

Analyse des variables continues au moyen d'une ANCOVA

Tom Poelman, Vakgroep Huisartsgeneeskunde en Eerstelijns-gezondheidszorg, Universiteit Gent

Une récente étude contrôlée randomisée (voir analyse p. 101) a évalué l'effet de l'acupuncture et celui d'un service de conseil sur la gravité des symptômes de dépression chez les patients dépressifs^{1,2}. Comme mesure de la gravité, les investigateurs ont utilisé le score **PHQ-9** (qui varie de 0 à 27, les scores 5 à 9 indiquant une dépression légère, les scores 10 à 14 une dépression modérée, les scores 15 à 19 une dépression modérée sévère, et les scores 20 et au-delà, une dépression sévère). Ce score a été déterminé chez tous les participants une première fois après la randomisation et une deuxième fois après 3 mois de suivi.

Score PHQ-9	Acupuncture + prise en charge habituelle	Service de conseil + prise en charge habituelle	Uniquement prise en charge habituelle
En début d'étude	15,3 (ET 5,33)	16,6 (ET 5,27)	16,2 (ET 5,09)
Après 3 mois	9,4 (ET 6,33)	10,9 (ET 6,45)	12,7 (ET 6,47)

À partir de ces chiffres, les investigateurs pouvaient analyser l'effet de l'acupuncture de 2 manières différentes :

- soit par comparaison du score final moyen (après suivi) dans le groupe acupuncture + prise en charge habituelle avec celui du groupe prise en charge habituelle uniquement, ce qui aurait donné comme résultat une différence au niveau du score PHQ-9 de $9,4 - 12,7 = 3,3$ points. Dans ce cas, on aurait pu dire qu'après 3 mois, le score PHQ-9 moyen était 3,3 points plus bas dans le groupe avec acupuncture que dans le groupe prise en charge habituelle uniquement
- soit par comparaison du score différentiel moyen (score après suivi moins le score de début d'étude) dans le groupe acupuncture + prise en charge habituelle avec celui du groupe prise en charge habituelle uniquement, ce qui aurait donné comme résultat une différence au niveau du score PHQ-9 de $(9,4 - 15,3) - (12,7 - 16,2) = -2,4$ points. Dans ce cas, on aurait donc pu dire qu'après 3 mois, le score PHQ-9 moyen avait diminué de 2,4 points en plus dans le groupe avec acupuncture que dans le groupe prise en charge habituelle uniquement.

De ce qui précède, nous pouvons déduire que les 2 méthodes ne donnent un même effet thérapeutique estimé que si les scores moyens de départ sont comparables dans les 2 groupes. En outre, la chance de trouver une différence statistiquement significative pour chacune des méthodes dépendra de la corrélation entre les valeurs initiales et finales. Si la corrélation est faible, une différence significative entre les scores finaux moyens sera plus facile à montrer, tandis qu'une différence significative entre les scores différentiels moyens sera plus difficile à montrer. Ce sera l'inverse en cas de forte corrélation entre les valeurs initiales et les valeurs finales. La méthodologie est incorrecte lorsque les investigateurs post hoc ont choisi, parmi ces 2 techniques d'analyse, celle qui donne le résultat le plus significatif.

Comme, dans notre exemple, la valeur de départ du score PHQ-9 n'était pas la même dans les 2 groupes, les scores finaux moyens ont moins de chance de montrer une différence statistiquement significative.

Même quand on décide d'utiliser les scores différentiels moyens, il faut encore tenir compte d'un effet statistique important, à savoir la **régression vers la moyenne** (regression to the mean), laquelle est responsable d'une diminution globalement plus importante des scores lorsqu'au départ ils sont plus élevés^{3,4}. Comme le groupe prise en charge habituelle a inclus un plus grand nombre de personnes dont le score de départ était plus élevé, l'effet de l'acupuncture sera donc sous-estimé. Une analyse de covariance (ANCOVA) permet d'éviter cette distorsion⁴. En fait, c'est une analyse par régression qui corrige le score final de chaque patient en fonction de son score de départ. Par conséquent, l'évaluation de l'effet est indépendante de la différence entre les groupes quant au score initial moyen, et, par la même occasion, le phénomène de régression vers la moyenne est neutralisé. Dans notre exemple, cette technique donne une différence de -2,46 points (IC à 95% de -3,72 à -1,21 points) entre le groupe acupuncture et le groupe prise en charge habituelle uniquement. Cette valeur correspond approximativement à l'estimation de l'effet obtenue à partir des scores différentiels moyens (voir plus haut). Car la corrélation entre les scores initiaux et les scores de suivi, dans notre exemple, était effectivement suffisamment forte (l'ET du PHQ-9 varie peu entre le début de l'étude et après 3 mois de suivi) et le phénomène de régression vers la moyenne n'est donc pas complètement neutralisé. Les investigateurs ont cependant tout de même opté pour cette technique ANCOVA car elle offre une plus grande puissance statistique pour montrer une différence entre les groupes⁴. Et comme l'analyse de covariance est une forme d'analyse par régression multiple, elle présente en outre l'avantage de permettre de corriger l'analyse pour tenir compte de certaines variables pronostiques, comme l'âge.

Conclusion

L'analyse de covariance (ANCOVA) est une analyse par régression multiple qui permet d'estimer correctement un effet thérapeutique pour comparer 2 groupes dont les résultats pré et post intervention ont été correctement pris en compte.

Références

1. MacPherson H, Richmond S, Bland M, et al. Acupuncture and counselling for depression in primary care: a randomised controlled trial. *PLoS Med* 2013;10:e1001518.
2. Declercq T, Poelman T. Acupuncture et conseil comme approche complémentaire de la dépression en première ligne ? *MinervaF* 2014;13(8):101-2.
3. Bland JM, Altman DC. Some examples of regression towards the mean. *BMJ* 1994;309:780.
4. Vickers AJ, Altman DC. Statistic notes: Analysing controlled trials with baseline and follow up measurements. *BMJ* 2001;323:1123-24.