



Network meta-analysis as a tool in clinical practice guidelines

Stavros A. Antoniou¹ · Dimitrios Mavridis^{2,3} · Sofia Tsokani² · Salvador Morales-Conde⁴ · the EAES Guidelines Subcommittee

Received: 21 August 2022 / Accepted: 19 November 2022 / Published online: 1 December 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

In this issue, the European Association for Endoscopic Surgery (EAES) and its Guidelines Subcommittee report on a rapid guideline for the management of choledocholithiasis [1]. Rapid guidelines or rapid recommendations are short guidelines, addressing one up to a few prioritized clinical questions. With rapid guidelines, efforts are invested in the rigor of development, both at the level of evidence retrieval and appraisal, as well as in the process of generating recommendations based on this evidence. Through a structured approach and in line with predefined methodological standards, an international, interdisciplinary panel provided a weak recommendation in favor of preoperative or intraoperative endoscopic retrograde cholangiopancreatography (ERCP), or laparoscopic common bile duct exploration (LCBDE).

A weak or conditional recommendation means that most patients would opt for the proposed intervention after informed decision making, and only a minority would not. For healthcare professionals and policy makers, it means that decision aids are likely needed to help individual patients make decisions consistent with their values and preferences, depending on the setting, resources, surgical expertise, and other parameters.

This guideline is somewhat unique, in that, it is one of a few guidelines in the field of surgery that applies a relatively novel evidence synthesis method, called network meta-analysis. Traditional (pairwise) meta-analysis is a statistical method that synthesizes the results of different studies; this allows a more precise estimate of the actual comparative effect between two interventions. However, it falls short in comparing multiple interventions simultaneously, rather than it compares two interventions at a time. Network meta-analysis can compare multiple interventions within the same model (in this guideline, LCBDE versus preoperative, intraoperative, or postoperative ERCP). This allows simultaneous comparison and prioritization of one out of many interventions [2, 3].

Network meta-analysis allows synthesizing studies comparing two or more interventions as long as these interventions form a connected network. In the example given in Fig. 1, LCBDE is being directly compared with preoperative ERCP, intraoperative ERCP, and postoperative ERCP (indicated by connecting arrows), meaning that there are randomized trials comparing LCBDE with either of these interventions (direct evidence, deriving from standard pairwise meta-analysis). But LCBDE is also indirectly compared with, e.g., intraoperative ERCP via a common comparator, which is preoperative ERCP (indirect evidence, indicated by dashed lines). This is a connected network as we can follow a path from any intervention to any of the remaining interventions.

Using sophisticated statistical methods, direct and indirect evidence are combined and provide the mixed, or network evidence. The main output of the network meta-analysis is the network treatment effects (e.g., odds ratio with the corresponding confidence interval), which show the relative efficacy between any pair of competing interventions. Another attractive feature of network meta-analysis is ranking metrics. One of the most common is the so-called P-score, a probability estimate for each intervention being better than the remaining interventions [4].

Collaborators of the EAES Guidelines Subcommittee are listed in “Acknowledgement”.

✉ Stavros A. Antoniou
stavros.antoniou@hotmail.com

¹ Department of Surgery, Papageorgiou General Hospital, Thessaloniki, Greece

² Department of Primary Education, School of Education University of Ioannina, Ioannina, Greece

³ Paris Descartes University, Sorbonne Paris Cité, Faculté de Médecine, Paris, France

⁴ Unit of Innovation in Minimally Invasive Surgery, Department of General and Digestive Surgery, University Hospital “Virgen del Rocío”, University of Sevilla, Seville, Spain

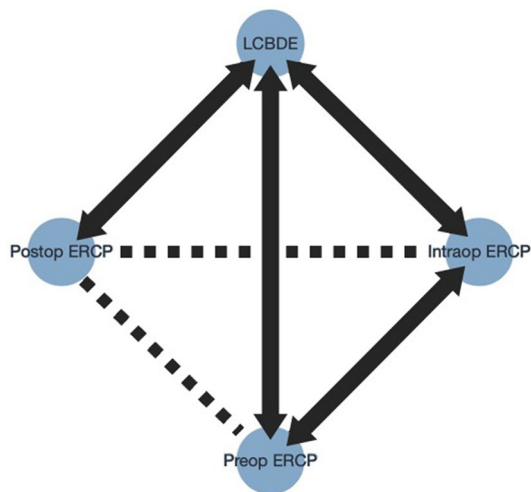


Fig. 1 Network plot for of interventions. Continuous arrows indicate direct comparisons; dashed lines indicate lack of direct (indirect) comparisons

A feature of network meta-analysis is that it provides more precise comparative effect estimates (how much better or worse is the intervention compared to the comparator), compared to a standard pairwise meta-analysis [5]. As an example, for the outcome ‘major complications’ (Clavien-Dindo ≥ 3), the pairwise (standard meta-analysis) 95% confidence interval of absolute effect difference between preoperative and intraoperative ERCP was between 9 fewer to 493 more patients. This means that the true comparative effect of preoperative versus intraoperative ERCP on major complications is somewhere between 9 fewer to 493 more patients; i.e., 9 fewer patients to 493 more patients will experience a major complication with preoperative ERCP. But the corresponding network estimates were between 0 and 62 more patients [1]. Network meta-analysis reduced the 95% confidence interval (which corresponds to the precision: the narrower the confidence interval, the higher the precision) by 88% compared to pairwise meta-analysis. These estimates allow us to be more confident about the actual comparative effect of interventions, which affects the panel’s decisions on which intervention should be recommended or not. Furthermore, it allows healthcare professionals’ and patients’ decision making based on more precise information.

The validity of the results of network meta-analysis lies on the assumption that patients in all intervention arms and in all trials have similar characteristics. In brief, we require that the distribution of effect modifiers is similar across treatment comparisons. That entails that potential effect modifiers (e.g., the proportion of patients with acute cholangitis, which would adversely affect outcomes) should be defined a-priori. Though this assumption can be approximated statistically, it is evaluated mainly clinically and epidemiologically [3, 6]. In this guideline, there was no substantial

variation in operative and interventional techniques, and most authors attempted transcystic stone extraction before choledochotomy. Furthermore, acute cholangitis and/or pancreatitis were exclusion criteria in 45%, 55%, 35%, and 33% of the preoperative ERCP, intraoperative ERCP, LCBDE, and postoperative ERCP cohorts, suggesting that these effect modifiers likely did not affect transitivity.

The summary evidence provided by the statistical analyses is subjected to rigorous appraisal before it can be used. For this purpose, the EAES Guidelines Subcommittee uses two evidence appraisal systems; the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) system, and the CINeMA (Confidence in Network Meta-analysis) system.

Different trials carry different risks of bias. CINeMA, through a semi-automated platform, weighs the contribution of the risk of bias of each study to the overall risk of bias, according to the contribution of this trial to each network effect estimate [7]. This method allows a modest assessment of the overall risk of bias for each comparison.

Moreover, factors other than risk of bias may contribute to the certainty of evidence. These are imprecision (how precise are the results, defined primarily by the confidence interval), heterogeneity (whether the results of different trials agree with each other), incoherence (whether the results of direct and indirect comparisons agree with each other), publication bias (whether any published studies have been missed, or whether there are any unpublished trials), and indirectness (whether the patients, interventions and outcomes of the trials fit within guideline-specific criteria). These parameters are summarized by the CINeMA and the GRADE systems to appraise the certainty of the evidence [7, 8].

Following a thorough appraisal of the evidence certainty, the evidence summaries can be used to inform the development of recommendations. However, going from evidence to recommendations is a challenging process in the context of multiple interventions. The GRADE approach includes 7 pillars (domains) in the so-called evidence-to-decision framework: balance between benefits and harms, certainty of the evidence, use of resources, patients’ values and preferences, applicability, feasibility, and equity [8]. In the previous steps, network meta-analysis has provided a comprehensive overview of the evidence on benefits and harms, and the certainty of this evidence. These are only two of the domains that inform the development of recommendations. The guideline development group (in our case, an interdisciplinary panel of surgeons, gastroenterologists, patient representatives, systematic reviewers, statisticians, and guideline methodologists) discusses this evidence, along with conceptual differences among the interventions in each of these domains. At the end of this process, the guideline development group has gained an overview of the developed

	Intraoperative ERCP	Preoperative ERCP	LCBDE	Postoperative ERCP
Benefits, desirable effects	Small benefits	Trivial/no benefits	Varies	Don't know
Harms, undesirable effects	Trivial/no harms	Small harms	Varies	Don't know
Certainty of the Evidence	Low	Very low	Low	Very low
Values and preferences	Substantial variability is expected or uncertain	Substantial variability is expected or uncertain	Substantial variability is expected or uncertain	Substantial variability is expected or uncertain
Resources	Varies	Moderate cost	Moderate savings	Moderate cost
Equity	Probably reduced equity	Probably increased equity	Varies	Probably increased equity
Acceptability	Probably not acceptable	Acceptable	Probably acceptable	Not acceptable
Feasibility	Varies	Feasible	Feasible	Feasible

Fig. 2 Summary judgements in the evidence-to-decision framework

GRADE evidence-to-decision framework (Fig. 2), which usually makes the consensus to recommend either intervention straightforward. In the discussed EAES guideline, a unanimous consensus was achieved across recommendations on the first Delphi round.

The landscape of the guideline development methodology has changed dramatically over the past few years. The European Association for Endoscopic Surgery and its Guidelines Subcommittee have embraced the latest advances in the field of medical statistics and guideline development methods, aiming to provide pertinent, trustworthy recommendations to improve patient care and experience. Surgical Endoscopy has supported this approach, hoping our readers will value EAES guidelines developed with rigorous, transparent and evidence-based methodologies.

Acknowledgement Collaborative authorship: EAES Guidelines Subcommittee: Monica Ortenzi, Nereo Vetoretto, Francesco Maria Carrano, Silviu Daniel Preda, Thanjakumar Arulampalam.

Funding This editorial was not funded.

Declarations

Disclosure Stavros A. Antoniou, Dimitrios Mavridis, Sofia Tsokani, Monica Ortenzi, Nereo Vetoretto, Francesco Maria Carrano, Daniel Silviu Preda, Thanjakumar Arulampalam and Salvador Morales-Conde declare no direct (financial) conflicts of interest related to this work. Stavros A. Antoniou is a certified guideline methodologist (INGUIDE level 2 certification) and a guideline methodologist credentialing instructor. He is chair of the EAES Guidelines Subcommittee. He served as a methodologist and co-chair of the guideline that is discussed in this editorial. Dimitrios Mavridis is the author of several publications on the statistical methods discussed herein. Authors Stavros A. Antoniou, Dimitrios Mavridis, Sofia Tsokani, Monica Ortenzi, Nereo Vetoretto, Francesco Maria Carrano, Daniel Silviu Preda, Than-

jakumar Arulampalam and Salvador Morales-Conde have no financial conflicts of interest to disclose.

Ethical approval Not applicable.

Patient consent Not applicable.

IRB approval Not applicable.

References

1. Boni L, Huo B, Alberici L, Ricci C, Tsokani S, Mavridis D, Amer YSAD, Andreou A, Berriman T, Donatelli G, Forbes N, Kapiris S, Kayaalp C, Kylänpää L, Parra-Membrives P, Siersema PD, Black GF, Antoniou SA (2022) EAES Rapid Guideline: Updated systematic review, network meta-analysis, CINeMA and GRADE assessment, and evidence-informed European recommendations on the management of common bile duct stones. *Surg Endosc* 36(11):7863–7876. <https://doi.org/10.1007/s00464-022-09662-4>
2. Salanti G (2012) Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 3(2):80–97. <https://doi.org/10.1002/jrsm.1037>
3. Seitidis G, Nikolakopoulos S, Hennessy EA, Tanner-Smith EE, Mavridis D (2022) Network Meta-Analysis Techniques for Synthesizing Prevention Science Evidence. *Prev Sci* 23(3):415–424. <https://doi.org/10.1007/s11121-021-01289-6>
4. Rucker G, Schwarzer G (2015) Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol* 15:58. <https://doi.org/10.1186/s12874-015-0060-8>
5. Nikolakopoulou A, Mavridis D, Furukawa TA, Cipriani A, Tricco AC, Straus SE, Siontis GCM, Egger M, Salanti G (2018) Living network meta-analysis compared with pairwise meta-analysis in comparative effectiveness research: empirical study. *BMJ* 28(360):k585. <https://doi.org/10.1136/bmj.k585>. PMID:29490922;PMCID:PMC5829520

6. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP (2014) Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE* 9(7):e99682. <https://doi.org/10.1371/journal.pone.0099682>. PMID:24992266;PMCID:PMC4084629
7. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, Salanti G (2020) CINeMA: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Med* 17(4):e1003082. <https://doi.org/10.1371/journal.pmed.1003082>. PMID:32243458;PMCID:PMC7122720
8. Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochweg B, Hazlewood GS, Alhazzani W, Mustafa RA, Murad MH, Puhan MA, Schünemann HJ, Guyatt GH, GRADE Working Group (2018) *J Clin Epidemiol*. 93:36–44. <https://doi.org/10.1016/j.jclinepi.2017.10.005>
9. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, Elbarbary M, Post P, Nasser M, Shukla V, Jaeschke R, Brozek J, Djulbegovic B, Guyatt G (2013) GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation’s direction and strength. *J Clin Epidemiol*. 66(7):726–35. <https://doi.org/10.1016/j.jclinepi.2013.02.003>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.