

# Estimation de modèles à équations structurelles par algorithme EM pour l'analyse longitudinale de la qualité de vie relative à la santé en cancérologie

M Tami<sup>1</sup>, A Barbieri<sup>1,2</sup>, X Bry<sup>1</sup>, D Azria<sup>3</sup>, S Gourgou<sup>2</sup>, C Bascoul-Mollevi<sup>2</sup>, C Lavergne<sup>1</sup>

(1) Institut de Mathématiques et de Modélisation de Montpellier (I3M), Université de Montpellier

(2) Unité de Biostatistique, Institut régional du Cancer de Montpellier (ICM) - Val d'Aurelle

(3) Département de Radiothérapie Oncologique, Institut régional du Cancer de Montpellier (ICM) - Val d'Aurelle

*Introduction* En oncologie, la qualité de vie relative à la santé (QdV) est devenue un critère essentiel dans les essais cliniques mais son analyse longitudinale reste complexe et non standardisée. Un des freins conceptuels de ce critère est son aspect multidimensionnel. En effet, la QdV comprend au minimum les fonctions physique, psychologique et sociale et les symptômes liés à la maladie ou au traitement. Actuellement, l'analyse longitudinale des données de QdV est réalisée par dimension sur l'étude des scores via des modèles linéaires mixtes. Nous proposons dans ce travail une approche qui permet de tenir compte conjointement de l'aspect longitudinal et de la nature multidimensionnelle de la QdV.

*Méthodes* Les données observées sont issues d'auto-questionnaires collectés à différentes visites du traitement et du suivi. Le questionnaire standard en Europe est le QLQ-C30 développé par l'EORTC. Il décompose la QdV en 15 dimensions : 5 dimensions fonctionnelles, 9 symptomatiques et le statut global de santé (*global health status, GHS*). A chaque temps, on construit un modèle à équations structurelles (*structural equation modeling, SEM*) où le *GHS* (unidimensionnel) est expliqué par deux variables latentes, chacune d'elles étant un facteur résumant respectivement les dimensions fonctionnelles d'une part, et les dimensions symptomatiques d'autre part. Pour tenir compte de l'aspect longitudinal, la variable *GHS* et les deux facteurs seront concaténés sur l'ensemble des visites. Nous pouvons ainsi par un modèle mixte expliquer la *GHS* par les deux variables latentes globales. Ce modèle prendra en compte l'effet individu (par un effet aléatoire) et un effet traitement. Ceci est rendu possible par la maximisation de la vraisemblance de chaque SEM utilisant l'algorithme EM. Cette approche, contrairement aux méthodes classiques dans la littérature, donne une estimation des facteurs ce qui permet de les réintroduire dans un autre modèle (ici le modèle linéaire mixte). Cette technique d'estimation est de plus computationnellement efficace.

*Résultats* Nous nous plaçons dans le cas d'une étude exploratoire où nous voulons comparer la QdV globale entre deux bras de traitement. Nous illustrerons cette méthode développée sous le logiciel R par l'intermédiaire de données issues de l'étude CO-HO-RT. Cet essai clinique de phase II randomisé est mené sur 150 patientes ménopausées

traitées pour un cancer du sein en situation adjuvante avec des récepteurs hormonaux positifs. Il évalue les séquences de traitement de radiothérapie (concomitante ou séquentielle) associée à une hormonothérapie. La mesure de la QdV est réalisée par les questionnaires EORTC-QLQ-C30 et EORTC-QLQ-BR23. Ce dernier est un module complémentaire développé spécifiquement pour le cancer du sein. Il permet la considération supplémentaire de quatre dimensions fonctionnelles et symptomatiques, spécifiques à la pathologie étudiée.

*Conclusions* Cette approche est proposée pour éviter la multiplicité des tests induite par l'analyse indépendante de chaque dimension. Le modèle proposé pour analyser la QdV globale à partir des dimensions sous-jacentes, est intuitif : il découle de la conception du questionnaire QLQ-C30. La méthodologie utilisée a l'avantage d'estimer les facteurs pour chacun des individus ce qui rend possible l'analyse longitudinale.