade observada e nos esforços que foram conduzidos para explorar as fontes dessa mesma heterogeneidade. De facto, como demonstrado, na presença de heterogeneidade substancial, os resultados meta-analíticos deverão ser cautelosamente interpretados até porque, mais do que uma única medida meta-analítica, a melhor resposta à questão de investigação pode passar, nesses casos, por várias medidas meta-analíticas distintas e pela adequada identificação e discussão das variáveis moderadoras da heterogeneidade.

Nota: Este artigo foi escrito de acordo com a antiga ortografia, não estando ao abrigo do novo Acordo Ortográfico.



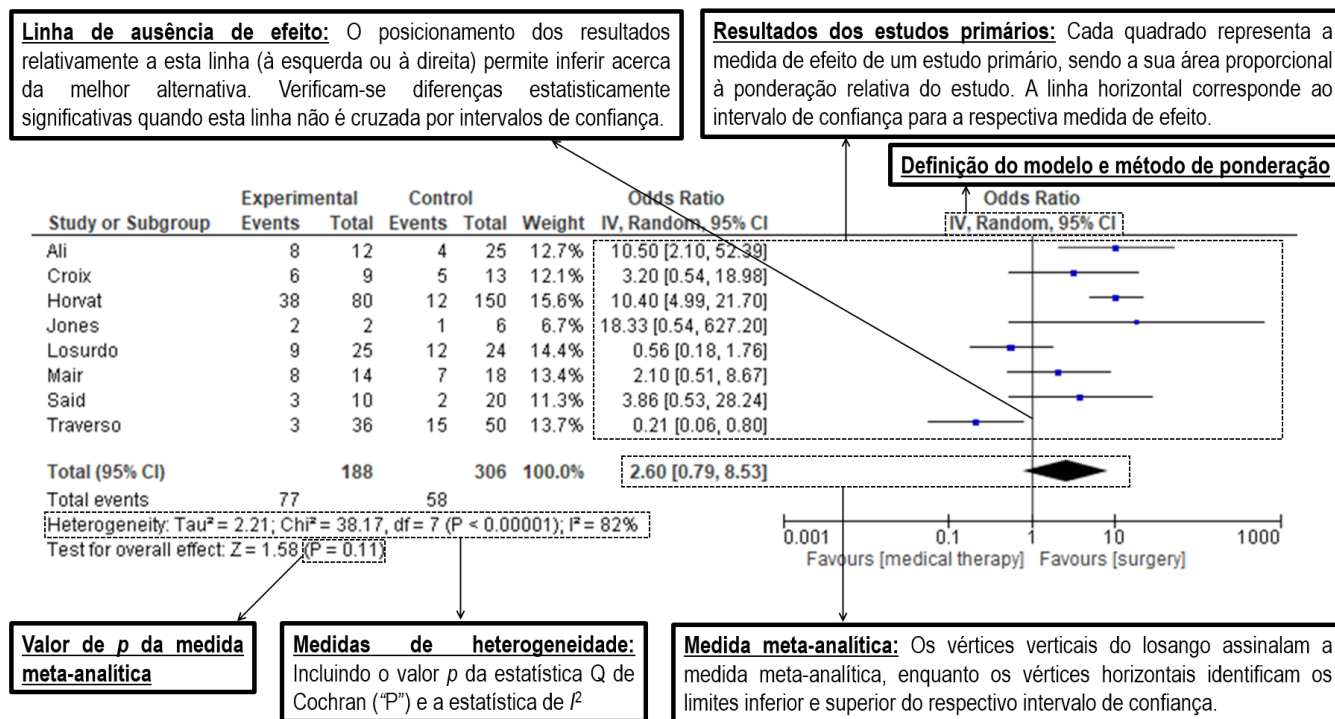


Figura 1. Exemplo de forest plot e respectiva interpretação

**Responsabilidades Éticas**

**Conflitos de interesse:** Os autores declaram não possuir conflitos de interesse.  
**Suporte financeiro:** O presente trabalho não foi suportado por nenhum subsídio ou bolsa.  
**Proveniência e revisão por pares:** Comissionado; com revisão externa por pares.

**Ethical Disclosures**

**Conflicts of interest:** The authors have no conflicts of interest to declare.  
**Financing support:** This work has not received any contribution, grant or scholarship.  
**Provenance and peer review:** Commissioned; with externally peer reviewed.

Submissão: 27 de agosto, 2019 | Aceitação: 03 de setembro, 2019  
 Received: 27<sup>th</sup> of August, 2019 | Accepted: 03<sup>rd</sup> of September, 2019

**REFERÊNCIAS**

- Sousa-Pinto B, Azevedo LF. Avaliação crítica de uma revisão sistemática e meta-Análise: Da definição da questão de investigação à pesquisa de estudos primários. Rev Soc Port Anestesiol. 2019;28:53-6.
- Sousa-Pinto B, Azevedo LF. Avaliação crítica de uma revisão sistemática e meta-Análise: Da selecção à avaliação da qualidade dos estudos primários. Rev Soc Port Anestesiol. 2019; 185:191-5.
- Shorten A, Shorten B. What is meta-analysis? Evid Based Nurs. 2013;16:3-4. doi: 10.1136/eb-2012-101118.
- Meng T, Zhong Z, Meng L. Impact of spinal anaesthesia vs. general anaesthesia on perioperative outcome in lumbar spine surgery: a systematic review and meta-analysis of randomised, controlled trials. Anaesthesia. 2017;72:391-401. doi: 10.1111/anae.13702.
- Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. Nature. 2018;555:175-82. doi: 10.1038/nature25753.
- Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. J Epidemiol Community Health. 2013;67:974-8. doi: 10.1136/jech-2013-203104.
- Nikolakopoulou A, Mavridis D, Salanti G. Demystifying fixed and random effects meta-analysis. Evid Based Ment Health. 2014;17:53-7. doi: 10.1136/eb-2014-101795.
- Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011. [accessed May 2019] Available from: <http://handbook.cochrane.org>.
- Kriston L. Dealing with clinical heterogeneity in meta-analysis. Assumptions, methods, interpretation. Int J Methods Psychiatr Res. 2013;22:1-15. doi: 10.1002/mp.1377.
- Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Millbank Q. 2016;94:485-514. doi: 10.1111/1468-0009.12210.