MÉTA-ANALYSE DES ESSAIS THÉRAPEUTIQUES

Principes, méthodes et lecture critique

J.P. PIGNON [1], T. POYNARD [2]

On peut définir une méta-analyse d'essais randomisés comme l'utilisation de techniques statistiques pour réaliser la synthèse d'un ensemble d'essais distincts mais similaires. La méta-analyse se distingue d'une banale revue d'essais randomisés par son caractère exhaustif et quantitatif (1, 2). Le but de la méta-analyse est de déterminer, par augmentation de puissance, le sens et l'importance de l'effet d'un traitement par rapport à un autre, dans le cas où les résultats des essais existants ne permettent pas de conclure. Par contre, la méta-analyse est moins puissante qu'un essai multicentrique de même taille et permet de répondre à une question moins précise du fait du caractère moins homogène des données analysées. En permettant une évaluation précise et non biaisée de l'efficacité d'un traitement, la méta-analyse peut être un préalable utile à la réalisation d'un grand essai multicentrique.

PRINCIPE ET MÉTHODES STATISTIQUES

L'effet observé d'un traitement est la somme de l'effet « vrai », d'erreurs systématiques, ou biais, et d'erreurs aléatoires. Quand l'effet « vrai » est faible, il est essentiel de réduire au maximum les erreurs aléatoires et les biais si l'on veut pouvoir détecter cet effet. Pour minimiser les erreurs aléatoires, il faut inclure le maximum d'essais et réaliser la meilleure analyse possible, c'est-à-dire la plus puissante. Afin de minimiser les biais (2, 3), il faut utiliser des essais strictement randomisés, idéalement en double insu ou évalués en insu, ou des essais utilisant un critère de jugement objectif ; de plus, il est nécessaire de conduire l'analyse « selon l'intention de traiter » (c'est-à-dire sur l'ensemble des sujets tirés au sort, en les laissant dans leurs groupes initiaux) et sur l'ensemble des essais existants.

La méthode de Mantel-Haenszel-Peto (2-4) est la méthode statistique la plus utilisée. Cette méthode n'est pas propre à la méta-analyse et correspond à un test du chi 2, stratifié sur l'essai. La mesure de l'efficacité d'un traitement par rapport à un autre utilise l'odds ratio, estimateur du risque relatif (cf. figure). Cette méthode donne aux résultats de chaque essai un « poids », dans le résultat global, fonction de la taille de chaque essai. Une autre méthode, basée sur la différence des pourcentages observés, a été proposée par Der Simonian et Laird (4, 5). Cette méthode donne un poids égal à chaque essai et permet de prendre en compte l'hétérogénéité. Le test qui en résulte est moins puissant (1, 5).

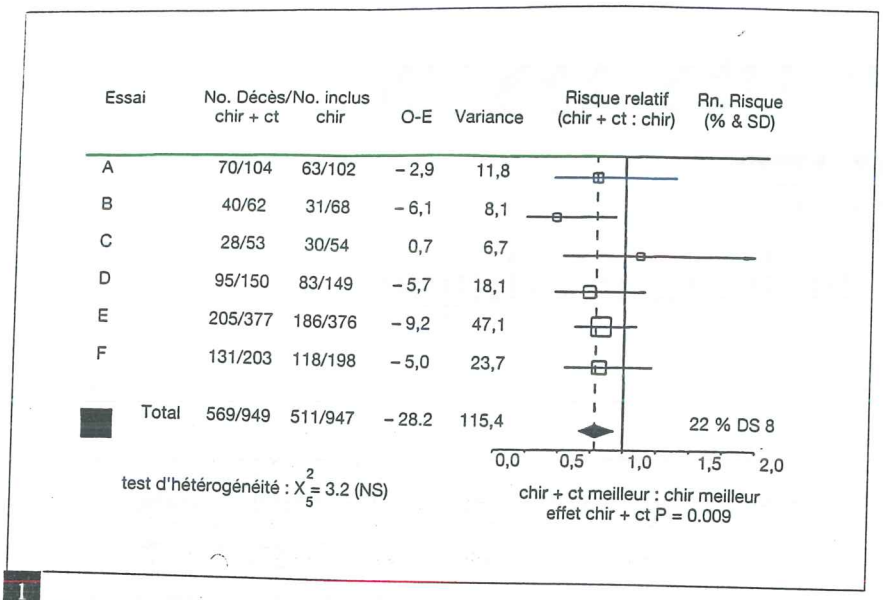
Il existe plusieurs types d'hétérogénéité dont les définitions doivent être bien connues pour éviter les confusions. La première hétérogénéité que peuvent remarquer les cliniciens est l'hétérogénéité clinique entre les essais. Par exemple, un essai peut n'inclure que des cancers du rectum de stade B de Dukes et un autre que des cancers du rectum plus graves de stade C. Certains traitements peuvent également être différents entre deux essais pour la dose ou la durée.

La deuxième hétérogénéité est l'hétérogénéité statistique. On parle d'hétérogénéité quand les différences d'efficacité entre les traitements varient d'un essai à l'autre de façon plus importante qu'une simple fluctuation due à l'échantillonnage. Cette hétérogénéité peut être mesurée et testée, mais les tests sont peu puissants (1). L'étude du graphique des odds ratios de chaque essai avec leur intervalle de confiance est la méthode la plus simple pour détecter l'hétérogénéité (cf. figure).

Les méta-analyses permettent des analyses de sous-groupes, par exemple, la comparaison des résultats des patients de plus de 50 ans à ceux des patients de moins de 50 ans. On peut également réaliser des comparaisons indirectes, par exemple, comparer les résultats de la méta-analyse des essais du tamoxifène pendant un an ou moins versus rien à ceux des essais du tamoxifène pendant plus d'un an versus rien. Les analyses de sous-groupes et les comparaisons indirectes sont très utiles pour la réalisation de futurs essais. Elles doivent être interprétées avec prudence du fait des nombreux biais possibles (3, 4, 6).

La majorité des méta-analyses publiées repose sur des données résumées extraites des publications de chaque essai, parfois complétées par des informations fournies par les auteurs de ces publications (6). Ce type d'étude convient parfaitement soit comme étape préalable à une étude plus poussée, soit quand la différence d'efficacité entre les deux traitements que l'on cherche à détecter est importante, comme la comparaison d'un anti-ulcéreux à un placebo. Plus rarement, les méta-analyses utilisent les données de chaque patient fournies par chacune des équipes ayant réalisé les différents essais inclus (3, 7, 8). La disponibilité des données individuelles permet une vérification directe de la qualité des essais, et donc l'inclusion d'essais non publiés, le

1. Département de biostatistique et d'épidémiologie, Institut Gustave-Roussy, rue Camille-Desmoulins, 94805 Villejuif Cedex
2. Service d'hépato-gastro-entérologie, hôpital Antoine-Béclère, 157, rue de la Porte de Trivaux, 92141 Clamart Cedex et CNRS URA 1484 Paris



1 *Méta-analyse des essais randomisés comparant chirurgie (chir) et chirurgie plus chimiothérapies adjuvantes (chir + ct ; données fictives) : analyses selon la méthode de Mantel-Haenszel-Peto (4).*
Le test de Mantel-Haenszel utilise la différence, dans le groupe traité, entre le nombre de décès observés O et le nombre de décès attendus E sous l'hypothèse d'une efficacité égale des traitements comparés. Dans cette méthode, l'odds ratio correspondant est estimé par l'exponentielle du rapport de la différence (O - E) par la variance de cette différence. Les différences (O - E) sont calculées pour chaque essai et simplement additionnées. Puis cette somme est divisée par la variance correspondante qui est la somme des variances des (O - E) de chaque essai. L'odds ratio (carré vide) de chaque essai est donné avec son intervalle de confiance à 95 %. La taille du carré est proportionnels à la quantité d'informations de chaque essai. Le losange correspond à l'odds ratio global et son intervalle de confiance à 95 %. La diminution relative de la mortalité due à la chimiothérapie (22 %) est donnée avec sa déviation standard (DS).

recueil des données des patients randomisés mais exclus de l'analyse, une actualisation des résultats des essais (suivi plus important des patients), une standardisation du critère de jugement (« hémorragie par rupture de varices œsophagiennes »), le recueil de facteurs pronostiques, essentiel pour les analyses de sous-groupes et parfois utile pour l'estimation de l'effet global. Il en résultera une analyse plus exhaustive et plus puissante (plus grand nombre de patients inclus, utilisation des méthodes d'analyse des données de survie).

LECTURE CRITIQUE

Le tableau ci-dessus donne les 23 critères de qualité d'une méta-analyse proposés par Sacks et al. (2). Ces critères ont été proposés pour les méta-analyses basées sur des données publiées. La qualité des méta-analyses rapportées sous une forme brève (lettres, résumés) ne pouvant être vérifiée, aucune conclusion pratique ne pourra être tirée de telles publications. Il faudra se méfier des méta-analyses réalisées par un seul auteur du fait des risques de biais de sélection et d'extractions des données (critères 11 et 12). Que cette étape soit réalisée, sans connaissance du nom du journal qui publie l'essai, ni connaissance du nom des auteurs de l'étude par les investigateurs, nous apparaît comme un point mineur par rapport à l'exhaustivité de la recherche bibliographique et la qualité des essais. Une étude de la concordance interinvestigateurs (critère 13) sera bienvenue dès que la sélection des études pose des problèmes et en cas d'inégalité importante de la qualité des essais. L'existence d'un protocole préalable (critère 1) permet de garantir l'absence d'influence des premiers résultats de la méta-analyse sur le choix des critères de sélection des essais et de la stratégie d'analyse.

La première question à se poser, devant une méta-analyse, est la pertinence du problème posé : quelle est la taille de l'effet attendu ? Est-il intéressant de démontrer un tel effet dans la pathologie concernée compte tenu de la fréquence de cette maladie et des effets secondaires connus du ou des traitement(s) étudié(s) ? S'il est répondu oui à cette question, il faut regarder la liste des essais analysés (critère 3) et la liste des essais non inclus, pour vérifier l'exhaustivité de l'étude. En effet, un recensement aussi complet que possible des essais réalisés doit être fait pour éviter le biais lié à la sélection d'un sous-ensemble d'essais dont les résultats seraient différents de ceux de l'ensemble des essais existants, en particulier celui résultant de la non-publication des essais négatifs [biais de publication, (9)]. La recherche de la publication des essais, à l'aide de banques de données informatisées, est souvent insuffisante, du fait d'une mauvaise indexation des articles (10). Cette recherche doit être complétée par d'autres méthodes : *Current contents*, résumés des congrès, etc. Les méthodes utilisées pour la recherche bibliographique doivent donc être précisées (critère 2). Vérifier si vous êtes d'accord avec les critères de sélection et d'exclusion (critère 4) des essais est également essentiel, car le choix de ces critères est une des étapes les plus délicates d'une méta-analyse. Il est indispensable de vérifier que, seuls, les essais randomisés ont été inclus (critère 5). L'utilisation de grilles d'évaluation (2) a été proposée par certains pour évaluer la qualité des publications ; mais les données publiées ne reflètent pas toujours fidèlement la qualité réelle d'un essai. Le nombre de patients exclus de l'analyse dans chaque essai constitue avec la vérification de l'absence de biais lors de la randomisation, l'un des points majeurs de l'évaluation de la qualité méthodologique des essais (critère 19). Ainsi, si la différence d'efficacité entre les deux traitements est faible ou si la proportion moyenne de patients exclus de l'analyse des essais est importante, il faut vérifier si les auteurs de la méta-analyse ont contacté les équipes ayant réalisé les essais pour obtenir des informations complémentaires sur les patients exclus pour pouvoir réaliser une analyse

selon l'intention de traiter. Si non, des biais peuvent très bien expliquer les résultats (7).

Une description soigneuse des essais inclus (critères 6 à 8) est indispensable pour juger de la « combinabilité » des essais. Le problème de l'hétérogénéité clinique (critère 9) nous paraît ici plus important que celui de l'hétérogénéité statistique. L'utilisation de graphiques, comme ceux donnant les odds ratios de chaque essai (cf. figure) est très utile pour apprécier globalement l'hétérogénéité (1-3). Les différences entre les populations des essais inclus ou entre les critères de jugement retenus permettent d'expliquer certaines discordances de la littérature. Nous avons vu les caractéristiques essentielles des méthodes statistiques les plus souvent utilisées (critère 15). Si la méthode décrite sort des méthodes classiques, elle devra être référencée et décrite au moins brièvement. L'utilisation d'intervalle de confiance est indispensable (critère 17) ainsi que la description des risques statistiques (critère 16) : degré de significativité en cas de résultat significatif, calcul de puissance en cas de résultat non significatif. Les problèmes d'hétérogénéité statistique (critère 10) devront être discutés. L'étude de la robustesse des résultats (2, 6) évalue l'impact, sur les conclusions de l'analyse, de certaines modifications : par exemple, un changement de méthode statistique, ou un changement de critère de jugement, ou l'inclusion ou non d'essais, soit de mauvaise qualité, soit concernant certaines modalités de traitements ou certains sous-groupes de patients (critère 20). L'étude de critères de jugement moins objectifs mais plus spécifiques, survie sans récidive au lieu de la survie par exemple, peut être envisagée pour confirmer l'analyse principale (3). L'inconvénient de tels critères est qu'ils peuvent correspondre à une réalité différente d'un essai à l'autre et qu'ils ne sont pas toujours disponibles. L'étude de la robustesse de l'analyse donne une plus grande validité aux résultats ; elle peut permettre de comprendre certains résultats discordants. Cette étape est surtout indispensable quand les résultats ne sont pas très significatifs, que la qualité des essais est variable, et que les critères de sélection des essais sont discutables. De même, la discussion du biais de publication (critère 21) est importante en l'absence de différence très significative. L'utilisation systématique de plusieurs méthodes statistiques n'est pas nécessaire quand les différences sont très significatives et qu'il n'existe pas d'hétérogénéité importante. La présence d'analyse de sous-groupe ou de comparaison indirecte pertinente (critère 18) peut renforcer les résultats globaux et être à l'origine de nouvelles hypothèses. Le résultat d'une méta-analyse doit être interprété en tenant compte de nos connaissances sur la maladie et sur le traitement étudié (3, 6). La mise en évidence d'un effet modeste ne sera intéressante que si le traitement étudié a peu ou pas d'effet secondaire, et qu'il est facile à mettre en œuvre. Les informations sur l'impact clinique et économique (critères 22 et 23) ainsi que l'existence ou non de support financier (critère 14) sont également des éléments utiles à la discussion, une fois écartée l'existence de biais pouvant expliquer les résultats.

TABLEAU

Critères de qualité d'une méta-analyse
[Adapté à partir de (2)]

A. Caractéristiques générales de la méta-analyse
1. Protocole préalable
2. Recherche bibliographique exhaustive
3. Liste des essais analysés
4. Liste des essais non inclus et raisons d'exclusion
5. Limitation de l'étude aux essais randomisés
6. Description des caractéristiques des patients inclus dans chaque essai
7. Description des traitements utilisés dans chaque essai
8. Description des formes cliniques de la maladie étudiée incluses dans chaque essai

B. « Combinabilité »
9. Hétérogénéité clinique
10. Hétérogénéité statistique

C. Prise en compte et discussion des biais potentiels
11. Biais de sélection
12. Biais d'extraction des données
13. Concordance interinvestigateurs
14. Existence d'une aide financière

D. Analyse statistique
15. Méthode statistique adéquate
16. Description des risques statistiques (α, β)
17. Utilisation d'intervalles de confiance
18. Analyse de sous-groupes

E. Robustesse de l'analyse
19. Évaluation de la qualité des essais
20. Analyses multiples en fonction des critères d'inclusion, des méthodes d'analyse, et des critères de jugement
21. Discussion du biais de publication

F. Implications des résultats
22. Impact clinique
23. Impact économique

CONCLUSION

Les difficultés de la méta-analyse résident dans le choix d'une question pertinente et dans la disponibilité des essais permettant de

répondre à cette question. Le succès de la méta-analyse réside dans sa capacité à résumer un grand nombre d'études. Des essais de meilleure qualité et la constitution de registres d'essais contribueront à la diminution des biais associés à toutes les revues d'essais randomisés. L'utilisation des données individuelles de chaque essai permet des méta-analyses de meilleure qualité (« supra-méta-analyse ») et plus puissantes qu'une méta-analyse n'utilisant qu'un résumé de l'information disponible dans chaque essai. La collaboration qu'elle implique entre différents groupes est très fructueuse dans la réalisation et l'interprétation (conclusions et perspectives) de la méta-analyse et peut conduire à la réalisation d'essais internationaux. ■

Remerciements
Nos remerciements vont à Mme S. Iacobelli pour son assistance technique dans la réalisation de ce manuscrit et à M. S. Koscielny pour l'écriture du logiciel graphique utilisé pour la figure.

RÉFÉRENCES

1. Ellenberg SS. Meta-analysis : the quantitative approach to research review. Semin Oncol 1988 ; 15 : 472-81.

2. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Charlmers TC. Meta-analysis of randomized controlled trials. N Engl J Med 1987 ; 316 : 450-5.

3. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31 000 recurrences and 24 000 deaths among 75 000 women. Lancet 1992 ; 339 : 1-75, 15-85.

4. Pignon JP, Poynard T. Méta-analyse d'essais thérapeutiques. Gastroenterol Clin Biol 1991 ; 15 : 229-38.

5. Der Simonian R, Laird N. Meta-analysis in clinical trials. Controlled Clin Trials 1986 ; 7 : 177-88.

6. Poynard T, Pignon JP. Duodenal ulcer. Analyses of 293 randomized clinical trials. London, Montrouge : John Libbey Eurotext, 1989.

7. Pignon JP, Arriagada R, Ihde DC et al. A meta-analyses of thoracic radiotherapy for small cell Lung Cancer. N Engl J Med 1992 ; 327 : 1618-24.

8. Simes J. Publication bias : The case for an international registry of clinical trials. J Clin Oncol 1986 ; 4 : 1529-41.

9. Poynard T, Calès P, Pasta L et al. Beta-adrenergic-antagonist drugs in prevention of gastrointestinal bleeding in patients with cirrhosis and esophageal varices. N Engl J Med 1991 ; 324 : 1532-8.

10. Poynard T, Conn HO. The retrieval of randomized clinical trials in liver disease from medical literature. Controlled Clin Trials 1985 ; 6 : 271-9.

MORCEAUX CHOISIS

La morphinomanie ravage le monde médical

Les profanes se figurent volontiers que les docteurs sont indemnes des tares qu'ils soignent et échappent aux tentations dont ils combattent les funestes conséquences. C'est là une grave erreur. Trop souvent les facilités qu'il doit à son métier tournent la tête du gynécologue, du psychiatre, du neuropathologiste et le font tomber dans un ou plusieurs des pièges contre lesquels il met en garde sa clientèle. Un savant n'est pas toujours un saint ni un sage, loin de là.

Si par exemple un Paul Sollier et un Erlenmeyer – spécialistes éminents du traitement des intoxications chroniques – n'étaient pas liés par le secret professionnel, ils étonneraient bien des gens en donnant la proportion des confrères qui viennent annuellement réclamer leurs soins. La morphinomanie ravage le monde médical. Elle est particulièrement grave chez les gens qui ont entre leurs mains la vie et l'honneur d'autrui, et qu'une erreur par excès dans une ordonnance ou une indiscrétion euphorique risquent de transformer en criminels. Les anecdotes comiques et terribles foisonnent sur ce sujet. Je me bornerai à celle-ci :

Il y avait en Allemagne un illustre praticien morphinomane, le doktor Westphal – il a donné son nom à un réflexe –, lequel avait lui-même comme élève un certain Levinstein, inventeur de la méthode de démorphinisation dite « suppression brusque ». Rien de plus simple que ce procédé, très tudesque, aujourd'hui complètement abandonné. On enferme le malade dans un cabanon capitonné, meublé d'un simple matelas, et on l'y laisse pendant soixante-douze heures. Au bout de ce temps, on le retire, et il a perdu, parfois avec le goût du pain, l'habitude du poison. Levinstein proposa à Westphal de le guérir par ce moyen. Après quelques hésitations, Westphal accepta, entra dans la chambre de torture. La porte se referma sur lui. Il avait été convenu que, sous aucun prétexte, on ne tiendrait compte de ses cris, ni de ses appels. Quand Levinstein ouvrit soixante-douze heures plus tard, il trouva son excellent maître mort, complètement mort. Dans le délire de sa douleur, la plus effroyable qu'on puisse imaginer, il avait déchiré son matelas avec ses dents. Les yeux dilatés, la contraction des mâchoires, la torsion des pieds et des mains disaient assez le supplice invraisemblable que le malheureux avait enduré... Je me hâte d'ajouter qu'aujourd'hui, grâce à une technique perfectionnée et plus humaine, Sollier et Erlenmeyer obtiennent le sevrage radical en très peu de jours. Néanmoins la convalescence est longue et les rechutes sont toujours à craindre. Ce n'est certes pas une petite affaire que d'exorciser le démon de l'opium ou de la coca.

La morphinomanie, quand on l'a étudiée de près, est un des vices les plus faciles à déceler, un de ceux qui avouent davantage. Pour ma part, je distingue à première vue, dans la rue, dans un salon, ou au théâtre, d'après leur regard, ceux qui s'adonnent à cette manie tyrannique, les pâles esclaves de la sinistre drogue. Combien, parmi mes premiers maîtres ou mes camarades, sont déjà descendus sous les ombres, leur seringue fatale à la main ! Combien ont interrompu net, à la stupeur générale, une carrière brillante et fructueuse et glissé, lentement d'abord, puis plus vite, à la déchéance ! Le monde est plein de ces fantômes qui font les gestes mécaniques de la vie, n'ayant plus que cette unique pensée, leur piqûre, et lui sacrifiant tout le reste.

Léon Daudet. *Souvenirs des milieux littéraires, politiques, artistiques et médicaux. Devant la douleur* (1914-1921). Édition établie par Bernard Oudin. Paris, Robert Laffont, Bouquins.