

Evaluation d'un traitement et fiabilité des outils de mesure

ou comment prendre des mesures pour la dépression quand on ne sait pas comment prendre les mesures...

Seuils pour l'ampleur d'efficacité mais non pour l'efficacité

Quand nous avons analysé dans Minerva¹ la méta-analyse de Kirsch et coll.² qui jetait un méchant pavé dans la mare de l'efficacité des antidépresseurs, nous avons abordé la discussion du seuil contestable de pertinence clinique de l'efficacité de ces médicaments. Le seuil de pertinence clinique choisi par les auteurs de cette méta-analyse pour montrer l'efficacité des antidépresseurs versus placebo n'était atteint que pour les formes sévères de dépression majeure. Cependant, le seuil arbitrairement choisi par Kirsch et coll. (et par NICE) n'était pas, à l'origine, destiné à montrer une différence entre efficacité et non efficacité clinique. A l'origine, ces valeurs repères avaient été proposées uniquement pour pouvoir comparer les différentes tailles d'effet de traitements³ : valeurs de DMS de 0,20 (peu d'effet), 0,50 (effet modéré) et 0,80 (effet important) selon Cohen. Kirsch a suivi les auteurs de guidelines de NICE dans leur choix d'un seuil arbitraire de 0,50 pour la DMS pour déclarer que le traitement se montre efficace. Ce choix d'un seuil d'efficacité clinique pertinente prêtant à discussion ouvrait le doute sur la fiabilité des conclusions des auteurs.

Fiabilité de l'instrument de mesure

Un pas supplémentaire dans l'approche critique des études concernant les antidépresseurs vient d'être franchi. Nous avons fait référence dans une analyse publiée dans la revue Minerva⁴ à la méta-analyse sur données individuelles de Fournier et coll.⁵. Cette synthèse montrait aussi que le bénéfice à attendre d'un antidépresseur versus placebo chez des patients présentant une dépression légère ou modérée était inférieur au seuil de 0,20 au score de Cohen. Elle montrait les seuils d'efficacité en fonction d'un score Hamilton Depression Rating Scale (HDRS) initial suivants : effet modéré (0,50 de Cohen) à partir d'un score HDRS initial de 25, effet plus grand (0,81 de Cohen) à partir d'un score HDRS initial de 27. La taille d'effet n'est modérée (= seuil d'efficacité pour NICE et KIRSCH) qu'à partir d'un score HDRS > 27, soit une dépression très sévère (≥ 25).

Une publication d'Isacsson et coll. fin 2011⁶ remet maintenant en cause l'outil de mesure HDRS utilisé dans les différentes études, dont celles reprises dans la méta-analyse de Fournier. Isacsson et coll. ont pu disposer des données individuelles des patients inclus dans 5 des 6 RCTs reprises dans la méta-analyse de Fournier (597 patients), notamment pour le score HDRS initial et son évolution. Ils montrent que la précision de l'évaluation au score HDRS diminue quand le niveau de sévérité de dépression diminue, donc au fur et à mesure que l'état du patient s'améliore, rendant une compa-

raison de l'évolution clinique de patients avec des niveaux de sévérité initiale de dépression différents pour ce score peu fiables. Ils notent aussi que les patients inclus dans les différentes RCTs avaient une perception différente des items du HDRS, ce qui introduit un biais supplémentaire. Ils concluent qu'une ampleur faible de l'effet peut être liée à la faible précision de l'outil de mesure et à sa faible sensibilité pour évaluer une modification, particulièrement en cas de dépression de sévérité légère ou modérée.

Lors du consensus de l'INAMI sur les antidépresseurs en 2006⁷, Van Praag⁸ avait relevé les incertitudes dans le domaine du diagnostic de la dépression, soulignant que les constructions diagnostiques sont hétérogènes d'un point de vue symptomatologie, se superposent entre elles, ne sont pas suffisamment circonscrites par rapport aux diagnostics proches et à la normalité. Il notait aussi que le facteur sévérité est insuffisamment inclus, que les symptômes ne sont pas « pondérés », qu'ils ne sont pas jugés en fonction de leur « valeur » diagnostique, que le facteur psychogenèse est négligé et que la valeur prédictive des constructions diagnostiques est minime. Il insistait sur la nécessité de prendre en compte, dans l'évaluation du facteur psychogenèse, la détermination du niveau de stress, de la sévérité des facteurs de stress, des capacités de résistance personnelle au niveau de la personnalité, l'influence possible d'éléments de personnalité sur le trouble dépressif. Il concluait à l'urgence d'un tel changement dans la démarche diagnostique.

Il n'a guère été entendu à ce jour...

Références

1. De Meyere M. Un lien entre l'efficacité des nouveaux antidépresseurs et la sévérité de la dépression ? *MinervaF* 2008;7(9):134-5.
2. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008;5:260-7.
3. Turner EH, Rosenthal R. Efficacy of antidepressants. [Editorial] *BMJ* 2008;336:516-7.
4. Declercq T. Dépression mineure : antidépresseurs et benzodiazépines efficaces ? *MinervaF* 2011;10(7):84-5.
5. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;303:47-53.
6. Isacsson G, Adler M. Randomized clinical trials underestimate the efficacy of antidepressants in less severe depression. *Acta Psychiatr Scand* 2011;1-7 doi: 10.1111/j.1600-0447.2011.01815.
7. L'usage efficient des antidépresseurs dans le traitement de la dépression. Consensus de l'INAMI Bruxelles 11/05/2006.
8. Van Praag HM. Onvrede met de diagnostiek van depressie. Implicaties voor het onderzoek van antidepressiva. Communication lors du Consensus de l'INAMI du 11/05/2006 L'usage efficient des antidépresseurs dans le traitement de la dépression.