We read with interest the network meta-analysis (NMA) by Tsiachris et al [1] that sought to compare and rank antidysrhythmic drugs for cardioversion of paroxysmal atrial fibrillation (AF). We congratulate the group for such an undertaking but also feel compelled to discuss additional limitations. To improve the power of their analyses, Tsiachris et al [1] analyzed data from studies that enrolled patients with AF duration up to 24 hours, 48 hours, 72 hours, and seven days. Indeed, the NMA generated some effect estimates with narrow confidence intervals but at a cost: potential network intransitivity and confidence in those estimates.

The NMA of data from studies that enrolled patients with such a wide range of AF duration can be problematic. Cardioversion rates may vary with duration of rhythm monitoring, and longer AF duration may be a predictor of unsuccessful cardioversion [2, 3]. Atrial fibrillation duration is likely to be a significant effect modifier if distributed differently across comparisons, and therefore, a likely source of intransitivity in Tsiachris et al's NMA. The authors provide evidence for treatment effect heterogeneity in Table 8 [1]. Among the studies that provided data for conversion to sinus rhythm within four hours, the trial heterogeneity for pooled, individual drug cardioversion rates across the ranges of AF duration varied from $I^2 = 0\%$ to $I^2 = 86\%$. Intransitivity within the analyzed network will result in biased evidence from indirect comparisons [4], thereby threatening the validity of Tsiachris et al's NMA results.

The face validity of the transitivity assumption can be assessed more systematically as part of the consideration of indirectness. Although it may be reasonable to apply the results of their analysis to the patients with paroxysmal AF (AF onset within seven days), Tsiachris et al [1] attempt to apply the results to the emergency department (ED) population with recent-onset AF (AF onset within 48 hours). In contrast to previously published NMA [5], the studied population in Tsiachris et al [1] included patients with AF duration greater than 48 hours and will differ from the population of interest (ED patients with AF duration less than 48 hours) - this is "indirectness". Tsiachris et al [1] apply the GRADE approach [4] using the CINeMA framework [6] to evaluate the confidence (or certainty) in their NMA results for the primary outcome. The GRADE approach [4] to rating the certainty of NMA estimates requires that reviewers downgrade a confidence rating by -1 for some concerns and -2 for major concerns in six domains including indirectness. Tsiachris et al [1] do not follow this direction and appear to overrate the confidence in a number of their NMA estimates without sufficient transparency.

We extracted the data from 25 of Tsiachris et al's 29 studies excluding one non-randomized study [7] and three others that should have been excluded by protocol [1] due to combined AF and atrial flutter patient data. (We cannot confirm from the paper or supplement that Tsiachris et al [1] obtained separated AF patient data from those studies.) We added data from Vogiatzis et al [8] and Kafkas et al [9] that did report separate AF patient data, both of which Tsiachris et al appear to have overlooked for the primary outcome. We independently performed a Cochrane Risk of Bias 2 (RoB 2) [10] quality reassessment and identified discrepancies with Tsiachris et

al's assessments in the overall risk of bias for 12 of 27 (44%) studies; 8 (30%) are overrated in quality. There are also five studies for which Tsiachris et al [1] deviated from the RoB 2 algorithm without providing justification for a higher-quality rating. For the domain of indirectness, we rated data from trials that included patients with AF duration up to 72 hours to have "some concerns" and data from trials that included patients with AF duration up to 7 days to have "major concerns". We redid the frequentist NMA of data from AF patients only (random-effects model using R linked with the CINeMA web application available at http://cinema.ispm.ch); "low-risk" reporting bias, "averaged risk" within-study bias and indirectness; clinically important effect size defined as odds ratio 1.1).

We re-applied the GRADE approach [4] through the CINeMA framework [7] with our own assessments of individual study quality (adhering to the standardized RoB2 algorithm [11]) and indirectness (appropriately considering the target ED population with recent-onset AF). Upon comparison of our NMA estimates with those by Tsiachris et al [1], we found that Tsiachris et al [1] report numerous estimates with overrated confidence. Our re-analysis generated only two estimates rated "moderate" in confidence with the remaining being rated "low" or "very low". "Low confidence" in an effect estimate signifies that "our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect" [4]. "Very low" confidence" signifies that "we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect" [4]. Furthermore, the primarily "low" and "very low" quality NMA estimates will subject the treatment ranking to misinterpretation.

We applied the GRADE approach using a minimally contextualized framework [11] to draw conclusions using a more conservative, standardized confidence assessment. Among the antidysrhythmic drugs used in clinical practice for cardioversion of AF, vernakalant *could be* among the most effective for cardioversion of recent-onset AF in the ED. Flecainide IV, propafenone IV, flecainide PO, ibutilide, propafenone PO, and amiodarone IV *could have* intermediate effectiveness. Procainamide and amiodarone PO *could be* among the least effective. The final classification of the eight interventions is in the Table. Furthermore, the claim by Tsiachris et al [1] that "class Ic antidysrhythmics, when administered intravenously, have higher conversion rates" also cannot be substantiated due to very uncertain evidence.

Our results from the network re-analysis are similar to those by Tsiachris et al [1]. However, when discussing the results in the context of the ED population with recent-onset AF, as in previous NMA [5], the evidence is uncertain and conclusions will remain speculative.

References

1. Tsiachris D, Doundoulakis I, Pagkalidou E, et al. Pharmacologic Cardioversion in Patients with Paroxysmal Atrial Fibrillation: A Network Meta-Analysis. Cardiovasc Drugs Ther 2021;35:293-308.
2. Roy D, Pratt CM, Torp-Pedersen C, et al. Vernakalant hydrochloride for rapid conversion of atrial fibrillation: a phase 3, randomized, placebo-controlled trial. Circulation 2008;117:1518-25.
3. Vardas PE, Kochiadakis GE. Amiodarone for the restoration of sinus rhythm in patients with atrial fibrillation. Card Electrophysiol Rev 2003;7:297-9.
4. Puhan MA, Schunemann HJ, Murad MH, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ 2014;349:g5630.
5. deSouza IS, Tadrous M, Sexton T, Benabbas R, Carmelli G, Sinert R. Pharmacologic Cardioversion of Recent-Onset Atrial Fibrillation and Flutter in the Emergency Department: A Systematic Review and Network Meta-analysis. Ann Emerg Med 2020;76:14-30.
6. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. CINeMA: An approach for assessing confidence in the results of a network meta-analysis. PLoS Med 2020;17:e1003082.
7. Pohjantahti-Maaroos H, Hyppola H, Lekkala M, Sinisalo E, Heikkola A, Hartikainen J. Intravenous vernakalant in comparison with intravenous flecainide in the cardioversion of recent-onset atrial fibrillation. Eur Heart J Acute Cardiovasc Care 2019;8:114-20.
8. Vogiatzis I, Papavasiliou E, Dapcevitch I, Pittas S, Koulouris E. Vernakalant versus ibutilide for immediate conversion of recent-onset atrial fibrillation. Hippokratia 2017;21:67-73.
9. Kafkas NV, Patsilinakos SP, Mertzanos GA, et al. Conversion efficacy of intravenous ibutilide compared with intravenous amiodarone in patients with recent-onset atrial fibrillation and atrial flutter. Int J Cardiol 2007;118:321-5.
10. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.
11. Brignardello-Petersen R, Florez ID, Izcovich A, et al. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework. BMJ 2020;371:m3900.

Certainty of the evidence and Classification[a] of intervention	Intervention[b]	Intervention vs Reference treatment (OR (95% CI))
High certainty (moderate to high certainty evidence)		
Category 2: among the most effective	---	---
Category 1: inferior to the most effective or superior to the least effective	---	---
Category 0: among the least effective	---	---
Low certainty (low to very low certainty evidence)		
Category 2: might be among the most effective	Vernakalant (L)	16.08 (9.39 – 27.53)
Category 1: might be inferior to the more effective or superior to the least effective	Flecainide IV (L)	7.70 (4.81 – 12.33)
	Propafenone IV (L)	6.67 (4.49 – 9.90)
	Ibutilide (L)	5.89 (3.12 – 11.08)
	Flecainide PO (L)	5.84 (3.19 – 10.68)
	Propafenone PO (L)	4.38 (2.63 – 7.30)
	Amiodarone IV (VL)	1.62 (1.04 – 2.50)
Category 0: might be among the least effective	Procainamide (VL)	1.04 (0.21 – 5.09)
	Reference treatment (Control)	---
	Amiodarone PO (VL)	0.07 (0.004 – 1.21)

CI: confidence interval; IV: intravenous; OR: odds ratio; PO: oral; vs: versus

[a]Categories do not inform value judgments about the importance of the effects.

[b]Letters in brackets represent the certainty of evidence for each intervention when compared with the reference treatment (Control): H=high; M=moderate; L=low; VL=very low