Corresponding author:
Alessandro De Cassai, M.D.
UOC Anesthesia and Intensive Care Unit,
University Hospital of Padua, Gallucci St,
35121 Padua, Italy
Tel: +39-0498213090
Fax: +39-0498213091
Email: alessandro.decassai@aopd.vene to.it
ORCID: https://orcid.org/0000-0002-9773-1832

Explanation of trial sequential analysis: using a post-hoc analysis of meta-analyses published in Korean Journal of Anesthesiology

Alessandro De Cassai[1], Martina Tassone[2], Federico Geraldini[1], Massimo Sergi[1], Nicolò Sella[1], Annalisa Boscolo[1], Marina Munari[1]

UOC Anesthesia and Intensive Care Unit, [1]University Hospital of Padua, [2]Department of Medicine-DIMED, University of Padua, Padua, Italy

Background: Trial sequential analysis (TSA) is a recent cumulative meta-analysis method used to weigh type I and II errors and to estimate when the effect is large enough to be unaffected by further studies. The aim of this study was to illustrate possible TSA scenarios and their significance using meta-analyses published in the *Korean Journal of Anesthesiology* (*KJA*) as working material.

Methods: We performed a systematic medical literature search for meta-analyses published in the KJA. TSA was performed on each main outcome, estimating the required sample size on the calculated effect size for the intervention, considering a type I error of 5% and a power of 90% or 99%.

Results: Six meta-analyses with a total of ten main outcomes were included in the analysis. Seven TSAs confirmed the results of the meta-analyses. However, only three of them reached the required sample size. In the two TSAs, the cumulative z-lines were not statistically significant. One TSA boundary for effect was reached with the 90% analysis, but not with the 99% analysis.

Conclusions: In TSA, a meta-analysis pooled effect may be established to assess if the cumulative sample size is large enough. TSA can be used to add strength to the conclusions of meta-analyses; however, pre-registration of the TSA protocol is of paramount importance. This study could be useful to better understand the use of TSA as an additional statistical tool to improve meta-analysis quality.

Keywords: Anesthesia; Medical education; Meta-analysis; Review; Sample size; Statistics.

Introduction

Traditional meta-analyses are only able to examine the pooled effect size rather than to evaluate whether the number of participants and the corresponding number of trials in a meta-analysis are sufficient to draw any conclusions. Moreover, the use of the traditional 95% CI or the 5% statistical significance threshold will lead to too many false-positive conclusions (type I errors) and too many false-negative conclusions (type II errors) [1].

Trial sequential analysis (TSA) is a recently described cumulative frequentist meta-analysis method [2] used to weigh type I and II errors and to estimate when the effect is large enough to unlikely be affected by further studies [3,4]. While TSA is based on frequentist thinking as it is founded on P value and type I and type II error methods, it in-

corporates elements of Bayesian thinking. Indeed, the calculated sample size in TSA is related to the pooled effect estimated in a meta-analysis.

TSA generates a graphical outcome divided into four areas by four lines: "benefit," "harm," "inner wedge," or "non-statistically significant," representing a statistically significant result for the first two areas ("benefit" and "harm") and a strong evidence that further studies will hardly be able to change the no-effect results for the "inner wedge" area (Fig. 1). Lying in the "non-statistically significant" area means that further studies are needed for a conclusion on the analyzed topic. The cumulative z-statistic line is drawn on this chart by adding the included studies with a chronological criterion, with the last study representing the end of the line and the area ("benefit," "harm," "inner wedge," or "non-statistically significant") [5].

The aim of this study was to illustrate the possible scenarios and possible significance of TSA using meta-analyses published in the *Korean Journal of Anesthesiology* (*KJA*) as working material.

Materials and Methods

We performed a systematic search of the medical literature following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Statement Guidelines for the identification, screening, and inclusion of articles. The search was performed by two researchers (ADC and MT) in close collaboration with the rest of the research team.

Search strategy

The search was performed on May 10, 2021, using the search tool in the KJA site and using the following terms: "meta-analysis," "metaanalysis," "meta analysis." In our search, we did not apply any restrictions on publication type or date, language or status.

Study selection

Two researchers (ADC and MT) independently screened the titles and abstracts of the identified papers to select those that were
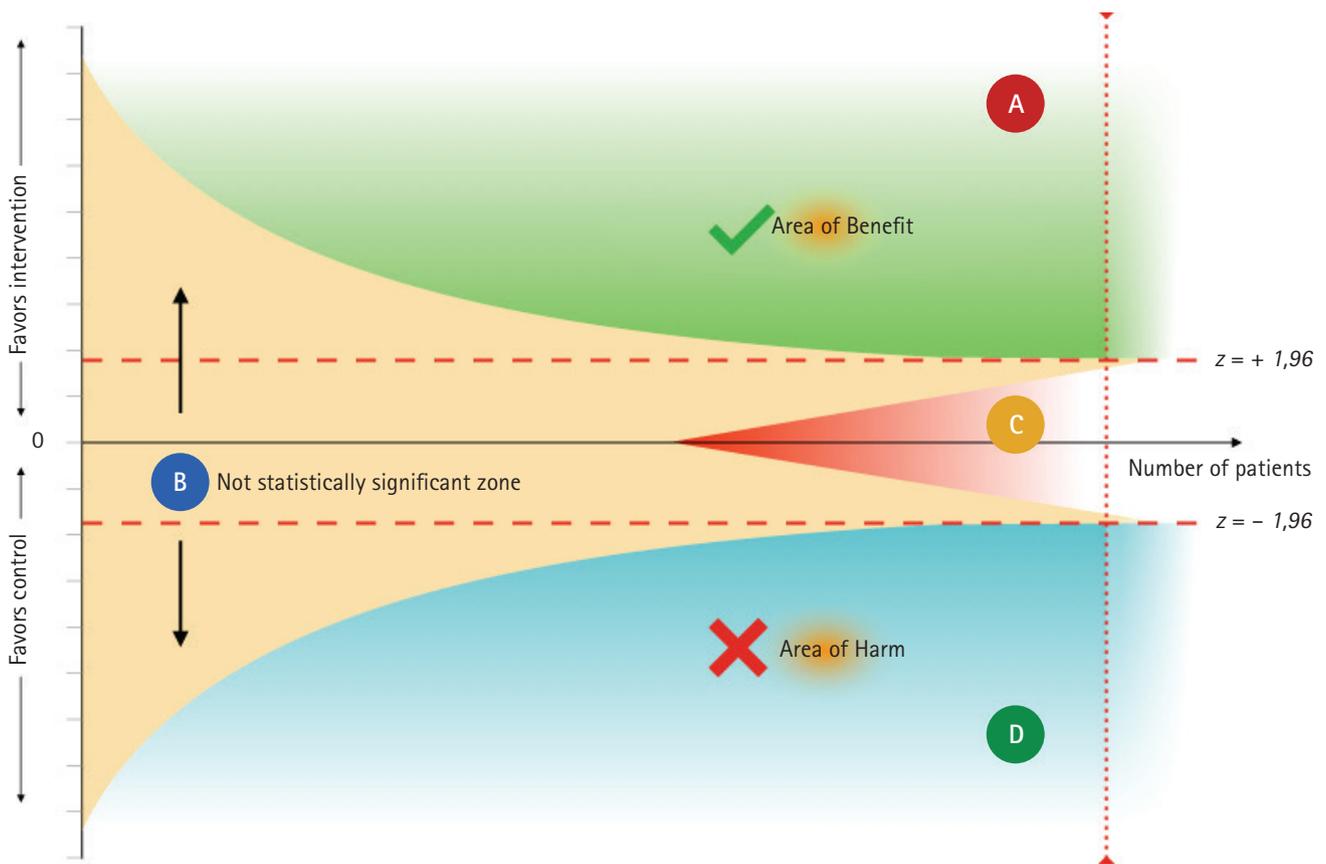
Fig. 1. Graphical representation of the trial sequential analysis (TSA) outcome. A: favors intervention (benefit), B: non-statistically significant, C: inner wedge, D: favors control (harm).

relevant. Only meta-analyses were considered eligible for analysis.

Data extraction and data retrieval

After identifying those studies meeting the inclusion criteria, two researchers (FG and AB) independently reviewed and assessed each of the included studies. The following information was collected: first author, year of the study, total number of patients per group, registration number, main outcome, and data for intervention and control relative to the main outcome.

If the main outcome was not clearly stated, it was retrieved by examining the registered protocol or by contacting the main author of the paper.

Statistical methods

TSA was performed on the main outcome for each paper using TSA software (Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen). The effect measure and model (mean difference, odds ratio, relative risk, risk difference, or Peto odds ratio) were used. A fixed effects model, random effects model using the DerSimonian–Laird method, random effects model using the Sidik–Jonkman method, or random effects model using the Biggerstaff–Tweedie method was selected according to the outcome measure and model. No continuity correction was applied in the case of a zero event. We estimated the required sample size on the calculated effect size for the intervention, considering a type I error of 5% and a power of 90%; benefit, harm, and inner wedge boundaries were drawn using the O'Brien–Fleming spending function.

Moreover, a more conservative approach, performing a second TSA with a type I error of 5% and a power of 99% was performed for each main outcome. This post-hoc conservative approach allowed us to assess whether the data provided convincing evidence of the true effect.

Results

We identified 11 papers [6–16] in our initial search (Table 1). However, four of them were excluded [6–9] because they were statistical rounds; the remaining seven were clinical meta-analyses. One of the meta-analyses [10] did not have sufficient information to perform a TSA and was therefore excluded, leaving six papers for the final analysis [11–16] (Fig. 2).

The topics of the meta-analyses were as follows: curare side effects [11], regional anesthesia [12,16], postoperative efficacy of ibuprofen [14], postoperative shivering [13], and postoperative

nausea and vomiting [15]. Notably, only two of them had a pre-registered protocol [12,14]. Four papers [11–14] had two main outcomes, and for this reason, a total of 10 TSAs were performed.

Choi et al. [11] evaluated the effect of pretreatment with lidocaine or opioids opioid pretreatments in the incidence of rocuronium-induced withdrawal movement. For both outcomes, the cumulative z-score line crossed the line to reach the required sample size in both the 90% and 99% analyses (Figs. 3 and 4).

Bailey et al. [12] evaluated the cumulative opioid consumption at 48 hours after midline laparotomy, comparing, on the one hand, continuous peripheral nerve blocks and multimodal analgesia and, on the other hand, continuous peripheral nerve blocks and epidural analgesia. In the TSA, the cumulative z-score line crossed the benefit boundary, but did not reach the required sample size for the outcome relative to the continuous peripheral nerve block in either the 90% or the 99% analysis (Fig. 5). For the other outcome, the cumulative z-score line did not reach any boundary and remained in the zone that is "non-statistically significant" area (Fig. 6).

Min et al. [13] evaluated meperidine and clonidine for the prevention of postoperative shivering. A TSA of the meperidine outcome revealed that the cumulative z-line crosses the 90% but not the 99% boundary for benefit (Fig. 7), while a TSA of the clonidine outcome revealed that the cumulative z-line crossed both the 90% and 99% boundaries for benefit without reaching the required sample size (Fig. 8).

The effect of a single dose of ibuprofen was evaluated by Kim et al. [14] on both postoperative opioid consumption and pain. While the cumulative z-line does not cross the 90% power boundary for effect but lies immediately below that in the opioid consumption outcome (Fig. 9), it crosses both the 90% and 99% boundary for benefit without reaching the required sample size in the analysis relative to pain scores (Fig. 10).

Kim et al. [15] evaluated the efficacy of ramosetron in preventing postoperative nausea and vomiting. A TSA revealed that the cumulative z-line crossed the boundary for benefit in both the
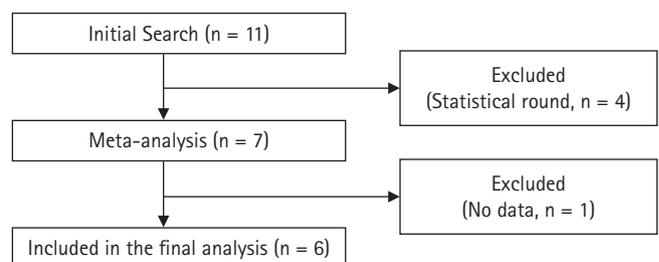
Fig. 2. Flow chart of study inclusion.
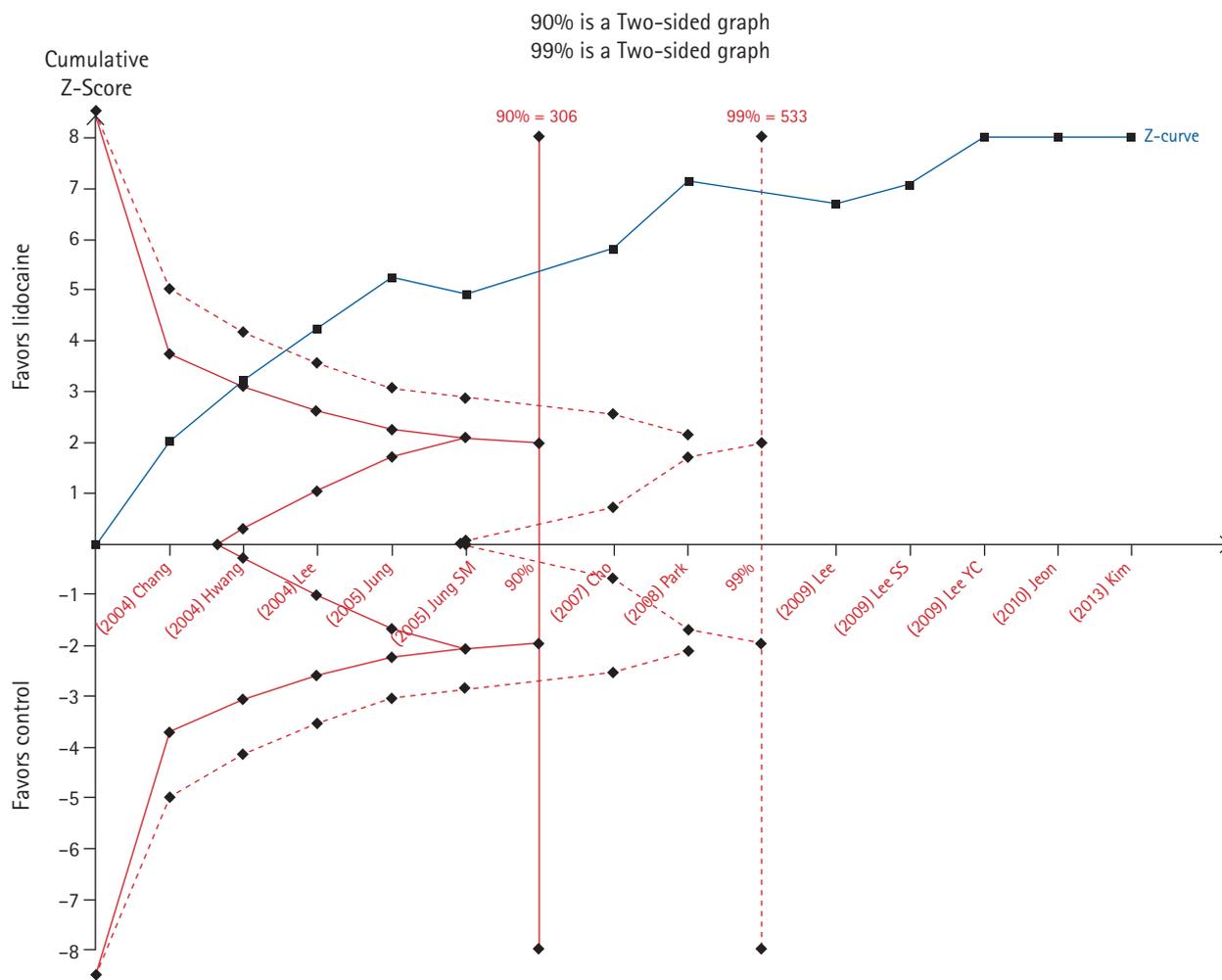
Fig. 3. Trial sequential analysis (TSA) of the effect of lidocaine in reducing rocuronium-induced withdrawal movement [11].

90% and 99% analyses, without reaching the required sample size (Fig. 11).

Another study by the same group of authors [16] investigated the pharmacological efficacy of lidocaine/tetracaine patches and peels on pain (Fig. 12). In the post-hoc analysis, the cumulative z-line crossed the boundary for benefit and the required sample size for both the 90% and 99% analyses.

Discussion

A TSA analyzes the cumulative evidence in a meta-analysis. Its output is represented by a cumulative z-line score that may lie in one out of four areas: benefit (labeled A in Fig. 1), harm (labeled D in Fig. 1), non-statistically significant (labeled B in Fig. 1), and inner wedge (labeled C in Fig. 1).

A pooled effect in favor of the intervention (benefit) or in favor of the control (harm), or the absence of any effect (inner wedge), may be established to assess if the cumulative sample size is large enough. On the contrary, when the cumulative z-line lies in the area that is not statistically significant, further studies with an increase in the overall sample size are deemed necessary.

Confirmation of the meta-analysis pooled effect

Seven out of ten TSAs confirmed the results of meta-analyses. However, only in three of them (Figs. 3, 4, and 12) the required sample size was reached. These TSAs suggest that the result is definitive and that other randomized controlled trials are unlikely to modify the effect on the outcomes.

On the contrary, in four TSAs (Figs. 5, 8, 10, and 11), the cumulative z-line, after crossing the boundary for effect, did not reach the required sample size. These TSAs suggest that, although the pooled effect is statistically significant, with regard to sample size, the result is not definitive, and future studies are necessary to be conclusive.
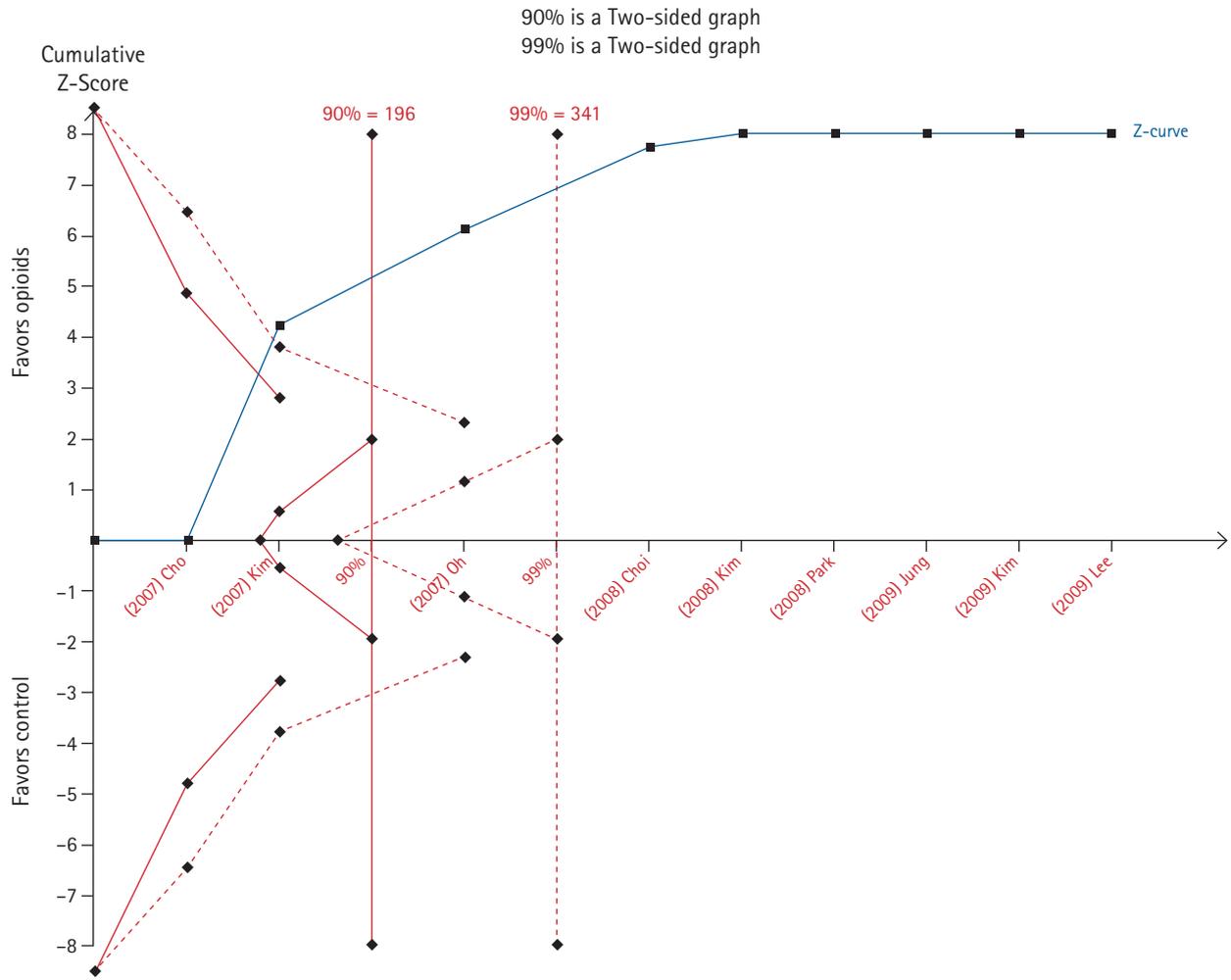
Fig. 4. Trial sequential analysis (TSA) of the effect of opioids in reducing rocuronium-induced withdrawal movement [11].

No confirmation of the meta-analysis pooled effect

In the two TSAs (Figs. 6 and 9), the cumulative z-line lies in the zone with no statistical significance. This implies that the sample size of the meta-analysis was too small, and it is therefore impossible to infer where the cumulative z-line will lie in future trials. If a TSA had been performed by the authors, more cautious conclusions could have been drawn.

Inner wedge

No studies have reported examples of the inner wedge zone. However, for completeness, we would like to briefly illustrate this eventuality. The inner wedge zone is delimited by the futility boundaries, creating an isosceles triangle with its base on the sample size line. If the cumulative z-score lies in the inner wedge zone, future studies on the argument must be considered futile because they will hardly be able to change the no-effect results.

Pre-registering TSA

The importance of registering the TSA protocol before conducting the analysis is depicted in Fig. 7). This TSA resulted in statistical significance using a power of 90%, but the statistical significance was lost using an analysis with a power of 99%. Despite no guidelines or clear recommendations regarding the choice of the power of the analysis, this example shows the limitation of a post-hoc analysis in which the power could be arbitrarily changed to confirm or not the recommended result.

Limitations

Our study has some limitations that we would like to discuss. A limited number of TSAs were included in the analysis, and no examples of a TSA lying in the inner wedge were available.

Other methods such as the law of iterated logarithm penalizing the z-value by the strength of the available evidence and number
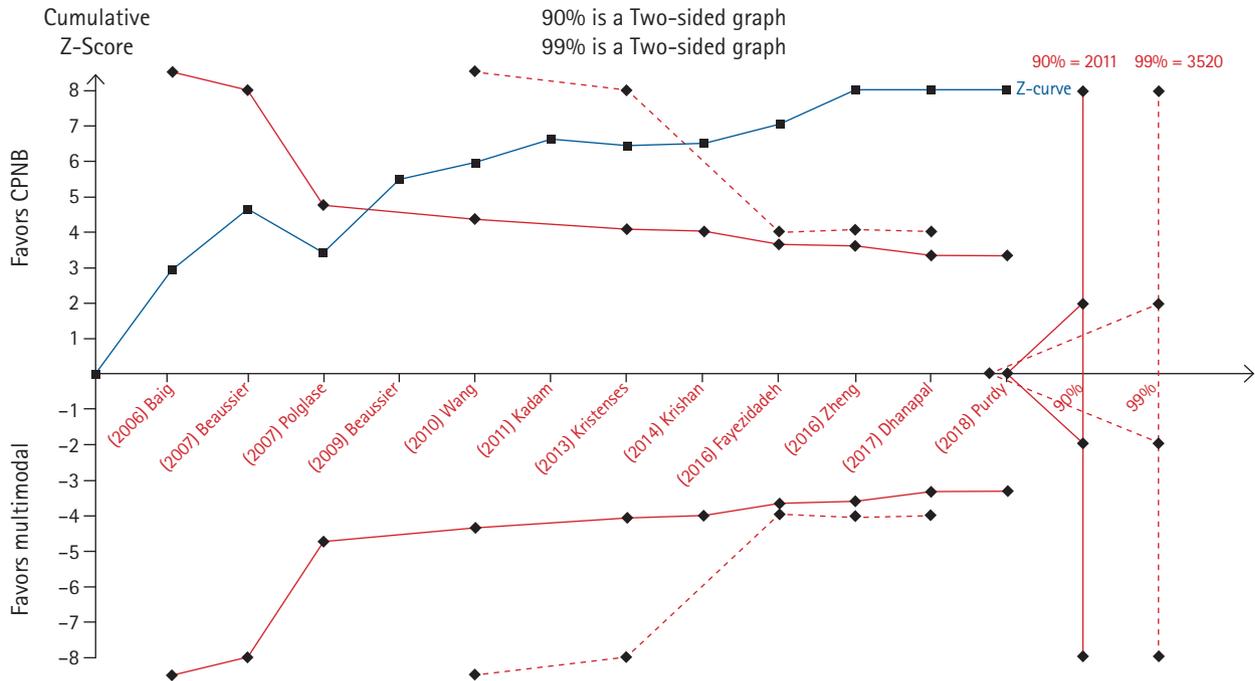
Fig. 5. Trial sequential analysis (TSA) of the effect of multimodal anesthesia compared to that of continuous peripheral nerve blocks on pain at 48 hours following midline laparotomy [12]. CPNB: continuous peripheral nerve block.
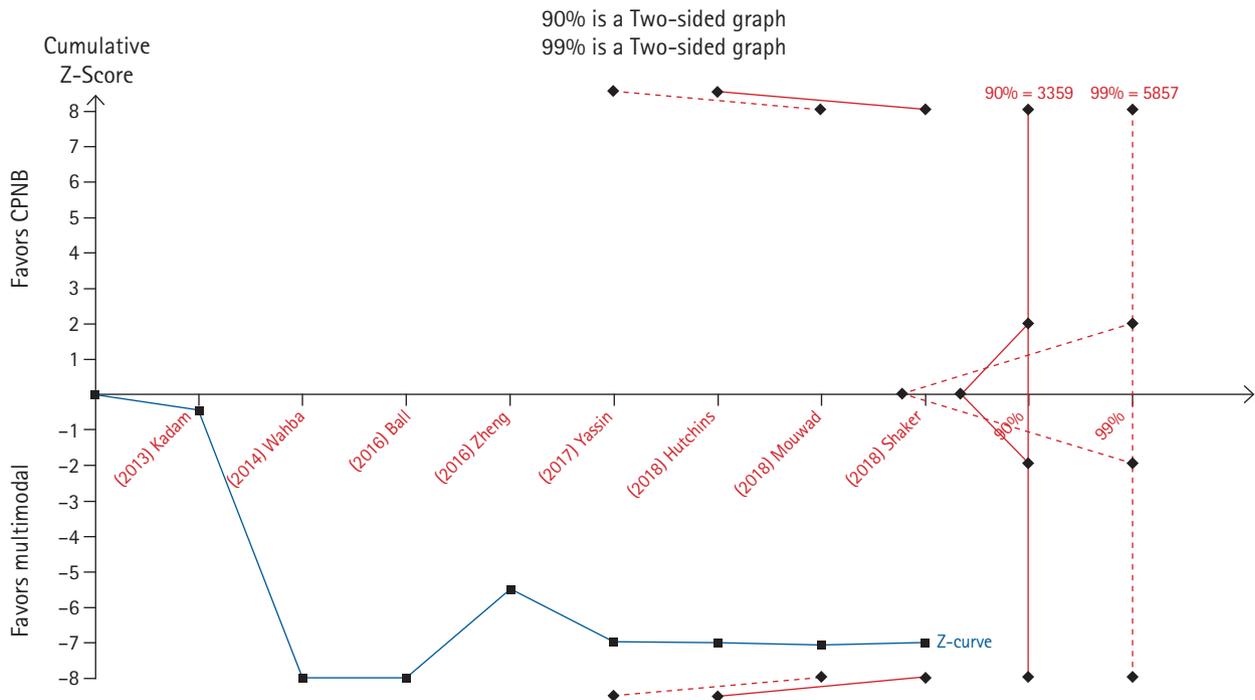
Fig. 6. Trial sequential analysis (TSA) of the effect of epidural anesthesia compared to that of continuous peripheral nerve blocks on pain at 48 hours following midline laparotomy [12]. CPNB: continuous peripheral nerve block.

of statistical tests could be used to adjust the issues of repeated significance testing. In our study, we chose the cumulative z-curve approach, but we recognize this was an arbitrary choice.

We also presented a guide to help clinicians interpret TSA; however, we recognize that we have not explained the statistical basis of this analysis and we recognize this as a limitation.
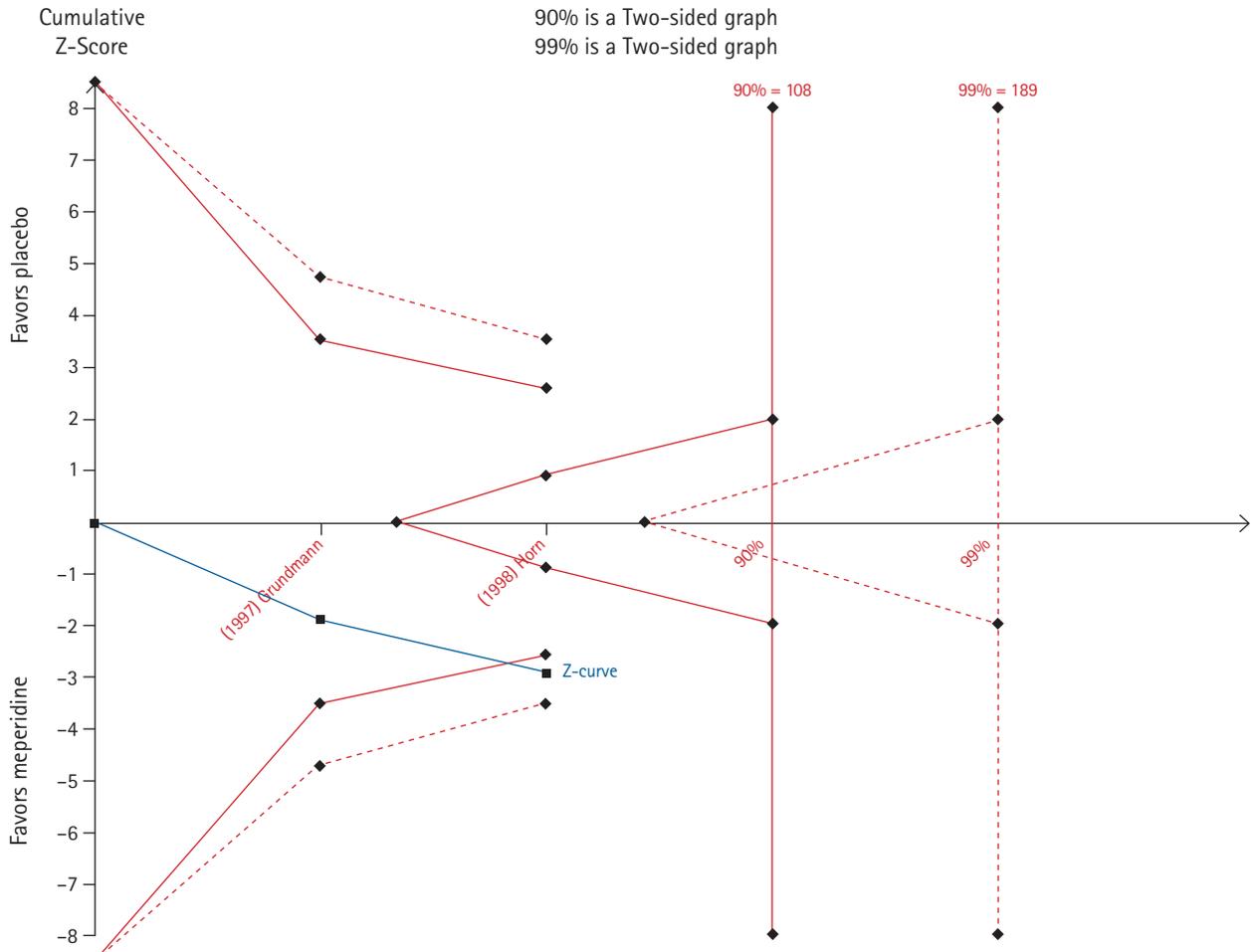
Fig. 7. Trial sequential analysis (TSA) of the effect of meperidine compared to that of placebo on postoperative shivering [13].
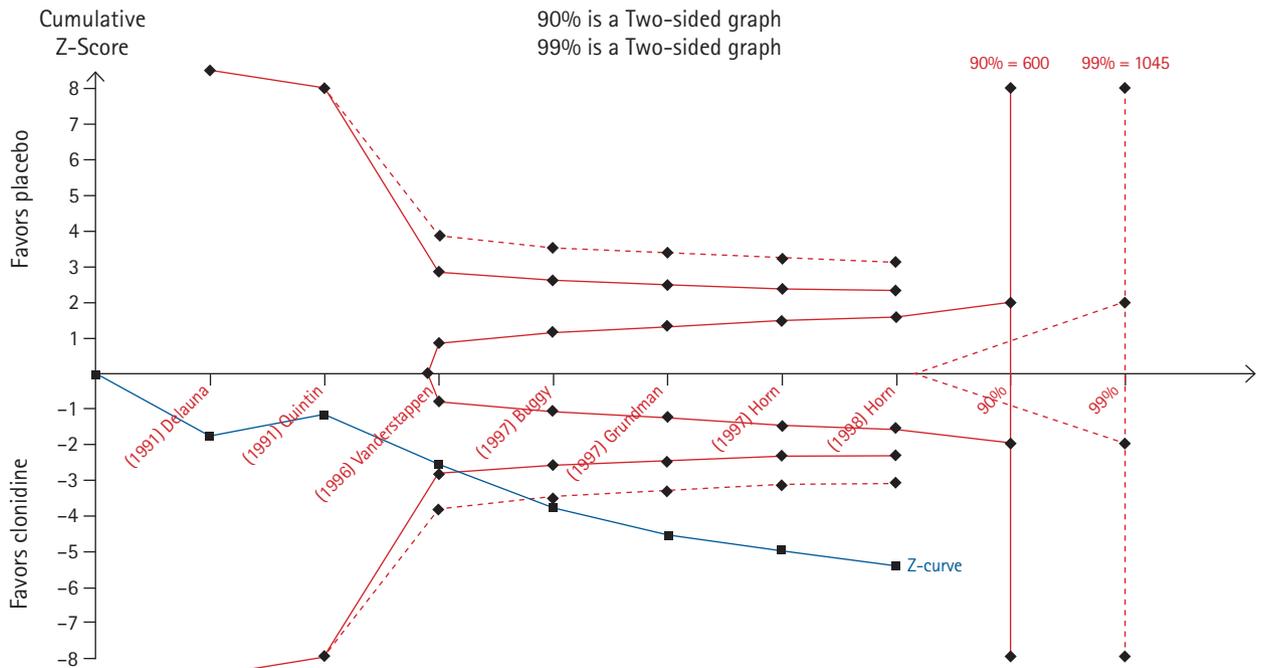
Fig. 8. Trial sequential analysis (TSA) of the effect of clonidine compared to that of placebo on postoperative shivering [13].
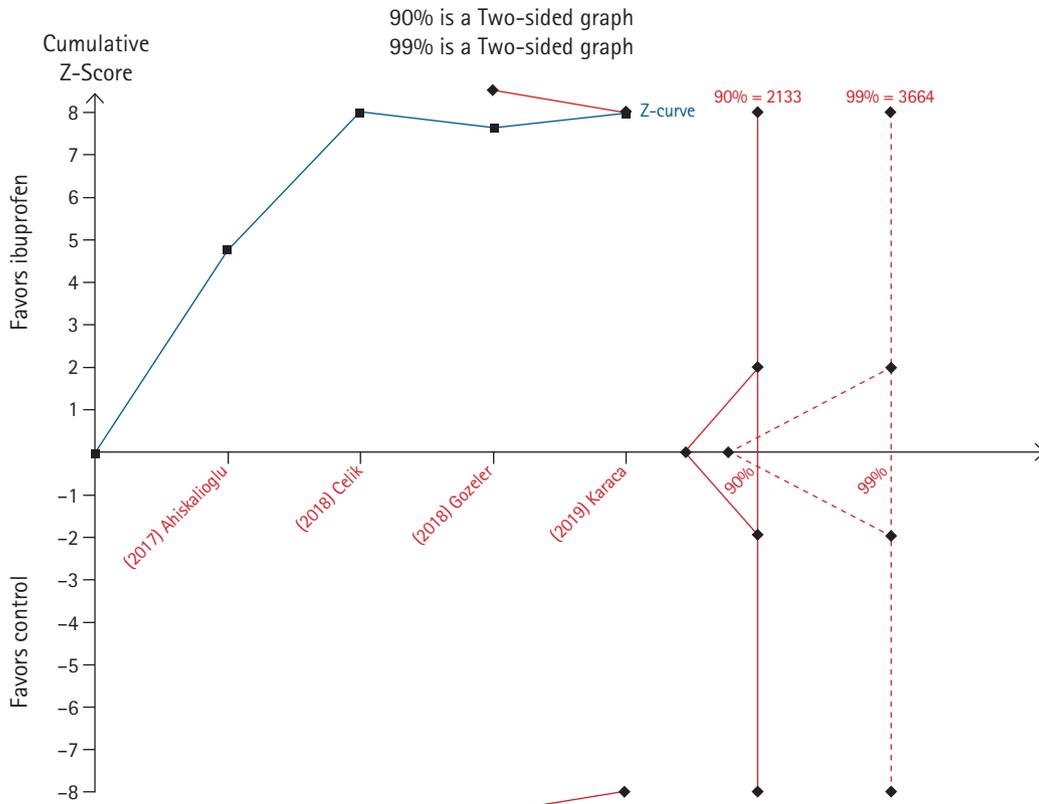
Fig. 9. Trial sequential analysis (TSA) of the effect of ibuprofen on postoperative opioid consumption [14].
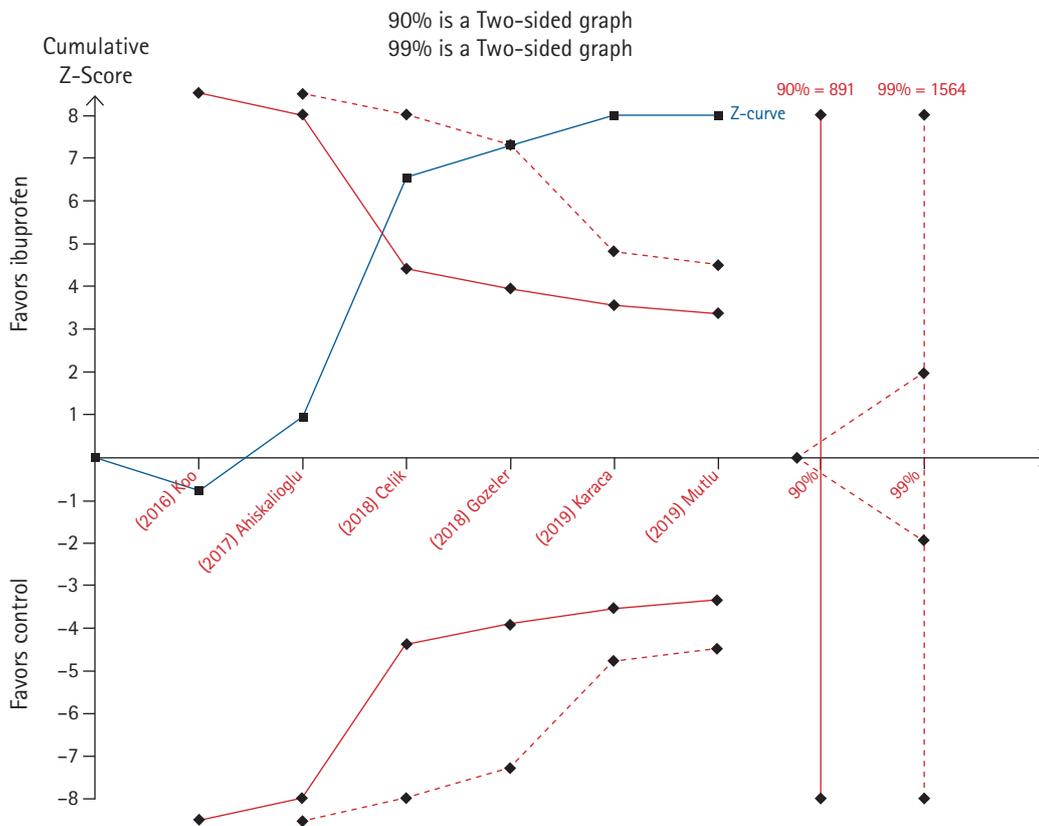
Fig. 10. Trial sequential analysis (TSA) of the effect of ibuprofen on postoperative pain [14].
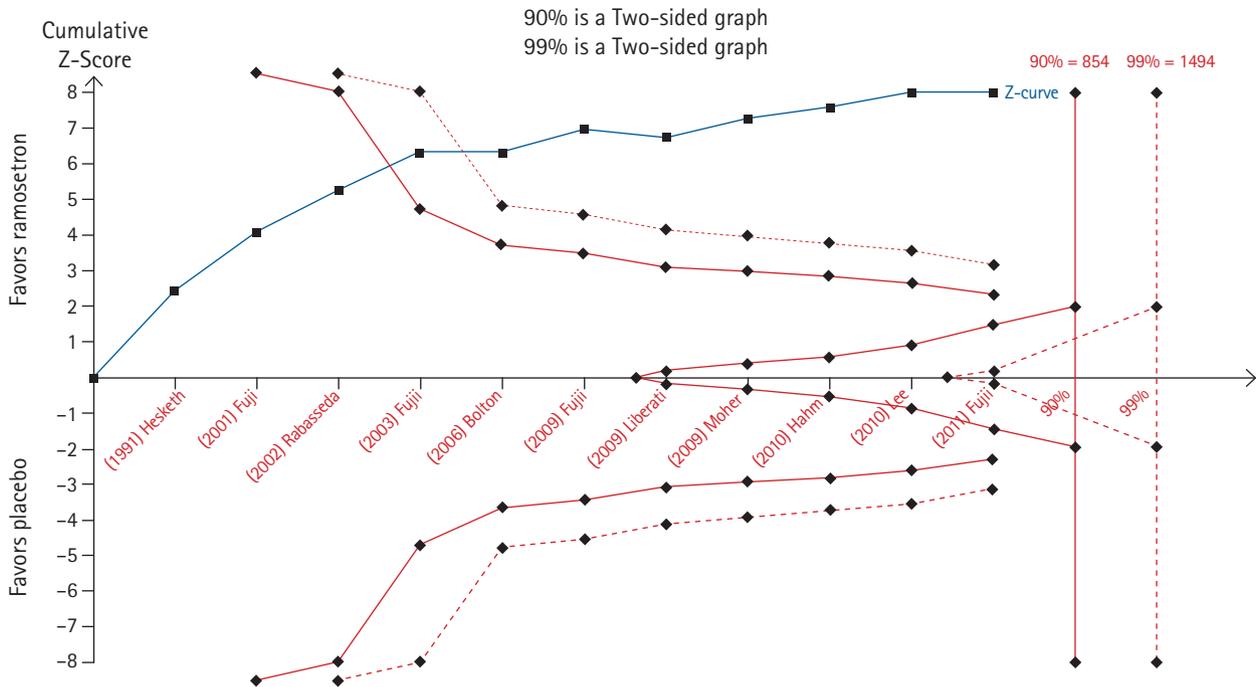
Fig. 11. Trial sequential analysis (TSA) of the efficacy of ramosetron in preventing postoperative nausea and vomiting [15].
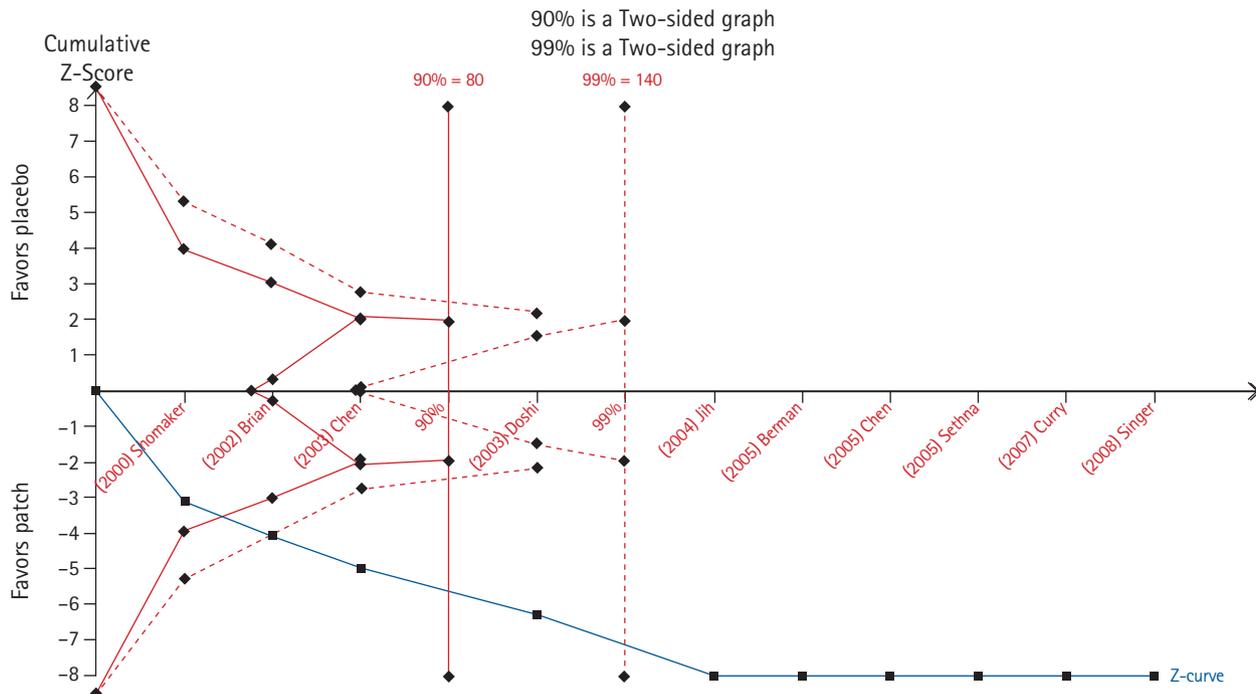
Fig. 12. Trial sequential analysis (TSA) of the efficacy of lidocaine/tetracaine patch and peel on pain [16].

We showed several examples of how a TSA can be applied to meta-analyses published in the KJA. We believe that this study provides useful insights to better understand the use of this statistical tool.

Table 1. Characteristics of the Included Studies

Author (yr)	Registration number	Main outcome	n	Intervention	Control	Overall effect (95% CI)
Choi et al. (2014) [11]	-	Incidence of rocuronium-induced withdrawal movement following pretreatment with lidocaine	905	223/480	316/425	Random effects using the M-H method:
						RR 0.60 (0.49, 0.74)
		Incidence of rocuronium-induced withdrawal movement following pretreatment with opioids	1016	146/582	353/434	Random effects using the M-H method:
						RR 0.28 (0.18, 0.44)
Bailey et al. (2020) [12]	CRD42017051770	Cumulative opioid consumption at 48 hours in patients undergoing midline laparotomy with continuous peripheral nerve blocks versus multimodal analgesia	1080	552	528	Random effects using the MD IV:
						−31.52 (−42.81, −20.22)
		Cumulative opioid consumption at 48 hours in patients undergoing midline laparotomy with continuous peripheral nerve blocks versus epidural analgesia	566	293	273	Random effects using the MD IV:
						16.13 (−0.10, 32.36)
Min et al. (1999) [13]	-	Meperidine for prevention of postoperative shivering	70	5/35	17/35	Fixed effects using Peto OR:
						0.2 (0.1, 0.5)
		Clonidine for prevention of postoperative shivering	518	99/259	161/259	Fixed effects using Peto OR:
						0.3 (0.2, 0.5)
Kim et al. (2021) [14]	CRD42020166141	Opioid consumption following treatment with ibuprofen	269	135	134	Random effects using MD IV:
						−170.70 (−265.64, −75.77)
		Postoperative pain scores following treatment with ibuprofen	266	185	181	Random effects using MD IV:
						−0.58 (−0.99, −0.18)
Kim et al. (2011) [15]	-	Incidence of postoperative nausea and vomiting following pretreatment with ramosetron	685	106/340	216/345	Random effects using RR IV:
						0.40 (0.27, 0.58)
Kim et al. (2012) [16]	-	Efficacy and safety of lidocaine/tetracaine patch and peel to treat pain	574	211/298	70/276	Fixed effects using RR IV:
						2.49 (2.01, 3.07)

n: number, M-H: Mantel–Haenszel, RR: relative risk, MD: mean difference, IV: inverse variance.

Acknowledgements

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Author Contributions

Alessandro De Cassai (Conceptualization; Formal analysis; Methodology; Project administration; Supervision; Writing – original draft; Writing – review & editing)

Martina Tassone (Conceptualization; Writing – original draft; Writing – review & editing)

Federico Geraldini (Writing – original draft; Writing – review &

editing)

Massimo Sergi (Writing – original draft; Writing – review & editing)

Nicolò Sella (Writing – original draft; Writing – review & editing)

Annalisa Boscolo (Writing – original draft; Writing – review & editing)

Marina Munari (Writing – original draft; Writing – review & editing)

ORCID

Alessandro De Cassai, https://orcid.org/0000-0002-9773-1832
Martina Tassone, https://orcid.org/0000-0002-2796-0488
Federico Geraldini, https://orcid.org/0000-0002-1837-0662
Massimo Sergi, https://orcid.org/0000-0002-6745-3309
Nicolò Sella, https://orcid.org/0000-0001-9338-4995
Annalisa Boscolo, https://orcid.org/0000-0003-3409-4159
Marina Munari, https://orcid.org/0000-0002-9843-7213

References

1. Wetterslev J, Jakobsen JC, Gluud C. Trial Sequential Analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 2017; 17: 39.

2. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol 2008; 61: 64-75.

3. De Cassai A, Boscolo A, Zarantonello F, Piasentini E, Di Gregorio G, Munari M, et al. Serratus anterior plane block for video-assisted thoracoscopic surgery: a meta-analysis of randomised controlled trials. Eur J Anaesthesiol 2021; 38: 106-14.

4. De Cassai A, Boscolo A, Geraldini F, Zarantonello F, Pettenuzzo T, Pasin L, et al. Effect of dexmedetomidine on hemodynamic responses to tracheal intubation: a meta-analysis with meta-regression and trial sequential analysis. J Clin Anesth 2021; 72: 110287.

5. De Cassai A, Pasin L, Boscolo A, Salvagno M, Navalesi P. Trial sequential analysis: plain and simple. Korean J Anesthesiol 2021; 74: 363-5.

6. Ahn E, Kang H. Introduction to systematic review and meta-analysis. Korean J Anesthesiol 2018; 71: 103-12.

7. Kang H. How to understand and conduct evidence-based medicine. Korean J Anesthesiol 2016; 69: 435-45.

8. Oh JH, Shin WJ, Park S, Chung JS. Reporting and methodologic evaluation of meta-analyses published in the anesthesia literature according to AMSTAR and PRISMA checklists: a preliminary study. Korean J Anesthesiol 2017; 70: 446-55.

9. Tantry TP, Karanth H, Shetty PK, Kadam D. Self-learning software tools for data analysis in meta-analysis. Korean J Anesthesiol 2021; 74: 459-61.

10. Kim WO, Kil HK, Shin YS, Ahn AK. Prevention of hemodynamic changes after tracheal intubation. Korean J Anesthesiol 1991; 24: 754-9.

11. Choi GJ, Lee S, Lee JH, Park SG, Kang H. Pharmacological and non-pharmacological intervention for rocuronium-induced withdrawal movement in the Korean population: a meta-analysis of 41 studies including 4,742 subjects. Korean J Anesthesiol 2014; 66: 419-32.

12. Bailey JG, Morgan C, Christie R, Ke J, Kwofie K, Uppal V. Continuous peripheral nerve blocks (CPNBs) compared to thoracic epidurals or multimodal analgesia for midline laparotomy: a systematic review and meta-analysis. Korean J Anesthesiol 2021; 74: 394-408.

13. Min SK, Kim WO, Nam YK, Han SG, Lee SJ, Lee YS. Pharmacological prevention of post-anesthetic shivering: clonidine vs meperidine: a meta-analysis of randomized controlled trials. Korean J Anesthesiol 1999; 37: 63-72.

14. Kim SY, Lee S, Lee Y, Kim H, Kim KM. Effect of single dose preoperative intravenous ibuprofen on postoperative pain and opioid consumption: a systematic review and meta-analysis. Korean J Anesthesiol 2021; 74: 409-21.

15. Kim WO, Koo BN, Kim YK, Kil HK. Ramosetron for the prevention of postoperative nausea and vomiting (PONV): a meta-analysis. Korean J Anesthesiol 2011; 61: 405-12.

16. Kim WO, Song BM, Kil HK. Efficacy and safety of a lidocaine/tetracaine medicated patch or peel for dermatologic procedures: a meta-analysis. Korean J Anesthesiol 2012; 62: 435-40.