



## Qualité méthodologique et biais dans les RCTs

P. Chevalier

Pour évaluer la qualité méthodologique des RCTs rassemblées dans une synthèse méthodique, nous faisons souvent référence au score de Jadad, donné par de nombreux auteurs. Nous avons déjà insisté sur la nécessité d'un score de Jadad d'au moins 3 points mais aussi de la présence supplémentaire d'une analyse en intention de traiter. Nous mentionnons aussi dans notre rubrique considérations sur la méthodologie si des biais sont possibles dans l'étude analysée, en évoquant à ce propos la validité interne de l'étude. Ces termes de validité, qualité et biais sont-ils équivalents ? De quels moyens disposons-nous pour les évaluer ? Quelle est la place du score de Jadad dans ce domaine ?

### • Validité, qualité et biais

Les termes de validité, qualité, biais sont souvent utilisés les uns pour les autres, mais à tort. En distinguer le sens est utile : la validité d'une étude est à la fois externe, évoquant les notions d'applicabilité et d'extrapolabilité, et interne, termes le plus fréquemment utilisés pour désigner la qualité méthodologique de celle-ci. L'évaluation de la qualité méthodologique d'une étude comporte la recherche de biais (évaluation de la fiabilité des résultats) mais aussi un jugement sur d'autres éléments (approbation éthique, calcul d'échantillon correct, rapport conforme à un consensus de référence,...). Pour rappel, un biais est une erreur systématique ou déviation de la vérité dans des résultats ou des conclusions. Ce biais peut aussi bien surestimer ou sous-estimer le résultat réel ou ne pas l'influencer. Il est donc plus correct de parler de risque de biais.

### • Score de Jadad et alternatives

Le score de Jadad évalue en fait le risque de certains biais dans l'étude mais il est de plus en plus considéré comme fort incomplet<sup>1</sup>. Il a cependant l'avantage de présenter un score chiffré, aisément reproductible. La Collaboration Cochrane propose une approche différente, sans établir de score ou de check-list, mais en explorant 6 domaines de risque de biais dans les études en groupes parallèles. Elle évalue si le risque est présent, absent ou incertain dans les domaines de l'élaboration de la séquence et du respect du secret de l'attribution pour la randomisation, de l'aveugle (insu), de données de résultats incomplètes, d'une mention sélective de résultats, d'autres biais potentiels. Quelques explications sur ces notions et leur évaluation.

Une *élaboration de séquence de randomisation* est correcte si elle propose une table de chiffre de référence aléatoire, une attribution par pile ou face ou technique correspondante. Elle est incorrecte si elle se base sur une donnée non aléatoire (date de naissance, jour ou numéro d'admission, préférence du patient ou du médecin).

Un *secret d'attribution* est correct si cette attribution est centrale, avec attribution séquentielle numérotée d'enveloppes ou du traitement. Il n'est pas correct si les participants ou les investigateurs peuvent le mettre à jour et modifier l'attribution : liste ouverte de numéros aléatoires, alternance ou rotation.

L'*insu* est évalué « correct » s'il est bien respecté (participants et chercheurs), s'il est jugé non nécessaire (critère et résultat non influencés par cette absence d'insu) ou si l'évaluation est en aveugle et que l'absence d'aveugle pour les participants et chercheurs n'introduit pas de biais possible. Il est jugé « incorrect » s'il est absent et que ce manque peut influencer les résultats.

Les *données sont jugées complètes* si elles sont toutes présentes, ou en partie manquantes mais de façon similaire et pour des raisons similaires dans les différents bras, ou en partie manquantes mais ne pouvant influencer l'ampleur du résultat, ou en partie manquantes mais imputées de manière appropriée. Elles sont jugées incomplètes dans les autres cas.

La *mention de résultats* est dite *non sélective* si des résultats sont donnés pour tous les critères spécifiés dans le protocole original. Elle est *sélective* en cas de résultats absents pour certains critères pré spécifiés, obtenus avec des instruments de mesure ou d'analyse non pré spécifiés, sur un critère primaire non pré spécifié.

De nombreux *autres risques de biais* doivent être écartés, parmi lesquels : biais dans le protocole d'étude, arrêt prématuré, caractéristiques initiales fort différentes, fraude, autre problème important.

Ce mode d'évaluation permet de mettre en évidence les biais les plus « classiques » : un biais de sélection (caractéristiques initiales entre les bras), de performance (soins complémentaires apportés ou expositions complémentaires à des risques), de migration (arrêts de traitement), de détection (évaluation des résultats), de notification (observations notifiées et non notifiées).

### • Domaines ou scores ?

Une publication récente<sup>2</sup> montre une faible corrélation entre différentes approches d'évaluation du risque de biais (domaines proposés par la Collaboration Cochrane, score de Jadad ou, encore, approche de Schultz (élaboration et secret d'attribution, ITT, double aveugle)<sup>3</sup>. Cette publication montre aussi un agrément inter observateurs variable, parmi 5 collaborateurs d'un même service universitaire, selon les domaines proposés par la Cochrane Collaboration, généralement pour ceux nécessitant un jugement plus pointu sur le potentiel de certaines méthodes de l'étude de provoquer un biais. Cette technique Cochrane semble donc nettement moins facilement reproductible, au moins dans certains domaines. Le problème, en l'absence d'un score avec seuil de validité chiffré pour l'évaluation d'une recherche, est la détermination finale d'une « valeur » décidant de l'inclusion ou non de ces résultats d'étude dans une méta-analyse. Il reste aussi à déterminer si ce problème en est réellement un : les divergences inter observateurs conduisent-elles à des choix différents qui remettent en cause la fiabilité des résultats sommés ?

### Références

- Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from <http://www.cochrane.org/resources/handbook/>
- Hartling L, Ospina M, Liang Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ* 2009;339:b4012.
- Schultz KF, Chalmers I, Hayes RJ, Altman DC. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.