# Frequency of use statistical methods for direct and indirect analysis in existing HTA reports in comparison to what is available – systematic review

Ewa Borowiack [1§], Anna Zapalska[1], Monika Kowalska[1], Mateusz Nikodem[2], Kamil Siedmiogrodzki[2], Jacek Walczak [2], Khalid Khan [3], Cathrine Meads[3], Ben Mol[4], Katrien Oude Rengerink [4], Shakila Thangaratinam[3]

[1]Arcana Institute, Cracow, Poland

[2]Caspolska, Cracow, Poland

[3] Queen Mary, University of London, Great Britain

[4] Academic Medical Center Amsterdam, Netherlands

[§]Corresponding author, eborowiack@inar.pl

# Abstract

### Background

There is still lack of complex, user-friendly statistical guide on choosing an appropriate statistical technique for direct or indirect comparison depending on different dataset. In order to collect information needed to prepare guide which will make statistical procedures more rational, we prepared two systematic reviews: methods potentially available to use and commonly used in HTA reports.

### Methods

Databases such as MathSciNet, Medline and Cochrane Library as well as textbooks and guidelines for systematic reviews were searched for materials regarding statistical methods which could be used for performing meta-analysis or indirect comparison. We focused on papers with mathematical descriptions of presented methods. In order to identify methods used in current practice, we performed searching in DARE Database.

### Results

Eleven statistical methods potentially available to use in systematic reviews were identified: 6 of meta-analysis (Inverse Variance, Mantel-Haenszel, DerSimonian-Laird, Peto, Bayesian and Generalized Linear Mixed Models)and 5 of indirect and mixed comparison (Bucher, Minimal Squares, Lumley, Bayesian Mixed Treatment Comparison and Generalized Linear Mixed Models). Moreover, some of these methods have different variants.

The most popular, currently used methods were Mantel-Haenszel for fixed models and DerSimonian Laird for random models. Reviewers prepared meta-analyses mostly based on RCTs (87%). In meta-analyses based on RCTs in around 40% of cases authors mentioned that to make strong conclusion more studies were needed. Employing Bayesian methods was generally very rare. For indirect comparisons the most frequent method was MTC Bayesian Model (53%).

### Conclusions

Bayesian models have essential advantage (additional data e.g. from observational studies can be included), but our analyses indicated that Bayesian approach is marginal for performing direct comparisons. However, it should be emphasized, that in the analyzed

random sample of systematic reviews studies other than RCTs were included in only 13%. More complex change in attitude to perform systematic reviews is necessary.

## Background

Proper way of analysis multi-source data is essential for obtaining reliable estimation of (relative or absolute) parameters of treatment efficacy. Poor knowledge about methods possible for use in data analysis and misinterpreting of known methods, can lead to low quality of data analyses presented in HTA reports.

The aim of the eBayesMet project supported by Leonardo Lifelong Learning Programme is to create freeware e-learning platform and user friendly guide, which will help researchers to provide data synthesis more appropriate from statistical point of view. First part of the project was to recognize all available statistical methods. There are many statistical models for meta-analyses which, depending on types of data from primary studies, are more or less credible and precise. Making the choice of optimal method for performing a meta-analysis of particular dataset is quite complex.

There are many factors which have influence on credibility of given statistical method: number of included studies, homo/heterogeneity of studies, total number of patients, total number of events, existence of some extra (prior) information.

The most common in indirect comparison is the Bucher's method which can be generalized to some cases of network analysis. However, opinions about accuracy of the Bucher's method are rather critical. There are alternative statistical constructions but, for sure, there is no universal method which suits best for every dataset. Bayesian approach can be adapted to meta-analysis problems and to all types of indirect, mixed and network comparisons. Bayesian models have one essential advantage: some extra information can be included (as a prior distribution). This extra information can be, for instance, data from observational studies, which cannot be just put together with results of RCTs studies.

In this publication we present results of two systematic reviews. The aim of the first one was to identify and describe possible statistical methods for performing meta-analyses and indirect comparisons. We restricted searching to methods designed for dichotomous (binary) data. In second review we focused on methods applied by researchers in published HTA reports. These two reviews will verify the hypothesis that even though some reliable methods exist, they are hardly ever used in HTA reports.

## Methods

Two systematic reviews were conducted – one to determine statistical methods potentially available to combine data into a meta-analysis (SR1) and second one to list methods of meta-analyses which are currently used in systematic reviews (SR2).

For SR1 following databases were searched: MathSciNet (American Mathematical Society) and The Cochrane Library Methods Database with English language restriction. All databases were searched from January 2000 to November 2009. Search terms included *Bayes, meta-analysis, odds, ratio, relative risk, risk difference, confidence interval, direct/ indirect comparison.* An additional search using keywords related to the meta-analysis for dichotomous data and indirect comparison was performed specifically in the scholar.google.pl website and on the Internet in general. Relevant materials about statistical methods were also collected from: Textbooks and Guidelines for Conducting Systematic Reviews. Bibliographies of included publications and related papers were checked for relevant articles.

Included were methodological studies where the method of meta-analysis of dichotomous data was described, in sufficient details to permit replication. The dichotomous data were analysed as odds ratios, relative risks or risk differences.

For SR2 searches were conducted in the following databases: Cochrane Database of Systematic Reviews (CDSR) and Databases of Abstracts of Reviews of Effects (DARE) from timeline January 2009 to April 2010. Separate searches were conducted for direct and indirect types of analyses. Already existing search strategies or filters for systematic reviews e.g. from The InterTASC Information Specialist' Sub-Group Search Filter Resource were used with some modifications if needed. Search terms included *meta-analysis, indirect comparison, pooling, Mantel Haenszel, Peto, Der Simonian, fixed/random effect, Bucher* and *network*. For direct comparison, a random sample of suitable 150 meta-analyses was chosen based on computerised random number list (assumption was that, if we make random sample of n=150 we can say with 95% probability that we do not avoid any statistical method which was used in minimum 2% of reviews). Different approach was used for systematic review of indirect comparison. After conducting search strategy number of identified records was below 200, so all of them were screened. Publications had to meet the following criteria: full paper available, drug or surgical procedures with quantity analysis (meta-analysis) based on systematic review, dichotomous data. Two reviewers independently extracted information about statistical issues like: type of model, statistical methods, EBM parameters. In our review the Bayesian models (used some information as prior distribution for example from

observational studies) were taken under consideration with special attention. That was the reason why some additional information about type of publications included to analysis were extracted.

# Results

### Systematic review of methods potentially available for use in meta-analysis

We examined 611 abstracts: 219 from Medline (through Pubmed) and 392 from MathSciNet Database. In figure 2 the flow chart of the selection of retrieved articles is shown. We excluded 543 articles based on abstracts, and another 17 after reviewing full texts. Additional seven articles were unable to obtain. We included one article from references list and another one publication identified on the Internet. In total eighteen articles met inclusion criteria.

We recognized six methods for conducting meta-analysis of dichotomous data: including 4 classical analytical methods (Inverse Variance, Mantel-Haenszel, DerSimonian-Laird, Peto) one Bayesian and one in-between method (Generalized Linear Mixed Models for direct comparison). Most of these methods have slightly different variants available for use. The Bayesian methods are more complicated in use, but they are the only ones, that may include some extra prior information (independent of head-to-head studies results). Note that the Bayesian approach can be applied for construction of both fixed effect model and random effects model of meta-analysis.

We found five main methods of conducting indirect or network analysis: Bucher, Minimal Squares, Lumley, Bayesian Mixed Treatment Comparison and Generalized Linear Mixed Models. Two of them are analytical, two are based on simulations and one is mixed. The most universal approach is the Bayesian Mixed Treatment Comparison, which can be applied in any scheme of network of head-to-head comparisons. The rest of methods are restricted to particular patterns of networks.

### Systematic review of methods used in existing HTA reports

The search for systematic reviews with statistical aggregations based on head-to-head studies identified 3178 citations. We chose the random sample of systematic reviews based on mathematical assumption, that if we make random sample of n=150 we can say with 95% probability that we do not avoid any statistical method which was used in minimum 2% of reviews. Separate search strategy was prepared for systematic review of indirect comparisons. Number of 192 records was obtained. Of these, 143 studies were rejected by abstracts. The

full text of the remaining 49 citations was examined in more detail. It appeared that 34 studies did not meet the inclusion criteria as described. Eventually, fifteen full texts studies were included in the analysis. Selection process is precisely described in flow chart 1.

Reviewers checked methodology of selected publications and collected information about statistical approach. Half of authors of publications (52%) took into account to use either fixed or random method, based on results of tests for heterogeneity. Researchers, who in advance assumed that only one approach will be adequate chose in 27% random method and in 21% fixed method. We need to remember that accepting limitations in methodology could affect on statistical analysis by restrain the best possible way of data synthesis.

One third (36%) of reviewers put information that fixed effect model with the Mantel-Haenszel method was chosen. Similar, but lower percentage of authors (27%) did not use fixed (only random) method or gave only simple, unspecified information that fixed effect model was adopted (26%). From those 26% (39 cases) with fixed effect model, 31 were calculated by RevMan, which means that either the Mantel-Haenszel or Inverse Variance method was used. Peto method was used in low percentage of all publications: 4% as the only fixed method and 4% as choice between fixed and Peto fixed. From whole sample, Bayesian approach was used only in one case. For random models most of reviewers (38%) provided information that DerSimonian Laird method was selected. From those 38% (57 cases), 31 were calculated by RevMan. More specified information was given by 23% reviewers, who precisely defined that it was DerSimonian Laird with the Mantel-Haenszel. In the one fifth (21%) random method was not applied and in 11% information was reduced to word ‘random’. DerSimonian Laird with Inverse Variance method was used only in 5%. In total sample, Bayesian approach was used only in 3% of cases.

The most popular in selected systematic reviews parameter for dichotomous data was risk ratio (63%). Odds ratio was calculated in one third part of the sample. Both parameters were estimated in only 3% of cases.

Meta-analyses from the sample were prepared mostly based on high quality studies like randomized controlled trials (87%). In some percentage also observational (6%) or controlled clinical trials (5%) were subject of searching. Meta-analyses using only observational studies were subject of estimation in 3%.

Authors were looking for information, if reviewers mentioned that amount of founded studies were not enough to make conclusions. Taking into considerations meta-analyses based on randomized controlled trials in 42% of cases more good quality studies were needed. In the other types of meta-analyses the amount was too low to make clear assumptions.

Of the results identified through CRD searching, six indirect methods were recognized. The most common type of indirect comparison method was Mix Treatment Comparison – Bayesian Model. The proportion of identified surveys was relatively higher compared to the incidence of other statistical methods and was equal 53%. In addition, collected data indicate 13% share of the two statistical methods in the total number of identified techniques of indirect comparison. These methods were: Bucher's method and Network Logistic Regression.

Among 15 identified systematic reviews the proportion of analyses that estimate relative effect for dichotomous endpoints was equal for both parameters odds ratio and relative risk (47%). Selection criteria in all identified reviews contained restriction to randomized clinical trials, therefore observational studies were automatically excluded from potential analysis. Noticed that 40% of the reviews contain information indicating the need for additional studies, observational studies could be good source of supplementary information.

In case of small sample size from randomized studies including observational trials should be consider as possibility of increasing the statistical power. Observational trials can help to examine outcomes that are ignored or poorly evaluated in efficacy trials, either because the trials are narrow in scope or because they have a short follow-up period. Including observational trials into indirect comparison can provide better evidence about an intervention's benefits and risks in everyday practice.

# Discussion

***Main Findings***

There is a variety of statistical methods potentially available for data aggregation. The majority of meta-analyses prepared in current practice are conducted using direct comparison. Around half of these used either fixed or random method, based on results of tests for heterogeneity; and a quarter of used only one approach without consideration for heterogeneity. The Mantel-Haenszel fixed effects model and the DerSimonian Laird random effects model were the most frequently used. Bayesian methods were infrequently/rare. The most often calculated dichotomous parameter was risk ratio. Meta-analyses were rarely conducted using only observational studies. In nearly half of the reviews the conclusion was that more studies were needed.

***Strength and weaknesses***

Our search of present systematic reviews focused on high quality databases (DARE, HTA and NHS EED). The risk of missing some relevant publications was limited, because these databases are assembled using systematic approach to searching in electronic bibliographic databases, hand-searching, scanning grey literature. Glenny et al. have pointed out that in this situation a comprehensive search of the research literature is unnecessary. We randomly sampled for reviews of direct comparisons at a prespecified statistical power and oversampled for indirect comparisons to ensure that we could comprehensively identify all available methods.

***Interpretation of findings***

We identified 4 systematic reviews presented findings of a survey of frequency of use of indirect comparisons in published systematic reviews. Systematic review with similar questions and aims connected with statistical methods for direct comparison was not recognized.

In review Song 2009 authors included only reports published between 2000 and 2007 in which the indirect comparisons were explicit in their titles and abstracts. Searching was performed in one database (MEDLINE by PubMed). Some differences in inclusion criteria between this and our review appeared, for example connected with type of intervention (restriction in our review) or type of methods used to performed analysis (naïve comparison without statistical calculation was not included to our review). Because in our review in contrast to Song 2009 we focused on systematic reviews prepared nowadays we were able to compare changes over the years in statistical approaches to indirect comparison.

The most commonly (more than 50%) used approach was the adjusted indirect comparison using classic frequents methods. According to results obtain in our review this method was used only in 13% reports published between 2009-2010. In contrast to our review more complex methods (including network or Bayesian hierarchical meta-analysis and mixed treatment comparison methods) were performed less frequently (20% vs. 66%).

Another report, Edwards 2009 was developed in order to perform assessment of different methodologies used to indirectly compare health technologies based on meta-analyses of randomized controlled trials. Authors conducted systematic searching in the following databases: the Cochrane Database of Systematic Reviews (CDSR); the Cochrane Methodology Register (CMR); the Excerpta Medica Database (EMBASE); the Index Medicus database (MEDLINE). Searching was completed in July 2007. Authors identified sixty-two reviews contained indirect comparisons of medical technologies. Mixed treatment comparison

or network meta-analysis was used in 11% of all identified systematic reviews in comparison to results obtain in our reviews – 53%. The remaining identified papers used following approaches: comparing point estimates (2%); comparing 95% confidence intervals (42%); performing statistical tests on summary estimates (13%); indirect comparison using a single common comparator (32%).

Methodologically, the third systematic review prepared by Glenny 2005 used similar approaches, for example database DARE (1994 to March 1999) was used as main sources of data. Twenty-three reviews included only indirect comparisons (15 used an adjusted method of comparison, and eight performed a naive indirect comparison).

Review of Li 2009 was designed in order to identify potential problems of commonly used statistical methods in indirect comparison. Authors conducted searching in 3 databases: The Cochrane Library Issue 4, 2007, MEDLINE (Dec 2007) and Chinese Biomedical Disk (CBM, Dec 2007). Fifty-eight papers with indirect comparison were identified: almost half of them performed Adjusted indirect comparison, 3 used naïve method, 6 used meta-regression, 4 used mixed models and nine used Bayesian methods. Differences between results obtained in our reviews are strictly connected with differences in the time span of the search.

The variety of approaches and methods for conducting meta-analyses and indirect comparisons is very large. For any model and any data set there is at least one method which can be applied and provides reliable results of calculations. Of course, there are extreme situations (e.g. no events in treatment or control group) where no precise and credible estimation is possible, but it is caused rather by non-sufficient data than the fact that there is no proper statistical method. Formally, it is possible to conduct analysis of data even for small data sets and even if there are no events – but one has to be more careful on interpreting results (accuracy is lower and a chance of statistically significant result is small). The problem is that there are several methods which (formally) can be used for one given data set, and researchers have to choose the one which is the most suitable for a particular case. The most flexible methods are those, which use the Bayesian approach. The ability of including extra information (data from observational studies, experts opinions, etc.) as a prior distribution is not possible in other methods. Moreover, Bayesian Mixed Treatment Comparison can be implemented for any type of network of direct comparisons. On the other hand, providing calculations with Bayesian approach is not very easy– it requires special, advanced (but freeware) software and results are obtained by simulating properly defined variables. Hence, if in particular case using Bayesian approach gives us no advantage, we recommend choosing one of analytical method, which makes calculation much simpler.

***Conclusions***

Our systematic review demonstrated a wide range of approaches and methods for aggregating data in current practice. There is a need for evaluating relative precision and credibility of the statistical methods of meta-analysis.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

participated in the conception and design of the study. Data collection, abstraction and analysis was carried out by and. and prepared the original draft of the manuscript, and all authors reviewed and critically revised the original and subsequent manuscript drafts and approved the final manuscript.

## Acknowledgements

This publication has been funded with support from the European Commission under the Lifelong Learning Programme.

## References

1. CRD's guidance for undertaking systematic reviews in health care. Available from http://www.york.ac.uk/inst/crd/
2. Edwards SJ, Clarke MJ, Wordsworth S, Borrill J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. Int J Clin Pract. 2009 Jun;63(6):841-54.
3. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, Bradburn M and Eastwood AJ. Indirect comparisons of competing Interventions. Health Technology Assessment 2005; Vol. 9: No. 26
4. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration. Available from www.cochrane-handbook.org, 2008.

5. Jing Li, Lin Wang, Deying Kang, Guanjian Liu, Youping Li. Indirect comparisons of competing interventions: a systematic review identified possible methods and current practice Abstract Cochrane Colloquium Abstracts 2009

6. Song F, Loke Y K, Walsh T, Glenny AMe, Eastwood A J, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ. 2009 Apr 3;338:b1147

7 H.C. Bűcher, G.H. Guyatt, L.E. Griffith, S.D. Walter, The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trials, J. Clin. Epidemiol. Vol. 50 (1997), No 6, pp. 683-691.

8 G. Woodworth „Biostatistics, a Bayesan Introduction", John Wiley & Sons, Incorporated (2004)

9 M. Borenstein, L.V. Hedges, J.P.T. Higgins, H.R. Rothstein „Introduction to Meta-Analysis", John Wiley & Sons, 2009.

10 J.P.T. Higgins, S. Green (editors), „Cochrane Handbook for Systematic Reviews of Interventions " Version 5.0.2. The Cochrane Collaboration, 2009.

11 Madhukar Pai, Michael Mcculloch, Jennifer D. Gorman, Nitika Pai, Wayne Enanoria, Gail Kennedy, Prathap Tharyan, John M. Colford, Jr. „Systematic Reviews And Meta-Analyses: An Illustrated, Step-By-Step Guide", The National Medical Journal Of India, Vol. 17, No. 2, 2004

# Figures

**Figure 1 - Flow chart of study selection process – methods used in existing HTA reports**

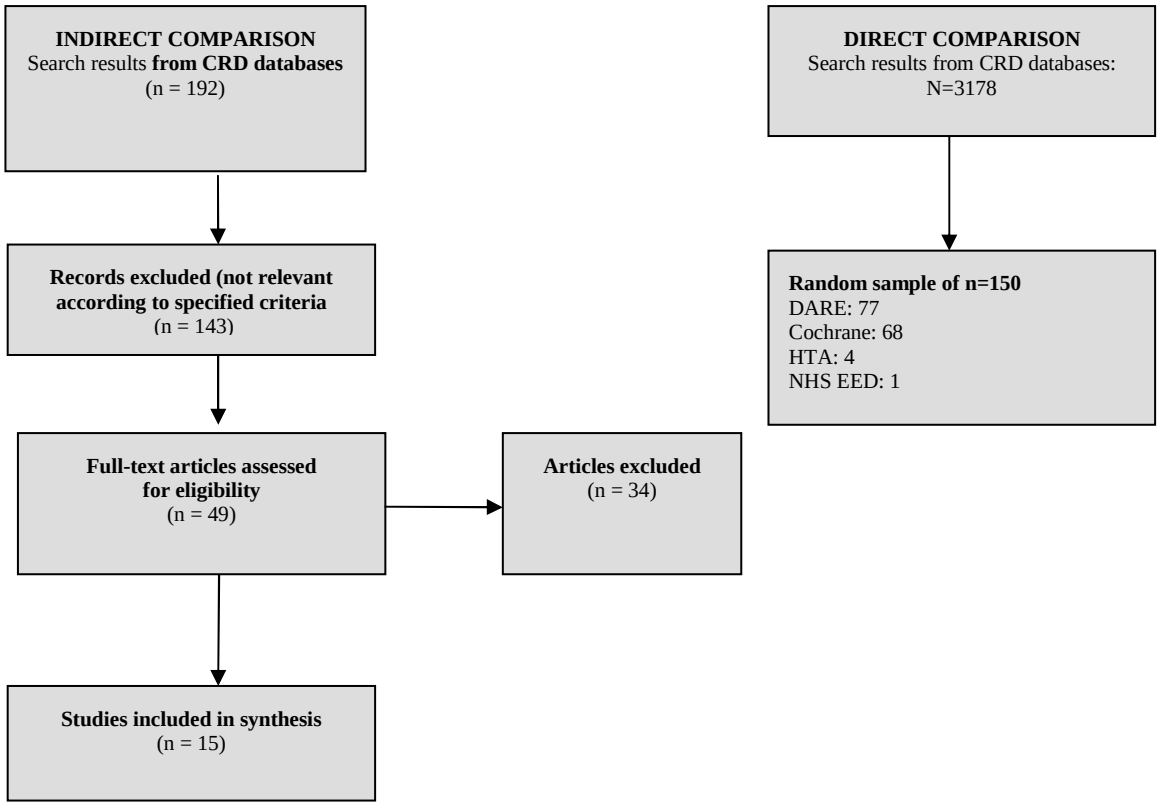

**Figure 2 - Flow chart of study selection process – methods potentially available to use in HTA reports**

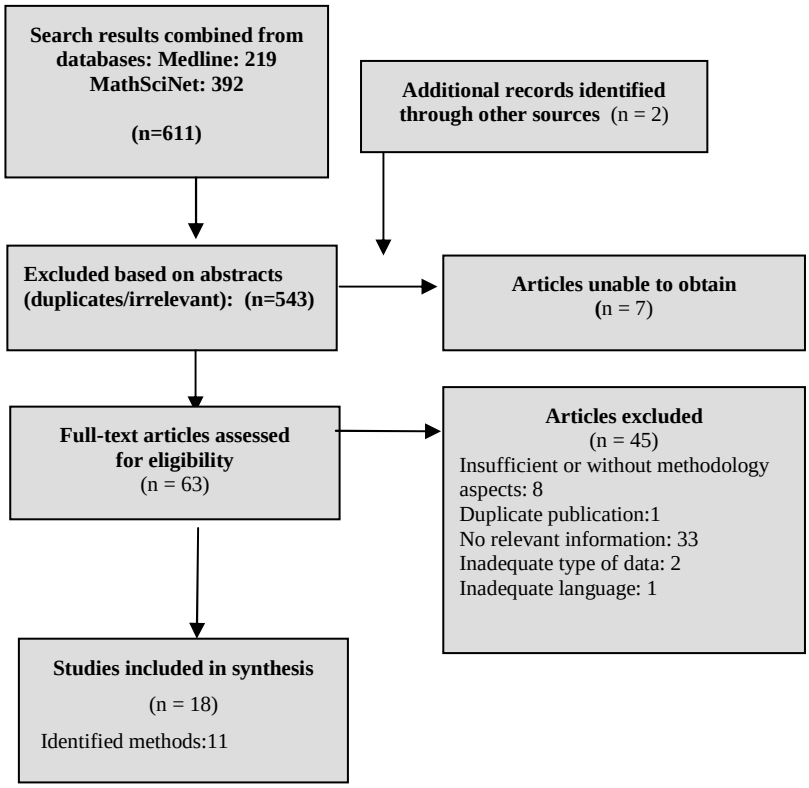

Figure 3 - Statistical methods for direct comparison

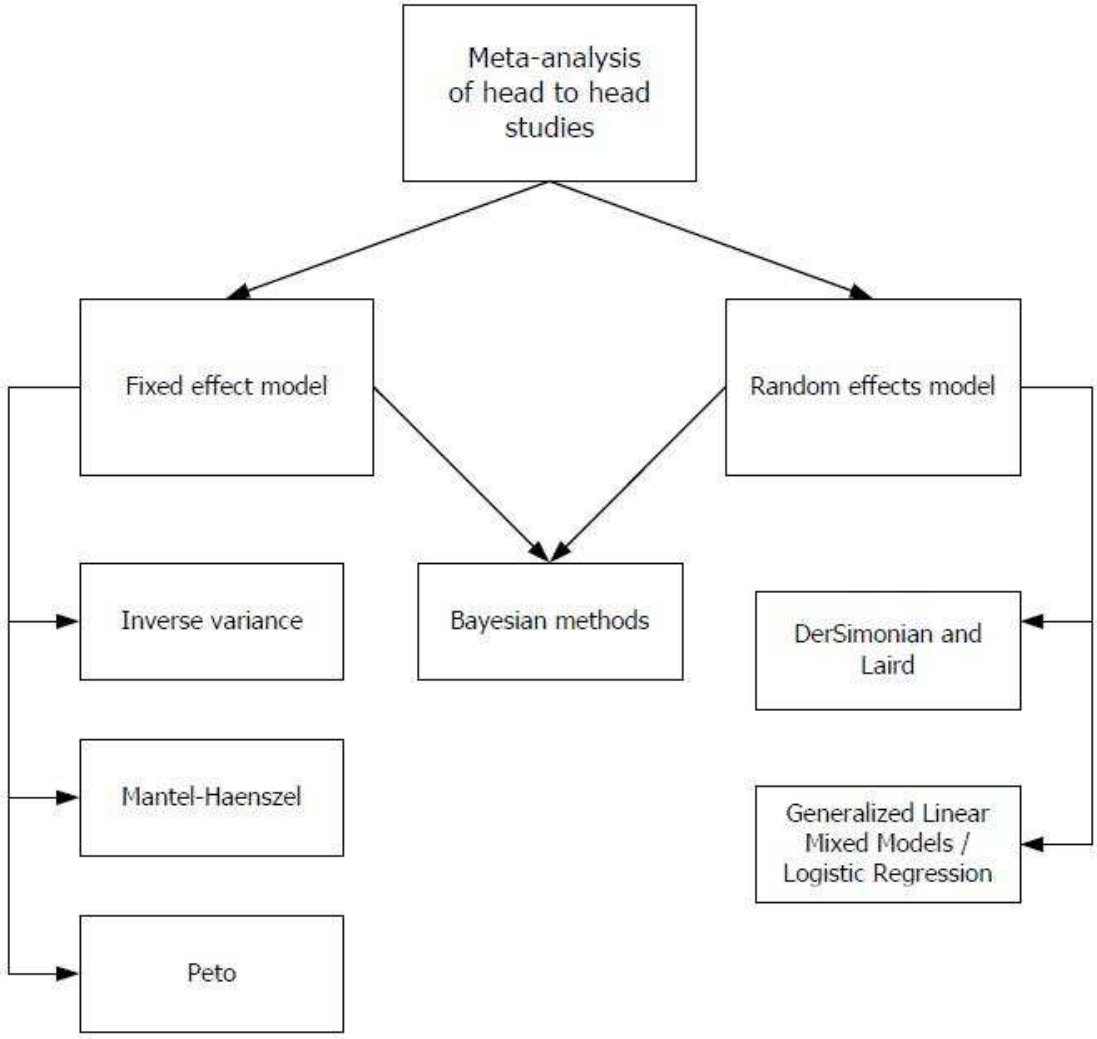

Figure 4 - statistical methods for indirect comparison