

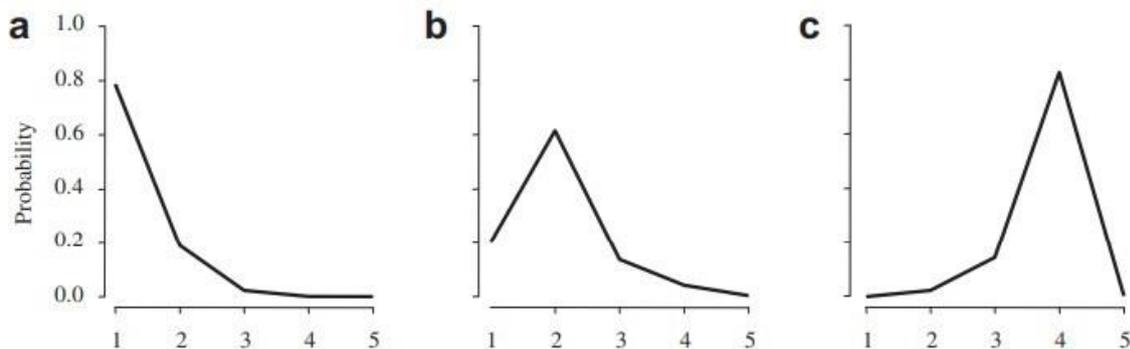


La surface sous la courbe de classement cumulatif (*Surface Under the Cumulative Ranking, SUCRA*) est-elle une manière fiable d'interpréter les résultats d'une méta-analyse en réseau sur le plan clinique ?

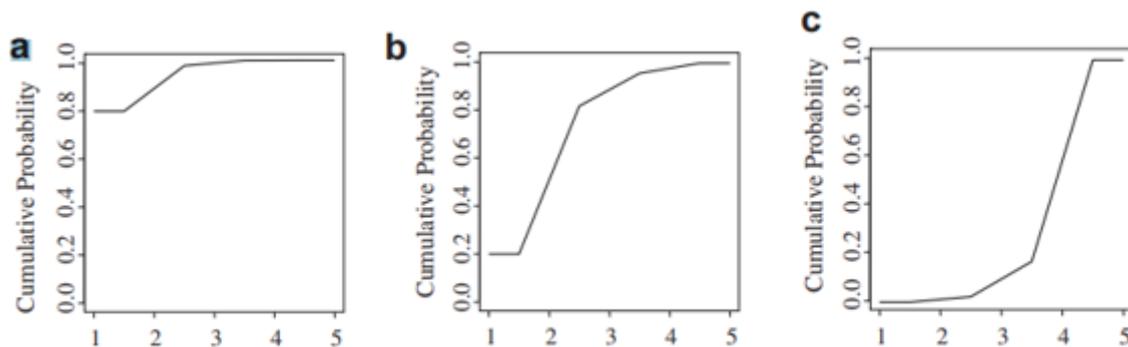
Tom Poelman, Vakgroep Volksgezondheid en Eerstelijnszorg, UGent.

Une méta-analyse « classique » combine les résultats de plusieurs études randomisées contrôlées qui comparent directement une intervention à un placebo ou à une autre intervention. Le but est d'obtenir une estimation plus précise de l'effet. En pratique clinique toutefois, souvent, plusieurs autres traitements peuvent entrer en ligne de compte pour la même indication sans qu'ils n'aient été comparés à un placebo ni entre eux dans des études comparatives (*head-to-head*). Les méta-analyses en réseau tentent de combler cette lacune en combinant des comparaisons directes et des comparaisons indirectes (1-3). Là aussi, l'objectif est de pouvoir estimer l'effet avec une plus grande précision. Minerva a déjà traité à plusieurs reprises de l'utilité et des pièges des méta-analyses en réseau (2,3). Pour permettre de tirer des conclusions fiables, les études incluses doivent être suffisamment homogènes, comparables et cohérentes (3).

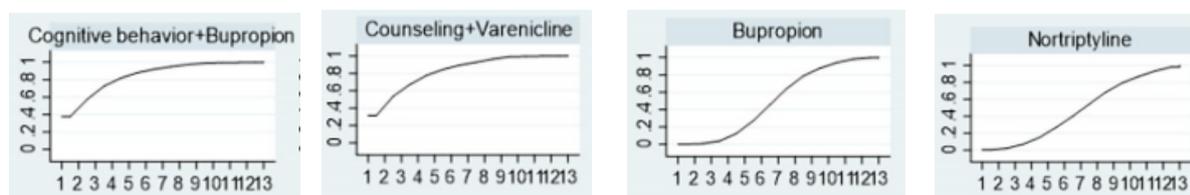
Il est souvent difficile d'interpréter correctement les résultats d'une méta-analyse en réseau parce qu'on est surchargé d'estimations de l'effet provenant des différentes comparaisons directes et indirectes (4). En outre, chaque estimation est elle-même associée à un certain degré d'incertitude. Pourtant, le clinicien doit pouvoir faire un choix ; il se demande avant tout : « Quel pourrait être le meilleur traitement dans cette indication ? » Les résultats de la méta-analyse en réseau permettent de représenter graphiquement pour chaque traitement la probabilité (0,0 à 1,0 sur l'axe des Y) qu'il soit le meilleur, le deuxième parmi les meilleurs, le troisième, etc. (axe des X). Dans l'exemple ci-dessous, le traitement A a 80% de chances d'être le meilleur traitement, et le traitement C a 0% de chances d'être le meilleur traitement (5).



Si l'on compare un grand nombre de traitements, à un certain moment, il devient difficile de comparer les traitements à partir de ces graphiques (4). Il est donc utile de classer les traitements étudiés par probabilités cumulées. Pour chaque traitement, on peut alors représenter graphiquement la probabilité (0,0 à 1,0 sur l'axe des Y) qu'il soit la meilleure option, mais aussi la probabilité qu'il fasse partie des deux meilleurs traitements, des trois meilleurs, etc. (axe des X). Dans l'exemple ci-dessous, le traitement A a 80% de chances d'être le meilleur traitement, mais il a également 90% de chances d'être parmi les deux meilleurs et 98% de chances d'être parmi les trois meilleurs (5).



Ainsi, dans le commentaire sur l'étude portant sur le tabagisme chez les patients BPCO (6,7), nous voyons que la combinaison de thérapie cognitivo-comportementale et de bupropion et la combinaison de thérapie cognitivo-comportementale et de varénicline ont respectivement 40% et 30% de chances d'être le « meilleur » traitement, tandis que le bupropion seul et la nortriptyline seule n'ont, face aux autres traitements, aucune chance (0%) d'être considérés comme meilleure option pour l'arrêt du tabagisme.



Ces courbes laissent entendre que la combinaison de thérapie cognitivo-comportementale et de bupropion est considérée comme la meilleure option, mais... ces graphiques permettent-ils aussi de savoir comment se classent le bupropion seul et la nortriptyline seule l'un par rapport à l'autre ?

Pour pouvoir répondre à cette question, on peut, pour chaque traitement, calculer la **surface sous la courbe** de classement cumulatif (*surface under the cumulative ranking*, SUCRA). Cela permet alors de déterminer le rang de chaque traitement dans le classement du meilleur au moins bon. La valeur varie de 0 à 100%. Plus la SUCRA est importante, plus forte est la probabilité que le traitement se range parmi les meilleurs. Plus la SUCRA est petite, plus forte est la probabilité qu'un traitement se range parmi les moins bons.

Comme on pouvait s'y attendre, le commentaire à propos de l'étude portant sur l'arrêt du tabagisme chez les patients BPCO (6,7) montre que la combinaison de thérapie cognitivo-comportementale et de bupropion a la SUCRA la plus importante, à savoir 85%, suivie par la combinaison de conseils et de varénicline, dont la SUCRA est de 82%. Le clinicien est-il alors tenu de privilégier la combinaison de thérapie cognitivo-comportementale et de bupropion plutôt que la combinaison de thérapie cognitivo-comportementale et de varénicline ? Il semble aussi que le bupropion (51%) se classe mieux que la nortriptyline (47,7%). Il est toutefois surprenant que la SUCRA du bupropion seul soit supérieure à celle de la combinaison de conseils et de bupropion (35,9%). Cependant, l'auteur du commentaire souligne qu'il s'agit d'un résultat hautement improbable qui va à l'encontre des recherches précédentes.

La section suivante portera donc sur la fiabilité de la SUCRA dans la justification du choix d'un traitement particulier (4).

1) La SUCRA ne tient pas compte du degré de preuve des comparaisons directes et indirectes.

Les scores SUCRA peuvent provenir, soit d'un grand nombre d'études avec peu de limites méthodologiques et donc un niveau élevé de certitude des données probantes, soit d'un petit nombre d'études avec un risque de biais élevé (sans préservation du secret de l'attribution, pas d'insu, taux d'abandons élevé et inégal), avec des résultats d'étude qui sont imprécis (larges intervalles de confiance), incohérents (hétérogénéité statistique importante) et indirects (hétérogénéité clinique

importante en termes de population étudiée ou de mesure des résultats), avec en outre un risque élevé de biais de publication et donc un niveau de certitude des données probantes faible à très faible. Il a en outre été démontré que le nombre d'études, la taille de l'étude et la consistance déterminent la validité des évaluations du SUCRA (8).

Le commentaire sur l'arrêt du tabagisme chez les patients BPCO (6,7) montre que, d'après les SUCRA, la combinaison de thérapie cognitivo-comportementale et de bupropion se retrouve en première position et la combinaison de conseils et de varécline obtient la deuxième place. Mais dans le supplément figure un tableau avec les comparaisons directes et indirectes qui montre que les deux comparaisons se basent sur un résultat indirect (aucune étude comparative disponible), avec d'importantes incohérences et imprécisions, ce qui fait que le résultat est associé à un très faible niveau de certitude des données probantes. Si nous tenons compte de ce tableau des données probantes, nous ne pouvons pas tirer de conclusion à propos de la comparaison entre le bupropion et la combinaison de bupropion et de conseils. La classification SUCRA pour la combinaison de thérapie cognitivo-comportementale et de bupropion, pour la combinaison de conseils et de varécline, pour le bupropion et pour la combinaison de bupropion et de conseils est donc très faiblement étayée. Par conséquent, il ne serait pas judicieux de tenir compte de ce classement lors du choix d'un traitement déterminé.

N° of studies	Study design	Certainty assessment					N° of patients		Effect OR with 95% CI	Certainty
		Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Intervention-1	Intervention-2		
Bupropion vs Counseling+Bupropion										
-	RCT	Not serious	Serious	Very serious	Not serious	Undetected	-	-	1.31 (0.42,4.11)	Very Low
Cognitive behavior+Bupropion vs Counseling+Varenicline										
-	RCT	Not serious	Serious	Very serious	Serious	Undetected	-	-	2.69 (0.49,14.80)	Very Low

2) La SUCRA prend en compte un seul critère de jugement, alors que le clinicien est souvent intéressé par plusieurs critères de jugement pertinents.

Un traitement apportant le gain le plus avantageux pour un certain critère de jugement peut apparaître comme le moins bon traitement eu égard à un autre critère de jugement (l'innocuité, par exemple).

Dans un commentaire de Minerva à propos de l'effet des médicaments dans les troubles paniques (9,10) deux critères de jugement pertinents étaient pris en compte. Sur la base de la SUCRA, les benzodiazépines, les antidépresseurs tricycliques et les ISRS ont été classés en tête pour la rémission. Mais, comme les benzodiazépines et les antidépresseurs tricycliques entraînaient plus d'effets indésirables, les ISRS ont été considérés comme constituant le « meilleur » traitement.

3) La SUCRA ne tient plus compte de l'importance de la différence d'effet entre deux traitements. Il est donc difficile pour le clinicien d'évaluer si la différence est cliniquement pertinente.

Par exemple, dans notre exemple d'arrêt du tabagisme chez des patients BPCO (6,7), on voit que la différence entre, d'une part, la combinaison de thérapie cognitivo-comportementale et de bupropion et, d'autre part, la combinaison de conseils et de varécline n'est pas statistiquement significative dans une comparaison indirecte, ce qui fait qu'il n'est même pas possible de déterminer la pertinence clinique.

4) La SUCRA reste le résultat d'une méta-analyse en réseau où les différences entre les traitements peuvent être dues au seul hasard.

La SUCRA ne tient donc pas compte de la largeur des intervalles de confiance des comparaisons directes et indirectes (8,11) On note également que le gain en termes de précision qu'apporte la mise en commun des comparaisons indirectes est plus faible qu'avec les comparaisons directes. Il n'est pas plus important avec quatre études randomisées évaluant des comparaisons indirectes qu'avec une seule étude randomisée contrôlée évaluant une comparaison directe (1). La précision diminue à mesure qu'augmente le nombre de liaisons incluses dans le réseau (1). Il est donc important que le médicament à l'étude soit relié au réseau de manière directe au moins une fois (1).

Conclusion

Les méta-analyses en réseau sont utiles pour comparer différents traitements. La surface sous la courbe de classement cumulatif (*Surface Under the Cumulative Ranking*, SUCRA) pourrait aider le clinicien à interpréter plus facilement le résultat d'une méta-analyse en réseau. Cependant, lorsque le nombre de comparaisons directes est limité et que le niveau de preuve des comparaisons directes et indirectes est faible à très faible, le clinicien ne doit pas prendre en compte le résultat de la SUCRA.

Références

1. Stolk LM. Netwerk meta-analyse van geneesmiddelen. GEBU 2020 ;54 :89-95. DOI: 10.35351/gebu.nl.2020.8.15|13-08-2020
2. Chevalier P. Méta-analyse en réseau : comparaisons directes et indirectes. MinervaF 2009;8(10):148.
3. Chevalier P. Comparaisons indirectes (suite). MinervaF 2011;10(6):77.
4. Mbuagbaw L, Rochweg B, Jaeschke R, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev 2017;6:79. DOI: 10.1186/s13643-017-0473-z
5. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163-71. DOI: 10.1016/j.jclinepi.2010.03.016
6. Boudrez. Efficacité des interventions pour l'arrêt du tabagisme chez les patients atteints de bronchopneumopathie chronique obstructive. MinervaF 2023;22(4):71-4.
7. Wei X., Guo K., Shang X, et al. Effects of different interventions on smoking cessation in chronic obstructive pulmonary disease patients: a systematic review and network meta-analysis. Int J Nurs Stud 2022;136:104362. DOI: 10.1016/j.ijnurstu.2022.104362
8. Daly CH, Neupane B, Beyene J, et al. Empirical evaluation of SUCRA-based treatment ranks in network meta-analysis: quantifying robustness using Cohen's kappa. BMJ Open 2019;9:e024625. DOI: 10.1136/bmjopen-2018-024625
9. Catthoor K. L'effet de différents médicaments dans le trouble panique. Minerva Analyse 21/03/2023.
10. Chawla N, Anothaisintawee T, Charoenrungrueangchai K, et al. Drug treatment for panic disorder with or without agoraphobia: systematic review and network meta-analysis of randomised controlled trials. BMJ 2022;376:e066084. DOI: 10.1136/bmj-2021-066084
11. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. BMC Med Res Methodol 2015;15:58. DOI: 10.1186/s12874-015-0060-8