Journal of Clinical Epidemiology 169 (2024) 111276

Journal of Clinical Epidemiology

OTHER GRADE PAPERS

The GRADE Working Group and CINeMA approaches provided inconsistent certainty of evidence ratings for a network meta-analysis of opioids for chronic noncancer pain

Atefeh Noori[a,b,c], Behnam Sadeghirad[a], Lehana Thabane[a], Mohit Bhandari[a,d], Gordon H. Guyatt[a], Jason W. Busse[a,e,f,*]

[a]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[b]*The Michael G. DeGroote National Pain Center, McMaster University, Hamilton, Ontario, Canada*
[c]*Hand Program, Division of Plastic, Reconstructive and Aesthetic Surgery, University Health Network, Toronto Western Hospital, Affiliated with University of Toronto*
[d]*Division of Orthopaedic Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada*
[e]*Department of Anesthesia, McMaster University, Hamilton, Ontario, Canada*
[f]*The Michael G. DeGroote National Pain Centre, McMaster University, Hamilton, Ontario, Canada*

Accepted 4 February 2024; Published online 8 February 2024

Abstract

Objectives: Assessment of the certainty of evidence (CoE) from network meta-analysis is critical to convey the strength of inferences for clinical decision-making. Both the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group (GWG) and the Confidence in Network Meta-Analysis (CINeMA) framework have been designed to assess the CoE of treatment effects informed by network meta-analysis; however, the concordance of results is uncertain.

Study Design and Setting: We assessed the CoE for treatment effects of individual opioids on pain relief and physical functioning from a network meta-analysis for chronic noncancer pain using the GWG approach and the CINeMA framework. Both approaches evaluate the CoE as high, moderate, low or very low. We quantified the number of discrepant CoE ratings between approaches and the magnitude of the difference (ie, one level, two levels, or three levels).

Results: Across 105 comparisons among individual opioids for pain relief, the GWG and CINeMA approaches provided different CoE ratings in 34% of cases (36 of 105). Across 66 comparisons for physical functioning, there was discordance in 17% of cases (11 of 66). All discrepancies were separated by one level. The CINeMA framework typically provided lower CoE ratings compared to the GWG approach, predominantly because of differences in the assessment of transitivity and heterogeneity.

Conclusion: Our findings suggest there are differences between the CoE ratings provided by the GWG and CINeMA approaches when applied to network meta-analyses. Further research is needed to replicate or refute our findings in other network meta-analyses and assess the implications for clinical decision-making.

Keywords: Certainty of evidence; GRADE; CINeMA; Network meta-analysis

Plain language summery

Network Meta-Analysis allows researchers to compare the effectiveness of multiple treatments against each other. The confidence in treatment effects (how likely they are to be true) is established by evaluating the CoE, which can be high, moderate, low or very low. We compared two systems for evaluating the CoE for a network meta-analysis of different opioids for chronic pain: the GWG and CINeMA. We found discrepancies between these approaches were common, with CINeMA typically providing lower assessments of the CoE than GRADE.

Funding source: This was an unfunded study. JWB is supported, in part, by a Canadian Institutes of Health Research Canada Research Chair in the prevention and management of chronic pain.

* Corresponding author. Department of Anesthesia, McMaster University, HSC−2V9, 1280 Main St. West, Hamilton, Canada, L8S 4K1. Tel.: 905-525-9140 (x21731); fax: 905-523-1224.

E-mail address: bussejw@mcmaster.ca (J.W. Busse).

https://doi.org/10.1016/j.jclinepi.2024.111276

What is new?

Key findings

- Discrepancies in CoE ratings were common between the GWG and CINeMA approaches when applied to the same network meta-analysis of opioids for chronic pain.

What this adds to what was known?

- In cases of discrepancy, the CINeMA approach generally yields a lower CoE compared to the GWG framework.
- Most discrepancies between CINeMA and the GWG approaches are the result of differences in the assessment of transitivity and heterogeneity.

What is the implication and what should change now?

- Systematic review authors conducting network meta-analyses should be aware of differences in CINeMA and GWG approaches when evaluating CoE, and justify their selected approach.

1. Introduction

Network Meta-Analysis (NMA), an extension of conventional meta-analysis, evaluates the relative effectiveness of multiple treatments by combining direct and indirect evidence [1,2]. With several available competing interventions and different outcomes to consider, the results of NMAs are often complex and challenging to interpret [3]. The surface under the cumulative ranking curve approach calculates the probability that a treatment is among the best options (with larger values representing higher ranking probabilities); however, this approach does not consider the certainty of evidence (CoE) and has the potential to mislead [4]. Recent guidance from the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group recommends a more comprehensive approach to rating competing interventions that considers the CoE of point estimates (including assessment of clinical relevance) [5].

The Cochrane Collaboration has endorsed the GRADE approach to evaluate CoE for meta-analyses [6]. Two approaches that apply GRADE principles for evaluating CoE of NMA are the GRADE Working Group (GWG) and the Confidence in Network Meta-Analysis (CINeMA) framework [7]. Both approaches incorporate risk of bias, publication bias, indirectness, imprecision, heterogeneity and incoherence; however, there are several conceptual differences between them. Moreover, CINeMA utilizes a contribution matrix that enables the partial automation of CoE evaluation, which has been incorporated into a freely accessible web application. Based on authors' judgment, the output can be partially or fully overridden in this software. In the current study, we investigated the concordance of CoE ratings between these systems in an NMA of opioids for chronic noncancer pain.

2. Methods

In this case study, we used the GWG approach [8] and CINeMA framework to evaluate the CoE for all network effect estimates for pain and physical functioning in patients with chronic noncancer pain receiving opioids vs. placebo or another opioid. Full methodologic details of our NMA and analysis, which used the GWG approach to assess the CoE, are described elsewhere [4,9].

2.1. Application of the GRADE principles on NMA output

Two reviewers applied the GWG approach for both pain relief and physical functioning independently and in duplicate and resolved any discrepancies by discussion. One reviewer also assessed CoE in CINeMA framework for each outcome, and the senior author confirmed the final assessment. The principles of both approaches are summarized in Table 1. Details regarding the application of each approach with an example are provided in Appendix A.

2.2. GWG approach

We assessed imprecision of the network estimate in relation to the anchor-based minimally important difference (MID) for outcome measures considered. We rated the network estimate as imprecise if the associated 95% confidence interval (95% CI) included half of the MID: 0.5 cm for pain on a 0−10 cm visual analog scale [10], and 2.5 points for physical function on the 0−100 points short form-36 physical component summary score [11]. We evaluated heterogeneity of all pooled estimates from direct comparisons using the I^2 statistic as per guidance from Cochrane [12]. We also visually inspected forest plots for consistency, given that there is no strict cut-off for heterogeneity interpretation, and adjudication depends on the direction and magnitude of effects and the strength of evidence for heterogeneity. The GWG does not recommend rating down twice in the presence of inconsistency and imprecision, as these issues are related. We either rated down the certainty of network by one level if there was evidence of incoherence between the direct and indirect evidence, or used only the direct or indirect evidence if one provided a higher CoE [13].

2.3. CINeMA framework approach

When the threshold for imprecision is entered as half of the MID, CINeMA uses two strategies to rate imprecision

Table 1. Comparison in approaches between GRADE Working Group (GWG) and CINeMA approaches to derive the overall certainty of evidence

	Direct evidence		Indirect evidence		Network evidence	
Domain assessment	**GWG**	**CINeMA framework**	**GWG**	**CINeMA framework**	**GWG**	**CINeMA framework**
Study limitations	Yes	Yes	No	Yes	-	-
Indirectness	Yes	Yes	No	Yes	-	-
Heterogeneity	Yes	Yes	No	Yes	-	-
Publication bias	Yes	Yes	No	Yes	-	-
Intransitivity[a]	No	No	Yes	No	-	-
Imprecision	-	-	-	-	Yes	Yes
Incoherence	-	-	-	-	Yes	Yes
Automatic compute	-	-	-	-	No	Yes
Overall rating across domains	Yes	No	Yes	No	Yes	Yes

Abbreviations: CINeMA, Confidence in Network Meta-Analysis; GRADE, Grading of Recommendations Assessment, Development and Evaluation; GWG, GRADE Working Group.

[a] Intransitivity is assessed as a part of the consideration of indirectness with the CINeMA framework.

of the network estimate: (1) if the point estimate is less than the threshold (in our case, half of the MID) then the estimate is considered imprecise if the associated 95% CI includes half of the MID or (2) if the point estimate is greater than the threshold (eg, half of the MID) then the estimate is considered imprecise if the associated 95% CI includes the null effect (Supplement Figure 1). For assessment of heterogeneity for each network estimate, CINeMA considers the agreement between 95% CI and the prediction interval, which is a range of values between which the true effect of a new study is likely to lie. For evidence with both direct and indirect estimates available, the incoherence is judged based on the agreement of direct and indirect 95% CIs and the *P* value from a side-split test [7].

2.4. Transitivity assumption

The transitivity assumption requires similarity across the contributing direct comparisons in terms of the population, intervention and control, and outcome measure. In the GWG approach for assessing the CoE of effect estimates from an NMA, indirect effect estimates may be rated down for intransitivity. We found a credible subgroup effect in a previous meta-analysis of opioids based on the duration of follow-up (<3 months vs. ≥ 3 months) [14]. Thus, when there was a large imbalance (50% or more difference) in the length of follow-up between the two direct comparisons, we rated down the CoE of indirect evidence one level for intransitivity.

CINeMA also suggests evaluating each included study according to its relevance to the research question in terms of effect modifying variables. The rationale for this recommendation is that the effect modifier distribution can only be evaluated in networks with enough studies available for a specific comparison. Based on CINeMA guidance [15], we downgraded the CoE for opioids that were assessed in a single study or only compared vs. placebo without direct comparisons with other opioids.

2.5. Sensitivity analysis for assessment of imprecision

We conducted a sensitivity analysis incorporating the imprecision assessment approach used by CINeMA into the GWG approach to evaluate the robustness of our results (ie, we reported no concerns with imprecision when the point estimate exceeded the threshold of half of the MID, and the CI included half of the MID but not the null effect) (Supplement Figure 1).

3. Results

3.1. Discrepancies for pain relief

Across 105 comparisons among individual opioids for pain relief, the GWG and CINeMA approaches provided different CoE ratings in 34% of cases (36 of 105, Table 2, Supplement Table 1). Each discrepancy was separated by one level; 35 were very low vs. low and one was moderate vs. low.

3.2. Types of discrepancies

Among these 36 discrepancies, 13 (36%) were related to intransitivity, 9 (25%) to heterogeneity, 7 (20%) to assessment of imprecision, and 6 (17%) to either intransitivity or heterogeneity. One discrepancy (2%) resulted because the GWG approach used the direct estimate with moderate certainty as the best evidence over the network estimate due to an inflated 95% CI [16]. For most discrepancies, the GWG resulted in higher certainty ratings compared with the CINeMA framework. Specifically, for 25 out of 36 discrepancies, GWG rated the CoE as low, but CINeMA rated very low; in one of them, the GWG rated the CoE as moderate and CINeMA rated low. For 10 comparisons, GWG rated the CoE as very low and CINeMA rated low (seven due to imprecision and three to intransitivity).

Table 2. Direct, indirect, and network estimates based on the GWG and CINeMA approaches for pain relief on a 10-cm visual analog scale among patients with chronic noncancer pain

GWG system						
Opioid vs. placebo	**Direct Estimate MD (95% CI)**	**#of Studies**	**I^2**	**#of patients**	**Direct CoE**	**Indirect Estimate MD (95% CI)**
BUP-sublingual	−0.87 (−1.11 to −0.63)	2	59	930	M	−0.92 (−2.09 to 0.24)
BUP-transdermal	−0.61 (−0.78 to −0.45)	6	0	1,471	L	−0.8 (−1.35 to −0.25)
COD-ER	−2.03 (−3.09 to −0.97)	1	NA	66	M	NA
FEN-transdermal	−0.73 (−1.06 to −0.39)	3	0	712	M	−0.83 (−1.47 to −0.19)
HMOR-ER	−0.41 (−1.1 to 0.27)	3	90	1,521	L	−0.64 (−1.29 to 0)
HYD-ER	−0.53 (−0.74 to −0.32)	3	0	1,260	M	NA
MPH-ER	−0.93 (−1.23 to −0.62)	9	0	880	M	−0.75 (−1.25 to −0.25)
OMOR-ER	−1.51 (−2.3 to −0.72)	3	73	619	L	NA
OXY-ER	−0.76 (−1.18 to −0.35)	13	85	3,579	L	−0.6 (−1.03 to −0.16)
TPN-ER	−0.73 (−1.02 to −0.43)	9	62	3,085	M	−1.2 (−1.9 to −0.49)
TPN-IR	NA	NA	NA	NA	NA	−1.09 (−2.22 to 0.04)
TRA-ER	−0.74 (−0.94 to −0.54)	11	37	4,202	M	−0.93 (−1.56 to −0.3)
TRA-IR	−1.13 (−1.76 to −0.5)	4	66	545	M	−0.97 (−2.03 to 0.1)
OXY-IR	NA	NA	NA	NA	NA	−0.99 (−1.81 to −0.17)

Results are mean difference (95% CIs) from DerSimonian and Laird random-effects meta-analysis.
Direct estimates were rated down in the presence of risk of bias (ROB), indirectness, publication bias, or heterogeneity.
Indirect estimates were rated down in the presence of intransitivity.
Network estimates rated down in the presence of incoherence or imprecision (either due to inclusion of the half MID in the 95% CI, or because the evidence was supported by a small number of participants that failed to meet the optimal information size ($\leq$300).
Abbreviations: H, high certainty of evidence; M, moderate; L, low; VL , very low. MPH, morphine; FEN, fentanyl; BUP, buprenorphine; OXY, oxycodone; TPN, tapentadol; TRA, tramadol; HMOR, hydromorphone; OMOR, oxymorphone; COD, codeine; HYD, hydrocodone. ER, extended release; IR, immediate release; CINeMA, Confidence in Network Meta-Analysis; GRADE, Grading of Recommendations Assessment, Development and Evaluation; GWG: GRADE Working Group; CoE, certainty of evidence; ROB, risk of bias; NMA, Network Meta-Analysis.
[a] Used the direct estimate as the best evidence over the network estimate because of the inflated 95% CI of network estimation.
[b] Downgraded due to intransitivity concern as this opioid was poorly connected to the network or assessed in a single study.

3.3. Discrepancies for physical functioning

Across 66 comparisons between individual opioids for physical functioning, the GWG and CINeMA approaches provided different CoE ratings in 17% of cases (11 of 66; Supplement Table 2); all discrepancies were separated by one level (eg, very low vs. low).

3.4. Type of the discrepancies

Six out of 11 discrepancies were related to assessment of heterogeneity (54%), four due to transitivity considerations (36%), and one was due to either heterogeneity or transitivity (10%). For all 11 discrepancies, GWG rated the CoE as low and CINeMA as very low.

3.5. Intransitivity

Using the CINeMA framework, all comparisons (a total of 58) involving extended release (ER) codeine, ER hydrocodone, ER oxymorphone, immediate release (IR) oxycodone, and IR tapentadol were rated down due to intransitivity for pain relief, as they were informed by a single study or were not connected to other opioids in the network. In the GWG approach, 36% (21 out of 58) of them were rated down for violating the transitivity assumption.

For physical function, all comparisons (a total of 41) involving ER codeine, ER hydrocodone, sublingual buprenorphine, and ER tramadol were rated down due to intransitivity in the CINeMA framework. Using the GWG approach, 32 of them (78%) were rated down due to intransitivity.

3.6. Heterogeneity

The GWG evaluates heterogeneity for direct evidence based on the I^2 statistic; however, CINeMA considers heterogeneity for each network estimate according to the agreement between the 95% CI and the prediction interval. For pain relief, CINeMA rated down 55 out of 105 comparisons (52%) for heterogeneity, however, GWG only evaluated 15 out of the 105 comparisons for heterogeneity (only comparisons for which at least two studies included the comparison). Among the 15 comparisons assessed by both the GWG and CINeMA approaches, (1) seven were assigned concerns based on heterogeneity by CINeMA but not the GWG, (2) the GWG assigned concerns to two that CINeMA did not, and (3) for six comparisons both the GWG and CINeMA were in agreement—three were assigned concerns and three were not.

GWG system			CINeMA framework		
Indirect CoE	NMA estimate MD (95% CI)	NMA CoE	NMA estimate	NMA CoE	Reasons
L	−0.86 (−1.35 to −0.38)	M[a]	−0.86 (−1.37 to −0.35)	L	ROB Heterogeneity
M	−0.71 (−1 to −0.41)	L	−0.71 (−1.02 to −0.40)	L	ROB Heterogeneity
NA	−2.03 (−3.28 to −0.78)	L	−2.03 (−3.31 to −0.74)	L	ROB Intransitivity[b]
L	−0.78 (−1.18 to −0.39)	L	−0.78 (−1.19 to −0.36)	L	ROB Heterogeneity
L	−0.52 (−0.88 to −0.16)	VL	−0.51 (−0.89 to −0.13)	L	ROB Heterogeneity
NA	−0.53 (−0.97 to −0.09)	L	−0.53 (−0.99 to −0.06)	VL	ROB Heterogeneity Intransitivity[b]
M	−0.86 (−1.17 to −0.56)	M	−0.86 (−1.18 to −0.55)	M	ROB
NA	−1.47 (−2.03 to −0.91)	L	−1.68 (−2.18 to −1.18)	L	ROB Intransitivity[b]
M	−0.66 (−0.89 to −0.44)	L	−0.66 (−0.89 to −0.43)	L	ROB Heterogeneity
L	−0.81 (−1.08 to −0.53)	M	−0.80 (−1.09 to −0.51)	M	ROB
M	−1.09 (−2.22 to 0.04)	L	−1.09 (−2.28 to 0.09)	VL	ROB Imprecision, Intransitivity[b]
L	−0.80 (−1.05 to −0.55)	M	−0.80 (−1.06 to −0.54)	M	ROB
L	−1.09 (−1.54 to −0.65)	M	−1.09 (−1.55 to −0.63)	M	ROB
M	−0.99 (−1.81 to −0.17)	L	−0.99 (−1.85 to −0.13)	VL	ROB Heterogeneity Intransitivity[b]

3.7. Imprecision

Using the GWG approach for pain relief, we rated down the network estimates for transdermal buprenorphine and fentanyl, ER hydromorphone, ER hydrocodone, ER oxycodone, and IR oxycodone vs placebo for imprecision, as the associated 95% CI included half the MID. Further, for the comparison of ER codeine vs placebo, the GWG approach rated the CoE down one level due to imprecision, as the direct evidence was informed by less than 300 observations. These comparisons were not rated down using the CINeMA framework.

3.8. Sensitivity analysis

By using the same approach for adjudicating imprecision for both the GWG and CINeMA approaches, the number of discrepancies between frameworks decreased to 31% of cases for pain relief (33 of 105 comparisons). Specifically, 13 (40%) of discordant ratings were related to intransitivity, 12 (36%) to heterogeneity, 6 (18%) to either heterogeneity or intransitivity, and two (6%) for both heterogeneity and intransitivity. Of the 33 discrepancies, CINeMA assigned a lower rating than the GWG approach for 91% (30 of 33): (1) GWG rated the CoE as low, but CINeMA rated very low for 24, (2) GWG rated the CoE as moderate and CINeMA rated low for four, and (3) GWG rated the CoE as moderate and CINeMA rated very low for two. For three comparisons, CINeMA assigned a higher CoE than GWG: low vs. very low (all due to intransitivity).

In our sensitivity analysis for physical function, there were discordant ratings in the CoE between frameworks for 12 of 66 of comparisons (18%). Seven related to assessment of heterogeneity (58%), four due to transitivity considerations (33%), and one was due to either heterogeneity or transitivity (9%). In all cases, the GWG approach rated the CoE higher than CINeMA (ie, 11 low vs. very low, one moderate vs. low) (Supplement Tables 3 and 4).

4. Discussion

We identified discrepancies in 27% of CoE ratings between the GWG and CINeMA approaches when applied to an NMA of opioids for chronic noncancer pain. Disagreements were separated by one level of magnitude in CoE ratings and were predominantly due to considerations around intransitivity or heterogeneity. When discordant, CINeMA tended to provide lower CoE ratings vs. the GWG approach. For imprecision, the GWG approach rated down the CoE more often than CINeMA.

Regarding the assessment of imprecision, CINeMA assigns no concerns if the 95% CI associated with the network estimate suggests benefit far from null and does not include the null effect. For example, the network estimation of the comparison between transdermal buprenorphine vs. placebo was −0.71 cm (95% CI −1.00 to −0.40 cm on a 10-cm visual analog scale for pain); the network estimate was not rated down in CINeMA because the 95% CI did not include the null effect. Thus, the CINeMA approach is effectively rating certainty in the null effect. Using the GWG approach, we rated our certainty with respect to the threshold of half of the MID whether the true effect was greater or less than the MID. Thus, we rated down one level for imprecision since the 95% CI included half of the MID and values below this were felt to be unimportant to patients [17].

The baseline CoE for the network estimate (before intransitivity, heterogeneity, and imprecision are assessed) was not a source of inconsistency between rating systems, likely because of very limited variability; however, GWG and CINeMA do use different approaches. Even after removing differences in the assessment of imprecision through our sensitivity analysis, there remained substantial discordance as a result of different approaches to the assessment of transitivity and heterogeneity.

The baseline CoE of some network estimates was informed by only indirect evidence and others by both direct and indirect evidence. Generally, in complex networks, there are multiple comparisons informing indirect evidence, including first-order loops and higher. GWG focuses on the most dominant first-order loop which usually contributes the most information to the indirect estimate. In cases in which the network estimate is informed by both direct and indirect evidence, the GWG approach uses the higher CoE of the two to inform the baseline network CoE. The rationale is two-fold: first, the higher rated evidence is typically more precise, second, in the absence of serious incoherence, the evidence (direct or indirect) associated with lower CoE is not likely to reduce the confidence of the network estimate [6]. The CINeMA approach does not choose a source of evidence to inform the baseline CoE estimate for the network estimate. Instead, CINeMA considers the CoE of all evidence based on contribution matrix to inform the baseline network estimate [18].

Ours is the first study to compare the concordance of CoEs between the GWG and CINeMA approaches for evaluating the certainty of network estimates, but there are limitations to our study. First, our findings are based on a single case study and the generalizability to other contexts is uncertain. Second, most of the evidence in our NMA of opioids for chronic pain was rated as low or very-low certainty and this limited variability may have attenuated differences between the GWG and CINeMA approaches. Despite this, discrepancies in CoE ratings were common. Finally, the interpretation and effect of using weights for rating the CoE in different NMA may change according to the network geometry, the amount of direct evidence available, and the degree of differences in risk of bias or indirectness across the comparisons of the network.

5. Conclusion

Our findings suggest that differences between CoE ratings provided by the GWG and CINeMA approaches when applied to network meta-analyses are common. When discrepancies arise, the CINeMA approach typically provides lower CoE ratings than the GWG approach predominantly because of differences in the assessment of transitivity and heterogeneity. Further research, ideally with greater variability in CoE ratings, is needed to confirm our findings in other NMA and assess the implications for clinical decision-making.

CRediT authorship contribution statement

Atefeh Noori: Writing − review & editing, Writing − original draft, Formal analysis, Conceptualization. **Behnam Sadeghirad:** Writing − review & editing, Methodology. **Lehana Thabane:** Writing − review & editing, Methodology. **Mohit Bhandari:** Writing − review & editing, Methodology. **Gordon H. Guyatt:** Writing − review & editing, Supervision, Methodology. **Jason W. Busse:** Writing − review & editing, Supervision, Methodology, Conceptualization.

Data availability

Data will be made available on request.

Declaration of competing interest

BS and GHG are members of the GRADE working group. GHG is a member of the JCE Editorial Board. There are no competing interests for any other author.

Acknowledgments

None.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2024.111276.

References

[1] Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons good research practices: part 2. Value Health 2011;14:429−37.

[2] Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013;159:130−7.

[3] Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev 2017;6:79.

[4] Noori A, Sadeghirad B, Wang L, Siemieniuk RAC, Shokoohi M, Kum E, et al. Comparative benefits and harms of individual opioids for chronic non-cancer pain: a systematic review and network meta-analysis of randomised trials. Br J Anaesth 2022;129:394−406.

[5] Brignardello-Petersen R, Florez ID, Izcovich A, Santesso N, Hazlewood G, Alhazanni W, et al. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework. BMJ 2020;371:m3900.

[6] Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ 2014;349:g5630.

[7] Papakonstantinou T, Nikolakopoulou A, Higgins JP, Egger M, Salanti G. CINeMA: software for semiautomated assessment of the confidence in the results of network meta-analysis. Campbell Syst Rev 2020;16:e1080.

[8] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924−6.

[9] Noori A, Busse JW, Sadeghirad B, Siemieniuk RA, Wang L, Couban R, et al. Individual opioids, and long- versus short-acting opioids, for chronic noncancer pain: protocol for a network meta-analysis of randomized controlled trials. Medicine 2019;98:e17647.

[10] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105−21.

[11] Ward MM, Guthrie LC, Alba MI. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. Arthritis Care Res 2014;66:1783−9.

[12] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.0. Cochrane; 2019. Available at: www.training.cochrane.org/handbook.

[13] Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. J Clin Epidemiol 2018;93:36−44.

[14] Busse JW, Wang L, Kamaleldin M, Craigie S, Riva JJ, Montoya L, et al. Opioids for chronic noncancer pain: a systematic review and meta-analysis. JAMA 2018;320:2448−60.

[15] Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, et al. CINeMA: an approach for assessing confidence in the results of a network meta-analysis. PLoS Med 2020;17:e1003082.

[16] Brignardello-Petersen R, Murad MH, Walter SD, McLeod S, Carrasco-Labra A, Rochwerg B, et al. GRADE approach to rate the certainty from a network meta-analysis: avoiding spurious judgments of imprecision in sparse networks. J Clin Epidemiol 2019;105:60−7.

[17] Goshua A, Craigie S, Guyatt GH, Agarwal A, Li R, Bhullar JS, et al. Patient values and preferences regarding opioids for chronic noncancer pain: a systematic review. Pain Med 2018;19:2469−80.

[18] Gimbel JS, Richards P, Portenoy RK. Controlled-release oxycodone for pain in diabetic neuropathy: a randomized controlled trial. Neurology 2003;60:927−34.