

Perte d'autonomie fonctionnelle des personnes âgées : aperçu des trajectoires d'autonomie à partir d'une revue de littérature

Loss of functional independence of elderly people: an overview of autonomy trajectories based on a literature review

Charlotte Bimou¹, Michel Harel², Thierry Dantoine¹, Achille Tchalla^{1,3}

¹ Laboratoire HAVAE, Université de Limoges, France, charlotte.bimou@etu.unilim.fr, tdantoine@gmail.com

² ESPÉ de Limoges, Université de Limoges, F-87036 Limoges Cedex, France, michel.harel@unilim.fr

³ Service de Médecine Interne Gériatrique et Biologie du Vieillissement ; CHU de Limoges, F-87042 Limoges Cedex, France, achille.tchalla@unilim.fr

RÉSUMÉ. Afin d'anticiper les mesures préventives pour le soutien et le maintien de l'autonomie fonctionnelle des personnes âgées suivies de façon prospective et par conséquent retarder l'entrée dans la dépendance, une compréhension du concept de trajectoires de perte d'autonomie fonctionnelle est nécessaire. L'objectif principal de cet article est de créer une synthèse lisible des ressources disponibles dans la littérature pour les différentes trajectoires de perte d'autonomie fonctionnelle chez des personnes âgées vivant en communauté et les méthodes de prévention proposées pour maintenir leur autonomie. Les ressources que nous abordons dans ce document sont d'ordre clinique, mathématiques et statistiques.

ABSTRACT. In order to anticipate preventive measures to support and maintain functional autonomy community-dwelling in elderly people followed prospectively and consequently delay loss of independence, a comprehension of the concept of functional independence trajectories is necessary. The aim of this paper is to create a readable synthesis of the resources available in the literature for different functional independence trajectories of older person and the methods of prevention proposed to maintain autonomy. The resources discussed in this document are clinical, mathematical and statistical.

MOTS-CLÉS. Modélisation mathématiques, Trajectoires d'autonomie, Perte d'autonomie fonctionnelle, Personnes âgées, ADL.

KEYWORDS. Mathematical modelling, Trajectories of autonomy, Loss of functional autonomy, Elderly, ADL.

Introduction

Les pays du sud de l'Europe, à l'exception de l'Allemagne, sont les plus vieillissants en raison de l'augmentation du nombre de personnes âgées. Une croissance qui pose de nombreux problèmes (MENDES et al 2016). L'évaluation de l'évolution de l'état d'autonomie fonctionnelle des personnes âgées est un domaine important à comprendre. Cependant, peu de recherches ont été entreprises dans ce domaine et nos connaissances restent limitées. Le présent document comporte un examen de la documentation portant sur l'évolution du score d'autonomie fonctionnelle chez des personnes âgées vivant en communauté. Au cours des dernières années, diverses études longitudinales ont démontré d'importantes différences interindividuelles dans l'ampleur du déclin de l'autonomie fonctionnelle chez les plus âgés. Différences qui pourraient, avec le temps, entraîner une variabilité accrue de l'état d'autonomie au sein d'une population de personnes âgées. On définirait les trajectoires d'autonomie fonctionnelle par l'évolution dans le temps d'un score ou d'un indicateur d'évaluation de l'autonomie. Selon Chen, la trajectoire d'autonomie fonctionnelle est définie comme étant l'évolution de l'état fonctionnel à plusieurs moments dans le temps (CHEN et al 2008). Selon les travaux de Bouvet, certains chercheurs ont pu établir des typologies des trajectoires de vieillissement en distinguant des trajectoires qui se produisent sans incapacité ou perte d'autonomie des trajectoires marquées par l'occurrence d'incapacités. Ils résument le premier

type de trajectoires comme étant un vieillissement normal ou réussi (BOUVET 2016). Le deuxième type de trajectoires est décrit par un vieillissement pathologique (D'EPINAY et SPINI 2007). Bien que de nombreuses études aient abordé cette question, il est difficile d'extraire une liste cohérente et fiable du nombre de groupes de trajectoires et des facteurs de risque pour élaborer des stratégies de prévention en raison de leur variabilité, notamment en ce qui concerne la conception et la méthodologie des études. Selon Chen et ses collaborateurs (CHEN et al 2008), la trajectoire fonctionnelle représente un processus dynamique et signale souvent une déficience ou une amélioration sous-jacente, en particulier chez les patients âgés. Nous avons donc effectué une analyse documentaire systématique de toutes les études longitudinales qui distinguent les trajectoires d'autonomie fonctionnelle et les facteurs ayant une incidence sur l'état fonctionnel chez les sujets âgés vivant en communauté.

Afin d'anticiper les mesures préventives pour le maintien de l'autonomie fonctionnelle des personnes âgées suivies de façon prospective, une compréhension du concept de trajectoires de perte d'autonomie fonctionnelle est nécessaire. L'objectif de notre étude était de fournir un aperçu systématique de la littérature fondée sur les trajectoires d'autonomie fonctionnelle des personnes âgées vivant en communauté ainsi que les méthodes mathématiques, statistiques employées pour identifier les trajectoires.

Le repérage de différents parcours d'autonomie fonctionnelle chez des personnes âgées permet non seulement d'explorer les facteurs de risque et de protection de la perte d'autonomie, de mettre en place une réflexion sur les politiques de santé publique comme le bien vieillir mais aussi de mettre en place des mesures de prévention précoce pour maintenir l'autonomie.

Matériels et méthodes

1. *Description de la méthode et sélection des études*

Les études utilisées dans ce travail ont été repérées grâce à des recherches documentaires dans les bases de données PubMed et Web of Science. Des mots-clés, des titres et des informations abstraites ont été utilisés pour trouver les articles. Les principaux termes de recherche étaient : "trajectoires", "données longitudinales", "trajectoires développementales", "autonomie", "autonomie fonctionnelle", "perte d'autonomie", "perte d'autonomie fonctionnelle", "déclin fonctionnel", "personnes âgées", "séniors". Le français et l'anglais étaient les deux langues de recherche des articles. A partir de ces mots-clés, nous avons intégré dans notre étude les publications incluant des études de trajectoires d'autonomie fonctionnelle de la personne âgée utilisant des indicateurs dont l'ADL (Activities of Daily Living) (KATZ et al 1970), l'IADL (Instrumental Activities of Daily Living) (LAWTON et BRODY 1969), le SMAF (Système de Mesure de l'Autonomie Fonctionnelle) (HEBERT et al 1988). Il existe d'autres outils cliniques pour évaluer la fonction dans les activités de la vie quotidienne. Parmi eux, l'indice de Barthel (MAHONEY et BARTHEL 1965) et le FIM (Functional Independence Measure) (GRANGER et al 1986) (HAMILTON 1987) sont des échelles également utilisées pour l'évaluation fonctionnelle (SHANGHA et al 2005). Une publication incluse dans notre étude présente les caractéristiques qui se rapportent à une population dont l'âge moyen est de 65 ans et plus. Nous avons inclus des études observationnelles longitudinales, des études de cohortes.

2. Définition du statut d'autonomie fonctionnelle

D'après l'examen de toutes les études, une définition similaire de l'autonomie fonctionnelle a été trouvée dans le cadre de notre étude. L'incapacité dans les activités de la vie quotidienne est fréquemment examinée à l'aide de l'indice ADL de Katz (KATZ 1963) parfois combinée avec les activités instrumentales de la vie quotidienne (IADL) (LAWTON et BRODY 1969). Certaines de ces études montrent qu'une personne âgée serait considérée comme dépendante pour les ADL si elle avait besoin d'aide dans l'une des six activités suivantes de l'indice : alimentation, continence, transfert, aller aux toilettes, s'habiller et prendre un bain. Si la personne âgée est en mesure d'exécuter les six activités, elle serait définie comme indépendante ou autonome (BUURMAN et al 2016). Cette définition est similaire à la définition de la perte d'autonomie fonctionnelle donnée par l'Organisation Mondiale de la Santé (OMS 1988).

3. Résultat de la recherche

Un examen manuel des titres de toutes les références identifiées a révélé un total de 11 articles (figure 1) qui ont dû faire l'objet d'un examen plus approfondi en vue de leur inclusion éventuelle dans le présent document. Bien que ces publications présentent différents outils d'évaluation de l'autonomie fonctionnelle de la personne âgée, elles fournissent des informations à peu près similaires. Dans notre travail, nous avons analysé des trajectoires d'autonomie des personnes âgées à partir de modèles statistiques explicites. Après évaluation des publications incluses, nous avons pu identifier les différentes catégories de trajectoires. Le tableau 1 donne un aperçu de ces trajectoires et le texte qui suit donne une description détaillée de nos constatations.

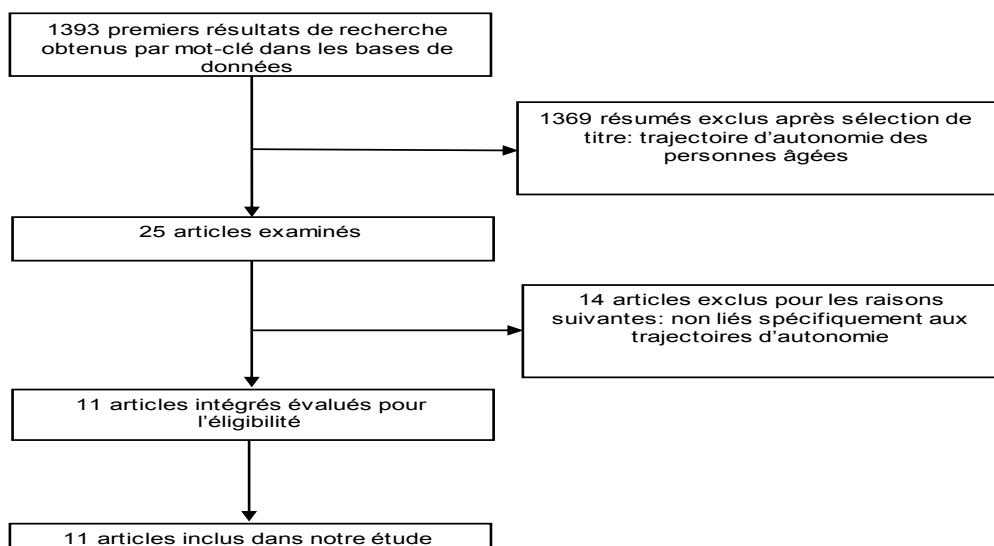


Figure 1.: Organigramme du processus de sélection des études de notre analyse documentaire

La figure 1 représente le processus de sélection pour notre analyse documentaire. La recherche par mot-clé et l'examen manuel de l'information sur les titres nous ont permis de trouver un total de 1393 titres dans PubMed. Après avoir sélectionné les articles qui nous ont fourni les informations les plus intéressantes (trajectoire d'autonomie fonctionnelle de la personne âgée), nous avons sélectionné 25 articles. Quatorze articles ont été exclus de notre étude car ils n'étaient pas particulièrement intéressés par l'autonomie fonctionnelle.

Tableau 3.1: Caractéristiques des études longitudinales de trajectoires d'autonomie prédisant l'issue fonctionnelle chez les personnes âgées

Auteur.s	Année	Titre	Indicateur	k	Description des groupes
Romoren et al	2003	Trajectories of Disability among the Oldest Old	ADL	4	1 : serious dementia 2 : severely dependent 3 : frailty 4 : fairly healthy
Wakefield et al	2007	Functional Trajectories Associated with Hospitalization in Older Adults	ADL	2	1 : stable trajectory 2 : functional decline
Chen et al	2008	Functional trajectory 6 months post-hospitalization : a cohort study of older hospitalized patients in Taiwan	Barthel Index	3	1 : good function 2 : moderate function 3 : poor function
Hybels et al	2009	Trajectories of mobility and IADL function in older patients diagnosed with major depression	IADL	3	1 : low scores of IADL 2 : higher scores of IADL 3 : few to no limitations on the IADL measures
Gill et al	2013	Comparisons Between Older Men and Women in the Trajectory and Burden of Disability Over the Course of Nearly 14 Years	ADL	5	1 : independent 2 : mild disability 3 : mild to moderate disability 4 : moderate disability 5 : severe disability
Han et al	2013	Dynamics of functional aging based on latent-class trajectories of activities of daily living	ADL	5	1 : independent 2 : mild disability 3 : mild to moderate disability 4 : moderate disability 5 : severe disability
Buurman et al	2015	Trajectories of Disability Among Older Persons Before and After a Hospitalization Leading to a Skilled Nursing Facility Admission	ADL	3	1 : minimal disability 2 : mild disability 3 : moderate disability
Nini et al	2017	Predicting Trajectories of Functional Decline in 60- to 70-Year-Old People	ADL	3	1 : no/little decline 2 : intermediate decline 3 : severe decline
Kuo et al	2017	Activities of Daily Living Trajectories Among Institutionalized Older Adults : A Prospective Study	ADL	3	1 : high dependency with gradually declining function 2 : low dependency with gradually declining function 3 : persistent independent function
Wei et al	2018	Impact of the disability trajectory on the mortality risk of older adults in China	ADL	3	1 : independent type 2 : low start, rapidly increasing in late life type 3 : high start, rapid development type
Saito et al	2018	Exploring 2.5-Year Trajectories of Functional Decline in Older Adults by Applying a Growth Mixture Model and Frequency of Outings as a Predictor : A 2010?2013 JAGES Longitudinal Study	LTCL	3	1 : slowly declining 2 : persistently disabled 3 : rapidly declining

Le tableau 1 décrit les articles sélectionnés pour notre étude de revue. La première colonne intitulée "Auteu.r.s" représente le nom des auteurs, la deuxième colonne "Année" correspond à l'année de publication de l'étude ; la colonne intitulée "Titre" illustre le titre de l'article ; "Indicateur" représente l'outil d'évaluation utilisée ; "k" illustre le nombre de groupes de trajectoires identifiés par l'étude ; "Description" représente la description de chaque groupe de trajectoires. Les périodes de référence des études incluses dans notre travail variaient de 2003 à 2019.

La plupart de ces études avaient pour outil de mesure d'autonomie fonctionnelle les ADL (8 ; 73%) ; une petite proportion utilisait les IADL (1 ; 9%), l'index de Barthel (1 ; 9%), Long-Term Care Insurance (LTCI) (1 ; 9%). L'indicateur ADL de Katz est utilisé de manière internationale et appréhende uniquement la perte d'autonomie physique. Ainsi les trajectoires développementales de perte d'autonomie fonctionnelle utilisant les ADLs se réfèrent aux fonctions essentielles pour l'auto-soin. Ces fonctions sont des tâches de base de la vie quotidienne cotées dans le questionnaire des ADLs comme suit ; 1 faire sa toilette, 2 s'habiller, 3 aller aux toilettes et les utiliser, 4 se coucher ou quitter son lit et s'asseoir et quitter son siège, 5 contrôler ses selles et ses urines cotée, 6 manger des aliments déjà préparés (KATZ 1963). Les activités de la vie quotidienne sont fondamentales pour maintenir l'autonomie et la qualité de vie des personnes âgées. La somme de ces scores ADL fait la distinction entre les personnes âgées ayant une autonomie fonctionnelle adéquate et les personnes dépendantes. Un score de 6 indique une autonomie complète ; 4 une dépendance partielle et 2 ou moins une dépendance fonctionnelle sévère (KATZ et AKPOM 1976).

Outre les activités de la vie courante, il existe d'autres activités appelés activités instrumentales de la vie courante (IADL) dont l'élaboration nécessite des capacités cognitives. Le test de Lawton (LAWTON et BRODY 1969) a été développé dans ce contexte et fait office de référence dans la littérature internationale. Pour ce texte, huit activités instrumentales sont évaluées dont 8 pour les femmes et 5 pour les hommes. La cotation se fait par 1 ou 0 que l'activité soit ou non réalisée. Les IADL (LAWTON et BRODY 1969) (GRAF 2008) sont fréquemment évaluées à l'aide de rapports d'auto-évaluation ou de rapport familiaux sur l'application de divers domaines (SUCHY et al 2011). Elles peuvent être aussi fondées sur le rendement qui nécessite que divers activités de la vie quotidienne soient effectuées devant l'examinateur (SUCHY et al 2011). Ces activités désignent des tâches quelque peu complexes comme la cuisine, les courses, l'utilisation du téléphone, l'entretien ménager ou la gestion de l'équipement de soins personnels qui demande des facultés cognitives. Sous cet angle, un grand nombre de travaux de recherche ont été portés sur la découverte des mécanismes de la déficience fonctionnelle, en examinant les associations entre le déclin cognitif et fonctionnel des personnes âgées (SUCHY et al 2011) (GLOSSER 2002) (JEFFERSON 2006). Ces travaux de recherche ont démontré que la capacité de s'engager avec succès dans les IADL repose sur le fonctionnement exécutif la mémoire (CAHN-WEINER et al 2007) (BURTON et al 2006) (GRIGSBY et al 1998) (JEFFERSON et al 2006) (KIOSSES et ALEXOPOULOS 2005) (MARIANI 2008). Ainsi, l'évaluation de l'IADL peut permettre d'identifier un début de déclin physique, cognitif ou les deux chez une personne âgée qui pourrait autrement sembler capable et en santé (WARD et al 1998). L'échelle IADL est composée de huit fonctions qui couvrent les capacités à utiliser le téléphone, faire ses courses, préparer le repas, entretenir la maison, faire la lessive, utiliser les moyens de transport, prendre ses médicaments, et les capacités à gérer son budget. Un score de 8 correspond à une autonomie haute, bonne ; Une autonomie basse correspond à un score de 0 (WADE et HEWER 1987).

L'indice de Barthel des activités de la vie quotidienne a été introduit par Dorothea Barthel en 1955

(MAHONEY et BARTHEL 1965) pour suivre les progrès de l'autonomie et de la mobilité des patients hospitalisés en réadaptation. Il s'agit d'une échelle pondérée simple qui a été bien validée (WADE et HEWER 1987) et qui pourrait être le meilleur parmi les indices ADL courants (GRESHAM et al 1980). La fiabilité des différentes façons d'utiliser l'indice a été étudiée (COLLIN et al 1988).

Le LTCI est un système d'évaluation de la capacité fonctionnelle de la personne âgée mise en place par le gouvernement Japonnais (SAITO et al 2018). Selon ce système, les personnes sont classées dans l'un des sept niveaux de soins en fonction de la gravité de leur déficience physique et cognitive au moyen d'une entrevue à domicile et d'un examen médical. Un niveau de soins plus faible implique une capacité fonctionnelle moins bonne. Par exemple, le niveau 3 de ce test fait référence au "soutien complet nécessaire pour la toilette, le bain, l'habillement et toutes les autres activités quotidiennes de base". Les détails de chaque niveau de soins sont fournis ailleurs (KONDO et al 2012).

Toutes ces études étaient de nature prospective. Le recueil des données a été fait à partir de questionnaires ou des entrevus, des observations directes ou des sources diverses.

4. Méthodes d'analyse des modèles de trajectoires

Les méthodes utilisées par les auteurs de ces publications varient entre l'analyse de trajectoire de classe latente (latent class trajectory analysis) (MCCUTCHEON 1987) également appelée modélisation de trajectoire (NAGIN 2005) utilisée dans les travaux de (CHEN et al 2008) (HYBELS et al 2010) (JONKMAN et al 2018) (WEI et al 2018) et la modélisation linéaire hiérarchique (hierarchical linear modeling) (HYBELS et al 2010) employée par (LIANG et al 2010), le modèle de mélange de croissance généralisé à effet aléatoire (generalized growth mixture model with random effect) (MUTHEN et al 2002) employée par (GILL et al 2013) (HAN et al 2013) (SAITO et al 2018).

Afin d'évaluer la corrélation entre les trajectoires de perte d'autonomie fonctionnelle et les facteurs qui affectent ces trajectoires et de développer des modèles de prévision, les auteurs utilisent la régression logistique (régression logistique multinomiale pour la plupart). Un aperçu des méthodes qui comprennent l'analyse de trajectoire de classes latentes, la modélisation linéaire hiérarchique, la modélisation générale des mélanges de croissance et la régression logistique multinomiale est présenté dans les sections suivantes.

L'objectif de ces méthodes est tout d'abord d'estimer les formes des trajectoires, ensuite d'estimer les probabilités des groupes de trajectoire, de relier les probabilités des groupes de trajectoire aux covariables et ensuite de classer les individus dans les groupes en se basant sur leurs probabilités à postériori.

4.1. Méthodes d'analyse des modèles de trajectoires

L'analyse de trajectoires de classes latentes (Latent class growth analysis (LCGA)) ou analyse de croissance de classes latentes est un type spécial des modèles de mélange (Growth Mixture Modeling (GMM)) dans lequel les trajectoires de croissance individuelle d'une classe sont homogènes (JUNG et WICKRAMA 2008). La technique relève de modèles de mélanges finis qui décompose l'hétérogénéité d'un échantillon en le divisant en nombre fini de sous-groupes latents en fonctions des profils des variables indicatrices (COLE et al 2012). La méthode utilise une seule variable de résultat mesurée à

plusieurs points dans le temps pour définir un modèle de classe latente dans lequel les classes latentes correspondent à différentes formes de courbe de croissance pour la variable d'intérêt (MUTHEN et al 2004). Le modèle est utilisé par Nagin et ses collègues (NAGIN et LAND 1993), (NAGIN et TREMBLAY 1999), (JONES et al 2001), (NAGIN et al 1999) sous l'appellation de Group-Based Trajectory Modeling (GBTM).

Le modèle d'analyse de trajectoires basé sur la méthode Group-Based Trajectory Modeling est une application des modèles de mélange finis fournie par Nagin (NAGIN 2005). Le Group-Based Trajectory Modeling est une méthode d'analyse de trajectoires de groupes qui permet d'étudier des phénomènes qui évoluent dans le temps. La conception de la méthode a été faite dans le but d'identifier des sous-groupes distincts de trajectoires parmi un ensemble de trajectoires suivant des progressions similaires de certains comportements ou résultats avec le temps (NAGIN et al 1999);(NAGIN 2005) (NAGIN et TREMBLAY 1999). Le modèle suppose que la population d'étude est constituée d'un mélange de groupes homogènes définis par les trajectoires (NAGIN et ODGERS 2010). La méthode est de plus en plus utilisée dans la recherche pour cartographier l'étiologie et le développement de nombreux types de symptômes ou de troubles psychologiques, notamment la dépression, l'anxiété, le stress (FAN 2013).

Modèle Statistique

Dans le modèle de trajectoire GBTM, le lien entre la variable explicative et le temps est en général modélisé par un lien polynomial (NAGIN 2005). La forme de la trajectoire et la probabilité pour chaque individu d'appartenir à l'un des groupes de trajectoires sont estimées. Chaque trajectoire est décrite comme une variable latente (variable dont la mesure ne peut pas être directement prise, mais qui est supposée exister à travers d'autres variables observées). Ainsi, le modèle identifie les groupes latents qui existent dans la population (NAGIN 2005). Lors de la modélisation, chaque individu est affecté au groupe pour lequel la probabilité d'appartenance est la plus faible. Le modèle définit également la probabilité pour chaque groupe d'appartenir à une trajectoire. Nagin définit la probabilité d'un des groupes comme étant la différence entre 1 et la somme des deux autres groupes ((NAGIN 2005). Cette probabilité est la probabilité moyenne à la probabilité postérieure. Elle est calculée à partir des paramètres estimés pour chaque individu afin d'examiner l'ajustement du modèle (JONES et al 2001). Le calcul de ces probabilités est basé sur l'application du théorème de Bayes.

Covariables du modèle

La nature longitudinale des données est modélisée en faisant dépendre le(s) paramètre(s) β_k du temps. Les covariables ou facteurs de risque à intégrer dans le modèle peuvent être stables ou non stables dans le temps. Les covariables stables dans le temps sont intégrées dans le modèle en supposant qu'elles influent sur la probabilité d'appartenir à un groupe particulier. Les covariables dépendantes du temps peuvent également affecter directement le comportement observé (JONES et al 2001). Cela est illustré sur la figure (figure 2) ci-dessous.

Lorsqu'elles sont intégrées dans le modèle, les covariables affectent la probabilité d'une trajectoire de données particulière. Soient $X_i = (X_{it}, \dots, X_{iR})$ les covariables pour l'individu i , $Y_i = (Y_{it}, \dots, Y_{iT})$ la

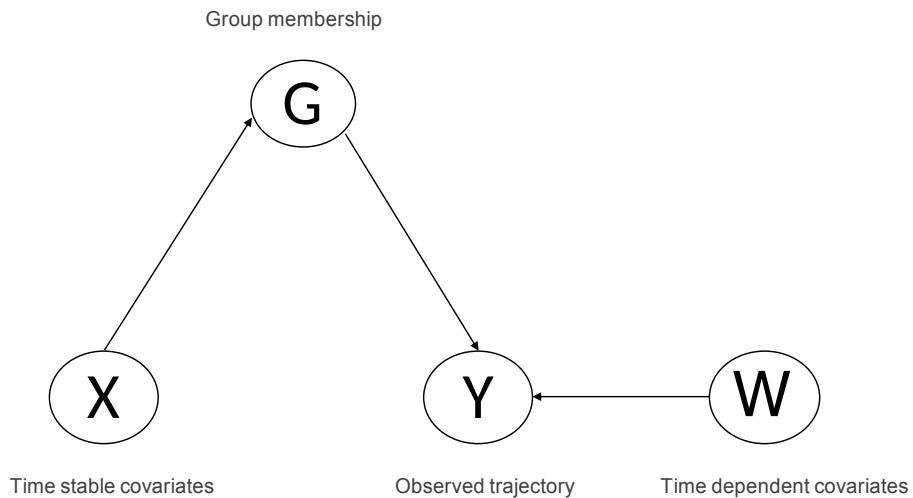


Figure 2.: Directed Acyclic Graph Representing the Independence Assumptions (Jones 2001, p 376)

trajectoire de données de l'individu sur T temps de mesure indépendants étant donné le groupe G_i . En supposant qu'il existe K groupes, nous pouvons écrire la distribution ou probabilité conditionnelle des données d'observation pour l'individu i , à partir des covariables stables du temps et d'une covariable dépendant du temps, $W_i = (w_{it}, \dots, w_{iT})$ comme suit (JONES et al 2001) :

$$f(y_i | x_i, w_i) = \sum_{k=1}^K Pr(G_i = k | X_i = x_i) Pr(Y_i = y_i | G_i = k, W_i = w_i) \quad [1]$$

Dans le modèle GBTM, la tendance de modélisation des trajectoires dépend du nombre de temps de mesure et peut prendre l'une des trois formes : une tendance linéaire, quadratique ou encore cubique. La forme dépend du nombre de temps de mesure (NAGIN et TREMBLAY 1999). Alors, les trajectoires peuvent être constantes, linéaires, quadratiques ou cubiques. L'équation du modèle cubique est :

$$Y_t^* = \beta_0^j + \beta_1^j t + \beta_2^j t^2 + \beta_3^j t^3 + \epsilon_t \quad [2]$$

où les $\beta_0^j, \beta_1^j, \beta_2^j, \beta_3^j$ sont les coefficients du modèle ou paramètres qui représentent les trajectoires de la variable dépendante Y au cours du temps T . Ces coefficients permettent de déterminer la forme de la trajectoire. L'indice j permet de désigner j sous-groupes de trajectoires, dont chacun est associé à une origine et une pente propre (NAGIN 2005).

Le modèle GBTM permet d'intégrer des facteurs de risques ou covariables afin d'estimer leurs effets sur les trajectoires. L'estimation des paramètres du modèle se fait par la méthode du maximum de vraisemblance (JONES et al 2001). La procédure utilise l'algorithme quasi-Newton (DENIS et al 1979) (DENIS et al 1981) pour trouver les minimas et maximas de la fonction. Lorsqu'on ajoute les covariables X_{Lt} à l'équation (1.2.5.1.2), elle devient :

$$Y_t^* = \beta_0^j + \beta_1^j t + \beta_2^j t^2 + \beta_3^j t^3 + \alpha_1^j X_{1t} + \alpha_2^j X_{2t} + \dots \alpha_L^j X_{Lt} + \epsilon_t \quad [3]$$

Dans l'équation (3) définie comme une fonction polynomiale du temps t (NAGIN 2005), β_0 représente l'intercept ou l'ordonnée à l'origine et correspond à la valeur de la variable dépendante quand celle-ci est égale à zéro ; On observe une tendance linéaire de la trajectoire lorsque le coefficient estimé du composant du temps est linéaire. L'augmentation ou la diminution de la variable dépendante pour chaque augmentation d'unité de la variable indépendante, c'est-à-dire passage du temps t_1 à t_2 est représenté par β_1 ; Pour une tendance quadratique, le coefficient estimé représente le coefficient du temps quadratique. Dans ce type de tendance, β_2 représente l'augmentation ou la diminution de la variable dépendante pour chaque augmentation d'unité du carré de la variable indépendante ; Le coefficient estimé pour la tendance cubique représente le coefficient du temps cubique. Pour ce dernier, β_3 représente l'augmentation ou la diminution de la variable dépendante pour chaque augmentation d'unité cubique de la variable indépendante (SY 2017).

La méthode GBTM est une méthode largement utilisée et applicable dans les études longitudinales. Mais, malgré sa large applicable, l'échelle de notation de l'instrument utilisé pour mesurer la variable dépendante décide la distribution de probabilité spécifique utilisée pour estimer les paramètres. La méthode utilise une stratégie de modélisation multinomiale pour approximer la distribution de la population avec différentes distribution (KWON 2014). Différentes distributions sont alors possibles pour la variable dépendante. Lorsque les données sont des fréquences, elles nécessitent l'utilisation d'une distribution de Poisson ; Le modèle logit est utilisé pour les variables dépendants dichotomiques et le modèle normal censuré pour les variables continues. Dans la littérature de la méthode GBTM, ces trois modèles sont appelés respectivement ZIP, LOGIT et CNORM (JONES et al 2001).

4.2. Modélisation linéaire hiérarchique

Les données longitudinales de nature hiérarchisée requièrent l'utilisation des modèles linéaires hiérarchiques qui permettent d'analyser facilement ces données à deux niveaux hiérarchiques. L'expression modèles linéaires hiérarchiques recouvre deux caractéristiques déterminantes (STEPHEN 1993). Premièrement, les données appropriées pour de tels modèles sont structurées de façon hiérarchique. Deuxièmement, les paramètres de ces modèles peuvent être considérés comme ayant une structure linéaire hiérarchique (STEPHEN 1993). La modélisation linéaire hiérarchique est une forme de modélisation de trajectoire permettant de décrire le changement au fil du temps d'un état (RAUDENBUSH et BRYK 2002). Aussi appelé modèle multiniveaux (MASON et al 1983) (GOLDSTEIN 1991), le modèle linéaire hiérarchique a pour principe de chercher à modéliser une variable d'intérêt continue y en fonction de variables descriptives x (caractéristiques observées au niveau individuel). Comme nous l'avons vu précédemment, les données recueillies pour ce type d'analyse ont une structure hiérarchique à deux niveaux. Ainsi, il existe des relations entre les variables de ces deux niveaux (niveau 1 et niveau 2). Selon la nature des relations, il existe trois principaux types de modèles. Le premier modèle s'intéresse aux relations existantes entre les variables de niveau 1 mais souhaite contrôler la variance existante au niveau 2 et gérer les problèmes de non indépendance des observations (ULAGA 2003). La structure du modèle linéaire hiérarchique est la suivante pour chaque niveau.

Au niveau 1 :

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij}, j = 1, \dots, J, i = 1, \dots, n_j \quad [4]$$

Au niveau 2 :

$$\beta_{0j} = \gamma_{00} + \mu_{0j}, j = 1, \dots, J \quad [5]$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}, j = 1, \dots, J \quad [6]$$

où Y_{ij} illustre un état. Il peut par exemple désigner le score ADL d'un individu i dans le groupe j . β_{0j} représente l'intersection estimée de l'état de l'individu pour le groupe j , β_{1j} représente le taux de variation ou la pente estimée de pour le groupe j . X_{ij} correspond à la variable explicative x pour l'individu i dans le groupe j (par exemple : age, sexe, etc) ; ϵ_{ij} représente une erreur aléatoire de l'individu i dans le groupe j .

Après avoir étudiée la façon dont les groupes j se différencient selon leurs origines β_0 et β_1 , on peut s'intéresser aux potentielles relations existantes entre les variables de niveau 1 mais aussi de prendre en compte l'impact des variables de niveau 2. Pour cela, certaines variables du niveau 2 caractérisant les groupes peuvent être analysées. Cette hypothèse nous permet d'écrire le deuxième modèle du modèle linéaire hiérarchique comme suit.

Au niveau 1 :

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij}, j = 1, \dots, J, i = 1, \dots, n_j \quad [7]$$

Au niveau 2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}, j = 1, \dots, J \quad [8]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + \mu_{1j}, j = 1, \dots, J \quad [9]$$

L'équation du niveau 1 (équation 7) demeure le même que celui du premier modèle. Cependant, pour le niveau 2, on constate l'ajout de la variable W_j qui représente la variable mesurée au niveau 2 pouvant expliquer les variations entre les modèles des différents groupes. W_j permet de prédire la façon incrémentale Y_{ij} une fois X_{ij} pris en compte. μ_{0j} et μ_{1j} représentent les termes d'erreurs pour le groupe j . Les pentes et les intersections des régressions de niveau 1 varient en fonction de certaines caractéristiques des individus.

Modèle de mélange de croissance généralisé à effet aléatoire

Le modèle mixte selon Cnaan (CNAAN et al 1997) est similaire à la régression multiple ordinaire et permet d'évaluer la corrélation entre les observations. Les modèles mixtes sont particulièrement utiles dans les situations où des mesures répétées sont effectuées sur les mêmes individus dans des études longitudinales. Ils permettent donc de prendre en compte l'aspect longitudinal des données.

Le modèle tient compte de la variabilité liée aux individus et de l'autocorrélation entre les variables. La partie fixe du modèle à effets mixtes est identique pour chaque individu et représente l'effet population. La partie aléatoire est traduite par la variabilité liée à chaque sujet (GAUDART et al 2010). Les trajectoires de données longitudinales sont modélisées à l'aide du modèle à effets mixtes généralisé suivant :

$$Y_{ti} = \beta_{0i} + \beta_{1i}X_{ti} + \beta_{2i}X_{ti}^2 + \dots + \beta_{pi}X_{ti}^p + \epsilon_{it} \quad [10]$$

où l'équation 10 est une équation du modèle polynomiale de puissance p ; Y_{ti} représente la variable d'intérêt du patient i pour une trajectoire j ($i = 1, \dots, n_j$, $j = 1, \dots, J$) à un temps donné t ; $X_{t1}, X_{t2}^2, \dots, X_{ti}^p$ représentent la variable indépendante ou variables explicatives entrées dans un terme linéaire, carré ou cubique; $\beta_{0i}, \beta_{1i}, \dots, \beta_{pi}$ représentent respectivement les coefficients des termes intercept, linéaire, quadratique; ϵ_{it} illustre pour l'individu i l'effet aléatoire spécifique donné par l'état (ADL par exemple) et le groupe de trajectoire attribué j . Les trajectoires peuvent ainsi être linéaires X_{t1} , quadratiques X_{t2}^2 ou cubiques X_{t3}^3 en fonction du nombre de mesures.

Des méthodes de classification non supervisées dont k-means et classification ascendante hiérarchique peuvent également être employées pour l'analyse de données longitudinales. Ces deux méthodes sont expliquées dans les sections ci-dessous.

4.3. Classification par la technique KmL utilisant l'algorithme k-means

Dans les études longitudinales, les mêmes variables sont fréquemment mesurées de façon répétée dans le temps et chaque suite de variable reflète l'évolution d'un phénomène (GENOLINI et al 2013). Pour effectuer l'analyse évolutive de ces données, plusieurs modèles (NAGIN et ODGERS 2010) (PICKLES et CROUDACE 2010) ont été développées. Dans ces analyses, on suppose que la trajectoire de chaque individu peut être reproduite par un mélange d'un petit nombre de trajectoires typiques avec des poids donnés à chacune de ces trajectoires d'un individu à l'autre. Ce poids définit la probabilité à posteriori que l'individu soit issu du groupe en question (CANOUIL). Parmi cet ensemble de méthodes, il existe la procédure KmL implémentée par Genolini et ses collaborateurs (GENOLINI et FALISSARD 2010) (GENOLINI et FALISSARD 2011) qui utilise l'algorithme de partitionnement k-means (MACQUEEN et al 1967). K-means est un algorithme de partitionnement qui appartient à la classe EM (Expectation-Maximization) (CELEUX et GOVAERT 1992). Genolini et Falissard dans leur travaux (GENOLINI et al 2013) expliquent les algorithmes EM comme suit.

Fonctionnement de l'algorithme k-means

L'objectif de l'algorithme k-means dans l'analyse de trajectoires est de déterminer la partition des individus et les trajectoires de sorte à minimiser une fonction de perte c'est à dire la fonction qui est minimisée dans la procédure d'ajustement d'un modèle. L'algorithme k-means est une méthode de classification considérée comme une approche alternative aux modèles de mélange (MACQUEEN et al 1967).

Initialement, le partitionnement assigne chaque observation à un groupe. Ensuite, le regroupement optimal est atteint en alternant deux phases : la phase d'attente et la phase de maximisation. Pendant la

phase d'attente, les centres de gravité de chaque groupe sont calculés. Ensuite, la phase de maximisation consiste à assigner chaque observation à son "groupe le plus proche". L'alternance des deux phases est répétée jusqu'à ce qu'aucun autre changement ne se produise dans les groupes. Pour illustrer cela, on considère un ensemble N de n individus. Pour chaque individu, une variable d'intérêt Y est mesurée à t périodes ou temps différents. La valeur de Y pour l'individu i au temps k est notée y_{ik} . Pour l'individu i , la séquence y_{ik} est appelée trajectoire ; elle est notée $y_i = (y_{i1}, y_{i2}, \dots, y_{it})$. L'objectif du regroupement est de partitionner N en sous-groupes homogènes g (GENOLINI et al 2013). Les moyennes k peuvent être calculées sur plusieurs distances. Pour l'implémenter sous un logiciel, Genolini et Falissard ont développé le package "KmL" sous le logiciel R (GENOLINI et FALISSARD 2010) (GENOLINI et FALISSARD 2011). Le package peut utiliser soit la distance Euclidienne ou la distance de Manhattan (KAUFMAN et ROUSSEEUW 2009).

L'algorithme du k-means est une méthode de classification qui nécessite que les individus aient des mesures effectuées aux mêmes temps t . La méthode standard ne permet pas de tenir compte de covariables pouvant influencer les trajectoires individuelles. Or, l'omission de certaines covariables dans la classification peut entraîner une mauvaise prise en compte de l'hétérogénéité inter-individuelle, et faire apparaître des groupes qui ne sont pas forcément reliés au phénomène d'intérêt l'autonomie fonctionnelle dan notre cas (CANOUIL).

Estimation des paramètres des trajectoires

L'estimation des paramètres des trajectoires est faite par la méthode du maximum de vraisemblance. On suppose que les mesures de l'individu i sont indépendantes les unes des autres conditionnellement au fait d'appartenir au groupe j .

Choix du nombre optimal de groupes

Un des problèmes rencontrés dans l'algorithme k-means est le choix du nombre optimal. La méthode KmL utilise le critère de Calinski et Harabasz (CHC) considéré comme l'un des meilleurs (KALINSKI et HARABASZ 1974) parmi tant d'autres. Les propriétés du critères ont été détaillées par les auteurs (MILLIGAN et COOPER 1985) (SHIM et al 2005). Basée sur le critère de Calinski et Harabasz, la procédure consiste à minimiser les variances à l'intérieur des groupes et de maximiser la différence entre les groupes.

Critère de Calinski Harabasz : le critère se base sur la division de la variance intergroupe $Trace(V_{Inter})$ par la variance intra groupe $Trace(V_{Intra})$. Il cherche à pondérer la variance intragroupe par le nombre de groupe et permet d'avoir un nombre optimal de groupe de manière objective. Lorsque les groupes sont bien distincts les uns des autres, la variance interclasse prend des valeurs plus élevées et si les groupes sont proches, la variance interclasse prend des valeurs plus faibles. La variance intragroupes quant à elle mesure la compacité des groupes. Lorsque les groupes sont compacts, la variance intragroupes prend des valeurs élevées. Elle prend des valeurs faibles lorsque les groupes sont dispersés (GENOLINI et FALISSARD 2010). Le critère a pour idée de calculer la somme des distances entre le k -ème groupe et les $(k - 1)^{mes}$ groupes. Par la suite cette somme est comparée avec la somme des distances à l'intérieur de chacun des k groupes. Le rapport entre les deux variances est noté $C(k)$ et s'écrit :

$$C(k) = \frac{\text{Trace}(V_{Intra})(n - k)}{\text{Trace}(V_{Inter})(k - 1)} \quad [11]$$

où n et k sont respectivement le nombre de groupes dans la partition étudiée. $\text{Trace}(V_{Inter})$ et $\text{Trace}(V_{Intra})$ représentent respectivement la somme des distances intergroupes et la somme des distances intragroupes. Lorsque les groupes sont compacts et bien séparés alors $C(k)$ est grand. Il est petit quand les groupes sont dispersés et mal séparés.

Critère d'inertie : Le critère d'inertie est lui aussi un indicateur qui permet une orientation vers la première partition à observer avant d'explorer des typologies aux nombres moins et/ou plus élevés de groupes.

4.4. Classification Ascendante Hiérarchique

Les progrès des méthodes de classification au cours des dernières années ont permis d'utiliser la classification ascendante hiérarchique pour déterminer des typologies de parcours (BARBARY 1996) (DE-GENNE et al 1996) (BARBARY et al 1998)). Pour cette méthode, le critère de Ward est généralement choisi pour le choix du nombre optimal de groupes. Le but étant de réduire l'hétérogénéité des données tout en conservant une part essentielle de l'information.

Méthode de Ward : L'idée du critère de Ward dans la classification hiérarchique ascendante est de minimiser l'inertie intragroupe et de maximiser l'inertie intergroupe afin d'obtenir les groupes les plus homogènes possibles. Pour le critère de Ward, les individus qui présentent des caractéristiques rares dans la population vont se retrouver éloignés des individus présentant des caractéristiques fortement représentées. Le critère de saut de Ward (WARD 1963) a été utilisé pour la méthode de classification hiérarchique ascendante.

Fonctionnement de l'algorithme de Classification Ascendante Hiérarchique

L'algorithme de Classification Ascendante Hiérarchique peut être décrit par 5 étapes principales :

1. Les groupes initiales sont les individus eux-mêmes ;
2. On calcule les distances entre les groupes ;
3. Les deux groupes les plus proches au sens du critère de Ward sont fusionnées et remplacées par une seule ;
4. Le processus reprend en 2 jusqu'à n'avoir plus qu'une seule classe, qui contient toutes les observations ;
5. A chaque étape les groupes les plus proches sont cherchées et fusionnées, et l'étape la plus importante dans l'algorithme réside dans le choix de la distance entre deux groupes.

Soit g le centre de gravité associé à n individus ; A et B deux groupes d'individus d'effectifs respectifs n_A et n_B , de centres de gravité associés respectifs g_A et g_B . La distance de Ward notée $Dist_W$ est :

$$Dist_W(A, B) = \frac{n_A n_B}{n_A + n_B} d^2(g_A, g_B) \quad [12]$$

Evaluation des critères de choix du nombre de groupes optimal des deux méthodes (k-means et CAH) par les critères AIC, BIC, ICL

En classification non supervisée, il est très courant de ne pas connaître à priori le nombre de groupes à former ou le nombre de composantes du modèle à ajuster (BAUDRY 2009). Alors, le choix du nombre de classe fait intervenir plusieurs critères tels que le critère AIC, BIC, ICL. Les critères AIC et BIC sont les plus classiques et sont tous les deux des critères de vraisemblance pénalisée. L'AIC se rapproche asymptotiquement de la vraisemblance intégrée du modèle et BIC utilise l'approximation de Laplace pour la vraisemblance intégrée du modèle. Dans un modèle de mélange gaussien, AIC a une forte tendance à surestimer le nombre de classe k tandis que BIC a un comportement plus satisfaisant mais quelquefois le BIC donne un nombre de groupes plus grand que le nombre jugé pertinent.

Un modèle de mélange gaussien GMM (Gaussian Mixture Model) est un modèle statistique qui sert à estimer de façon paramétrique des variables aléatoires en les modélisant comme une somme de plusieurs gaussiennes. Dans une étude faite Biernacki (BIERNACKI 2009), il a été trouvé que le modèle gaussien apporte une réponse rigoureuse, flexible et interprétable pour les multiples besoins de la classification : classification supervisée ou non, nature des données, choix du nombre de groupes.

Critère AIC : il est un critère simple et populaire de choix de groupes. Proposé par Akaike (AKAIKE 1974), le critère est connu par son choix d'un grand nombre de composantes dans le cadre des modèles de mélanges (BAUDRY 2009). Dans le cadre de sélection de modèle, il consiste à minimiser un critère de la forme vraisemblance pénalisée. Le critère s'écrit :

$$AIC = 2df - 2\ln L \quad [13]$$

où L est le maximum de la fonction vraisemblance du modèle et df le nombre de paramètres du modèle estimé. Avec le critère AIC, la déviance du modèle $2\log(L)$ est pénalisée par 2 fois le nombre de paramètres. Dans un cadre de sélection de modèle, le critère AIC choisit le modèle qui donne le plus faible AIC (AKAIKE 1974).

Critère BIC : inspiré du critère d'information Akaike, le critère BIC critère d'information bayésien pénalise plus fortement le nombre de paramètre du modèle que le critère AIC. Le critère s'écrit :

$$BIC = -2\ln(L) + \ln(n)df \quad [14]$$

où L est le maximum de la fonction vraisemblance du modèle, n le nombre d'observation dans l'échantillon et df le nombre de paramètres du modèle estimé. De façon standard k prend la valeur de 2 par défaut comme observé dans les formules de l'AIC et BIC. La valeur de k peut varier en fonction du besoin.

Critère ICL : dans le but d'améliorer le BIC, Biernacki et ses collaborateurs imitent l'approche BIC puisque ce dernier présente des difficultés dans le choix du nombre de groupes pertinentes (BIERNACKI et al 2000). Idéalement, ils remplacent la vraisemblance L du modèle par la vraisemblance complète. Le critère ICL repose sur un terme dit d'entropie. Contrairement au critère BIC, le critère ICL n'est pas consistant. ICL peut alors être expliqué comme une approximation d'un critère analogue à BIC. Toutefois, les deux critères se comportent de la même façon pour les modèles fournissant des groupes bien séparées. Plusieurs études ont été faites pour comparer ces critères entre eux dans des contextes différents. L'une d'elles a été faite par Yang en 2005 qui compare l'ACI et le BIC dans le contexte de la régression (YANG 2005). L'auteur a montré qu'en particulier, l'AIC est asymptotiquement optimal lorsque l'on souhaite sélectionner le modèle avec l'erreur quadratique moyenne (si l'on fait l'hypothèse que le modèle générant les données n'est pas parmi les candidats, ce qui est en fait presque toujours le cas en pratique) ; ce n'est pas le cas du BIC (YANG 2005). L'auteur montre aussi que la vitesse de convergence de l'AIC vers l'optimum est, dans un certain sens, le meilleur possible. Dans une étude menée par Kuha (KUHA 2004), il a été montré que malgré le fondement différent du critère AIC et BIC, on observe certaines similitudes entre et les deux informations. Contrairement à l'ICL, le BIC pénalise la complexité des modèles(BIERNACKI et al 2009).

Principe de fonctionnement de l'AIC et du BIC

Les deux critères fonctionnent avec l'algorithme du maximum de vraisemblance. Selon Olivier et collaborateur (OLIVIER et al 1997), le maximum de vraisemblance a pour principe de conduire à une surparamétrisation du modèle pour obtenir un bon résultat. Idéalement, le critère de pénalisation de type log-vraisemblance le plus utilisé est l'AIC. Cependant, ce critère reste moins satisfaisant quelquefois car il a des capacités de surparamétrisation. Le BIC quant à lui assure une meilleure estimation en pénalisant justement le surdimensionnement du modèle. Placé dans un contexte Bayesien, le principe du BIC est présenté comme suit : soit $X = (X_1, X_2, \dots, X_i)$, un n-échantillon de variables indépendants de densité inconnue notée f (SCRUCCA et al 2016). L'objectif est d'estimer la densité f en se donnant une collection finie de modèle allant de 1 à m notée (M_1, M_2, \dots, M_m) . Ces modèles sont implémentés dans le packages mclust du logiciel R avec les dénominations suivantes : "E" pour modèle à variance égale, "V" pour modèle à variance variable, "EII" modèle sphérique à variance égale, ... (SCRUCCA et al 2016). Ces modèles sont représentés sur les figures Un modèle M_i a pour densité g_{M_i} de paramètre α_i et K_i . L'objectif du BIC est de choisir un modèle parmi cette collection de modèles. Dans la démarche, α_i et M_i sont vus comme des variables aléatoires et sont munies d'une distribution a priori notée $P(M_i)$ pour le modèle M_i , $P(\alpha_i|M_i)$ pour α_i . Ainsi, le BIC cherche à sélectionner le modèle M_i qui maximise la probabilité à postériori $P(M_i|X)$ notée :

$$M_{BIC} = \arg \max P(M_i|X)$$

[15]

4.5. Modélisation par la méthode de régression logistique multinomiale

La régression logistique multinomiale a été utilisée pour analyser la dépendance entre la variable réponse et les variables explicatives afin d'identifier les facteurs prédictifs. L'inclusion des facteurs de

risque dans le modèle a été conçue pour déterminer si les facteurs de risque affectent les trajectoires. Cela permet de savoir si le facteur de risque diminue ou augmente la probabilité d'appartenir à un groupe de trajectoire. Dans la régression logistique multinomiale, le groupe de trajectoire ayant le plus grand nombre d'individu est souvent choisi comme groupe référent pour la variable réponse. Selon, pour faciliter l'interprétation dans un modèle de régression logistique, on choisit souvent comme catégorie de référence la catégorie associée au plus faible risque de présenter l'évènement défavorable. Lorsque la variable explicative est qualitative binaire, la modalité de référence correspond à la modalité codée 0 dans le modèle.

Équation du modèle de régression logistique multinomiale

Pour obtenir les probabilités d'appartenance d'un individu à une modalité ou catégorie k , on peut écrire $K - 1$ équations logits. La stratégie de modélisation consiste à modéliser $K - 1$ rapports de probabilités c'est-à-dire qu'on prend une modalité comme référence et on exprime $K - 1$ logits par rapport à cette référence. On a $J + 1$ paramètres à estimer pour chaque modalité k . Sachant que la modalité de référence est la modalité K , l'équation logit d'appartenance à la modalité k de l'individu x_i , $1 \leq i \leq n$ où n est le nombre d'individus est donc :

$$\text{logit}_k(x_i) = \ln \left[\frac{\pi_k(x_i)}{1-\pi_k(x_i)} \right] = \beta_{0,k} + \beta_{1,k}X_1(x_i) + \dots + \beta_{j,k}X_j(x_i) + \dots + \beta_{J,k}X_J(x_i) + \epsilon_k \quad [16]$$

L'équation (11) représente l'ensemble des logits du modèle multinomial à K modalités. Les x_i correspondent aux individus, $\pi_k(x_i)$ symbolise la probabilité à posteriori d'un individu x_i ou probabilité associée à la modalité d'intérêt ; \ln désigne la fonction logarithme népérien ; $\beta_k = (\beta_{0,k}, \beta_{1,k}, \dots, \beta_{J,k})$ correspond au vecteur des paramètres que l'on souhaite estimer à partir des données. Le vecteur $\mathbf{X}_J(x_i) = (1, X_1(x_i) + X_2(x_i) + \dots + X_J(x_i))$ symbolise les variables explicatives. Les x_i , $1 \leq i \leq n$ où n sont les individus de notre étude avec $n=221$. Il y'a donc $(K - 1)(J + 1)$ paramètres à estimer. Les ϵ_k sont des variables aléatoires gaussiennes centrées. Ces termes d'erreur représentent la déviation entre le modèle prédict et la réalité. La première composante ($X_0(x_i) = 1$) désigne la constante.

Estimation des paramètres

L'estimation des paramètres β_k nécessite de maximiser la vraisemblance. La fonction logarithmique étant croissante, il est équivalent de maximiser la log vraisemblance. Comme on a besoin de calculer les valeurs où la dérivée de la vraisemblance s'annule, il est souvent plus facile de calculer les dérivées de la log vraisemblance (RAKOTOMALALA 2011). De plus, la recherche du maximum de vraisemblance nécessite généralement de calculer la dérivée de la vraisemblance, et cela est beaucoup plus simple avec la log-vraisemblance. La fonction de log-vraisemblance à maximiser s'écrit alors comme la somme des logarithmes des vraisemblances ou probabilités non conditionnelles (BONNET 2004). La vraisemblance L de l'individu x_i est définie par :

$$L = \prod_{i=1}^n [\pi_1(x_i)]^{y_1(x_i)} \dots [\pi_k(x_i)]^{y_k(x_i)} \quad [17]$$

où $\pi_1(x_i) = P[y_k(x_i) = y_k(x_i)/X(x_i)] = (1, X_1(x_i), X_2(x_i), \dots, X_J(x_i)); y_k(x_i) = 1$ si et seulement

si $y(x_i) = y_k$.

La log-vraisemblance LL de l'individu x_i est définie par :

$$LL = \sum_{i=1}^n y_1(x_i) \ln(\pi_1(x_i)) + \dots + y_k(x_i) \ln(\pi_k(x_i)) \quad [18]$$

avec $\beta_k = (\beta_{0,k}, \beta_{1,k}, \dots, \beta_{J,k})$. En dérivant la log-vraisemblance par rapport aux éléments du vecteur β_k , pour chaque valeur de k , on obtient un vecteur de dimension $(K - 1)(J + 1)$ de dérivées appelé vecteur du gradient et noté \mathbf{G} . L'obtention de ce vecteur se fait par annulation de toutes les dérivées partielles de la log-vraisemblance LL (équation 18).

Calcul du vecteur gradient

Pour calculer G , on va commencer par estimer les $(K - 1)(J + 1)$ paramètres en annulant $(K - 1)(J + 1)$ dérivées partielles. Cela revient à résoudre l'équation matricielle $\mathbf{G} = \mathbf{0}$ où G est une matrice de dimension $(K - 1)(J + 1) + 1$ exprimée comme suit.

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_{K-1} \end{pmatrix} \quad [19]$$

avec G_k de dimension $(J + 1)1$ égal à :

$$\mathbf{G}_k = \begin{pmatrix} g_{0,k} \\ \vdots \\ g_{j,k} \\ \vdots \\ g_{J,k} \end{pmatrix}$$

Annuler \mathbf{G}_k revient à résoudre l'équation matricielle :

$$\mathbf{G}_k = \begin{pmatrix} g_{0,k} \\ \vdots \\ \vdots \\ g_{j,k} \\ \vdots \\ \vdots \\ g_{J,k} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}; 1 \leq k \leq K - 1 \quad [20]$$

où $g_{j,k} = \frac{\partial LL}{\partial \beta_{j,k}}$, $1 \leq j \leq J, 1 \leq k \leq K - 1$

Avant de résoudre cette équation, nous allons calculer la dérivée de $\pi_k(x_i)$ par rapport à $\beta_{j,k}$.

Calcul de la dérivée de la log-vraisemblance

Au vu de l'expression de la log-vraisemblance (équation 18), le calcul de la dérivée de la log-vraisemblance nécessite tout d'abord de calculer la dérivée de $\pi_k(x_i)$. Pour cela, considérons l'équation (16) et posons Y comme variable d'intérêt. Considérons la moyenne conditionnelle de la variable Y étant donné $X = x$ qu'on notera $E(Y = y_i/X = x_i)$. La moyenne conditionnelle correspond à la valeur espérée de la variable d'intérêt Y pour une valeur donnée de la variable prédictive X. Le terme d'erreur ϵ_k dans l'équation (16) est normalement distribuée de moyenne nulle et de variance constante (LAROSE et VALLAUD 2012). Ainsi l'équation (16) devient :

$$E(Y = y_i/X = x_i)) = \ln \left[\frac{\pi_k(x_i)}{1-\pi_k(x_i)} \right] = \beta_{0,k} + \beta_{1,k}X_1(x_i) + \dots + \beta_{j,k}X_j(x_i) + \dots + \beta_{J,k}X_J(x_i) \quad [21]$$

Pour simplifier les expressions, notons $E(Y = y_i/X = x_i)$ la moyenne conditionnelle comme égale à $\pi_k(x_i)$ donc $E(Y = y_i/X = x_i) = \pi_k(x_i)$, $\beta_{j,k} = a$ la somme des paramètres sans le terme d'erreur. Avec ces expressions, la moyenne conditionnelle de la régression logistique prend une forme autre que linéaire une forme en S appelée fonction sigmoïde (LAROSE et VALLAUD 2012). De la formule (21), on déduit facilement que la fonction sigmoïde est définie par :

$$\pi_k(x_i) = \frac{1}{1 + e^{-(\beta_{0,k} + \beta_{1,k}X_1(x_i) + \dots + \beta_{J,k}X_J(x_i))}} \quad [22]$$

Pour continuer les calculs, on a besoin du lemme suivant :

Lemme 1 : Soit la fonction $U(x) = \frac{1}{1+e^{-a(x)}}$ où $a(x)$ est une fonction dérivable, sa dérivée est égale à :

$$U'(x) = a'(x)U(x) [1 - U(x)]$$

Démonstration :

Posons

$$U(x) = \frac{1}{1 + e^{-a(x)}}$$

$$U'(x) = \frac{u'(x)}{u^2(x)}; u(x) = 1 + e^{-a(x)}$$

$$u'(x) = -a'(x)e^{-a(x)}; u^2(x) = (1 + e^{-a(x)})^2$$

$$\frac{u'(x)}{u^2(x)} = \frac{a'(x)e^{-a(x)}}{(1 + e^{-a(x)})^2} = a'(x) \frac{1}{1 + e^{-a(x)}} \frac{e^{-a(x)}}{1 + e^{-a(x)}} = a'(x)U(x) [1 - U(x)]$$

$$U'(x) = a'(x)U(x) [1 - U(x)]$$

Du lemme 1, on déduit que :

$$\frac{\partial \pi_k(x_i)}{\partial \beta_{0,k}} = \pi_k(x_i) [1 - \pi_k(x_i)]; j = 0$$

$$\frac{\partial \pi_k(x_i)}{\partial \beta_{j,k}} = X_j \pi_k(x_i) [1 - \pi_k(x_i)]; 1 \leq j \leq J \quad [23]$$

où $x = (1, x_1, x_2, \dots, x_J)$

Pour déduire l'expression de la dérivée de la log vraisemblance, on remplace $\frac{\partial \pi_k(x_i)}{\partial \beta_{j,k}}$ par son expression dans l'équation (18). Rappelons une autre expression de l'équation (18) qui est :

$$LL = \sum_{i=1}^n y_1(x_i) \ln(\pi_1(x_i)) + \dots + y_k(x_i) \ln(\pi_k(x_i)) \quad [24]$$

Ainsi :

$$g_{j,k} = \frac{\partial LL}{\partial \beta_{j,k}} = \sum_{i=1}^n [y_k(x_i) [\ln(\pi_k(x_i))] + (1 - y_k(x_i)) \ln [(1 - \pi_k(x_i))]]'$$

$$\begin{aligned}
&= \sum_{i=1}^n \left[y_k(x_i) \left[\ln(\pi_k(x_i)) \right] \right]' + \left[(1 - y_k(x_i)) \ln \left[(1 - \pi_k(x_i)) \right] \right]' \\
&= \sum_{i=1}^n y_k(x_i) \left[\ln(\pi_k(x_i)) \right]' + \ln(\pi_k(x_i)) \left[y_k(x_i) \right]' + (1 - y_k(x_i)) \ln \left[(1 - \pi_k(x_i)) \right]' + \\
&\quad \ln \left[(1 - \pi_k(x_i)) \right] \left[(1 - y_k(x_i)) \right]' \\
&= \sum_{i=1}^n y_k(x_i) \frac{\left[\pi_k(x_i) \right]'}{\pi_k(x_i)} + (1 - y_k(x_i)) \frac{\left[1 - \pi_k(x_i) \right]'}{(1 - y_k(x_i))}
\end{aligned}$$

En remplaçant l'expression de la dérivée de $\pi_k(x_i)$, on obtient :

$$g_{j,k} = \frac{\partial LL}{\partial \beta_{j,k}} = \sum_{i=1}^n y_k(x_i) \left[\frac{\pi_k(x_i)(1-\pi_k(x_i))}{\pi_k(x_i)} \right] X_j(x_i) + (1 - y_k(x_i)) \left[\frac{-\pi_k(x_i)(1-\pi_k(x_i))}{1-\pi_k(x_i)} \right] X_j(x_i);$$

Après simplification, on trouve :

$$\begin{aligned}
g_{0,k} &= \frac{\partial LL}{\partial \beta_{0,k}} = \sum_{i=1}^n [y_k(x_i) - \pi_k(x_i)]; j = 0 \\
g_{j,k} &= \frac{\partial LL}{\partial \beta_{j,k}} = \sum_{i=1}^n X_j(x_i) [y_k(x_i) - \pi_k(x_i)]; 1 \leq j \leq J
\end{aligned} \tag{25}$$

Avec cette équation du gradient, on va chercher à annuler l'expression du gradient pour trouver les valeurs des estimations des paramètres $\beta_k = (\beta_{0,k}, \beta_{1,k}, \dots, \beta_{J,k})$.

Pour cela, il faut résoudre le système de $J + 1$ équations à $J + 1$ inconnues pour chaque valeur de k .

$$\mathbf{G}_k = \begin{pmatrix} g_{0,k} \\ \vdots \\ g_{j,k} \\ \vdots \\ g_{J,k} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n [y_k(x_i) - \pi_k(x_i)] \\ \vdots \\ \sum_{i=1}^n X_j(x_i) [y_k(x_i) - \pi_k(x_i)] \\ \vdots \\ \sum_{i=1}^n X_J(x_i) [y_k(x_i) - \pi_k(x_i)] \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad 1 \leq k \leq K - 1 \tag{26}$$

Les $(K - 1)(J + 1)$ paramètres à estimer sont les solutions des $K - 1$ systèmes de $J + 1$ équations à $J + 1$ inconnues pour chaque valeur de k . Dans ce système non linéaire, il y a autant d'équations que de paramètres à estimer. En raison de sa non linéarité, le système n'admet pas de solution analytique. Sa résolution nécessite des méthodes numériques telles que l'algorithme de Newton-Raphson. L'algorithme de Newton-Raphson recherche pas à pas les valeurs paramètres ($\hat{\beta}_k$).

Le déroulement de l'algorithme se fait de la façon suivante. On part de valeurs initiales des $(J + 1)$ paramètres du modèle (par exemple, $\beta_k = 0$ et on $k = 1, \dots, K$). Puis on remplace chaque équation de par son approximation linéaire autour de ces valeurs initiales. On résout le système ainsi formé et on obtient un premier jeu de valeurs des paramètres. On répète l'opération en remplaçant chaque équation de par son approximation linéaire autour de ce premier jeu de paramètres. On résout le système ainsi formé, et ainsi de suite jusqu'à ce que les valeurs des paramètres ainsi déterminées ne changent (quasiment) pas lorsqu'on itère l'opération. On arrête alors la recherche des valeurs des β_k et les dernières obtenues sont les valeurs estimées des paramètres (AFSA 2016).

Soit β_k^0 un paramètre initial du paramètre β_k . On part de l'étape l à l'étape $l + 1$ par l'expression suivante :

$$\beta_k^{l+1} = \beta_k^l - \left(\frac{\partial^2 LL}{\partial \beta_k \partial \beta_k} \right) \mathbf{G}_k \quad [27]$$

où $\frac{\partial^2 LL}{\partial \beta_k \partial \beta_k}$ est la matrice hessienne $H_{k,k}$ de dimension $(J + 1)(J + 1)$, $1 \leq k \leq K - 1$. La matrice hessienne s'obtient en dérivant chacune des dérivées partielles. Elle est définie par :

$$\mathbf{H}_k = \begin{pmatrix} \frac{\partial^2 LL}{\partial \beta_{0,k} \partial \beta_{0,k}} & \cdots & \frac{\partial^2 LL}{\partial \beta_{0,k} \partial \beta_{J,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 LL}{\partial \beta_{j,k} \partial \beta_{0,k}} & \cdots & \frac{\partial^2 LL}{\partial \beta_{j,k} \partial \beta_{J,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 LL}{\partial \beta_{J,k} \partial \beta_{0,k}} & \cdots & \frac{\partial^2 LL}{\partial \beta_{J,k} \partial \beta_{J,k}} \end{pmatrix} = \begin{pmatrix} H_{0,0}^k & \cdots & H_{0,J}^k \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ H_{J,0}^k & \cdots & H_{J,J}^k \end{pmatrix} \quad [28]$$

On sait que :

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = \frac{\partial}{\partial \beta_{j_1,k}} \left(\frac{\partial LL}{\partial \beta_{j_2,k}} \right) = \frac{\partial}{\partial \beta_{j_1,k}} g_{j_2,k} \quad 0 \leq j_1 \leq J; 0 \leq j_2 \leq J$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = \frac{\partial}{\partial \beta_{j_1,k}} \sum_{i=1}^n X_{j_2}(x_i) [y_k(x_i) - \pi_k(x_i)]$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = \frac{\partial}{\partial \beta_{j_1,k}} \sum_{i=1}^n [X_{j_2}(x_i) y_k(x_i)]' - [\pi_k(x_i) X_{j_2}(x_i)]'$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = \sum_{i=1}^n -X_{j_2}(x_i) \frac{\partial}{\partial \beta_{j_1,k}} \pi_k(x_i)$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = - \sum_{i=1}^n X_{j_2}(x_i) \pi_k(x_i) (1 - \pi_k(x_i)) X_{j_1}(x_i)$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{j_2,k}} = - \sum_{i=1}^n \pi_k(x_i) (1 - \pi_k(x_i)) X_{j_1}(x_i) X_{j_2}(x_i) \quad 1 \leq j_1 \leq J; 1 \leq j_2 \leq J$$

$$\frac{\partial^2 LL}{\partial \beta_{0,k} \partial \beta_{j_2,k}} = - \sum_{i=1}^n \pi_k(x_i) (1 - \pi_k(x_i)) X_{j_2}(x_i) \quad 1 \leq j_2 \leq J$$

$$\frac{\partial^2 LL}{\partial \beta_{j_1,k} \partial \beta_{0,k}} = - \sum_{i=1}^n \pi_k(x_i) (1 - \pi_k(x_i)) X_{j_1}(x_i) X_{j_2}(x_i) \quad 1 \leq j_1 \leq J$$

$$\frac{\partial^2 LL}{\partial \beta_{0,k} \partial \beta_{0,k}} = - \sum_{i=1}^n \pi_k(x_i) (1 - \pi_k(x_i)) \quad [29]$$

Ainsi l'hessienne dans l'équation (28) s'écrit :

$$\mathbf{H}_{k,l} = \begin{pmatrix} \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) & \dots & \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{j}}(x_i) & \dots & \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{j}}(x_i) \mathbf{X}_{\mathbf{J}}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}(x_i) & \dots & \sum_{i=1}^n \pi_{\mathbf{k}}(x_i) (1 - \pi_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}^2(x_i) \end{pmatrix} \quad [30]$$

$$\mathbf{H}_{k,l} = \begin{pmatrix} 1 & x_1 & \dots & x_1 \\ \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n \end{pmatrix}^t \begin{pmatrix} \pi_k(x_i)(1 - \pi_k(x_i)) & & & 0 \\ & \ddots & & \\ 0 & & \pi_k(x_i)(1 - \pi_k(x_i)) & \\ & & & \end{pmatrix} \begin{pmatrix} 1 & x_1 & \dots & x_1 \\ \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n \end{pmatrix}$$

[31]

$$\mathbf{H}_{k,l} = X' V X$$

Avec X une matrice contenant les données pour chaque individu. La première colonne de matrice X correspond à la constante 1.

$$\mathbf{V} = \begin{pmatrix} \pi_k(x_i)(1 - \pi_k(x_i)) & & 0 \\ & \ddots & \\ 0 & & \pi_k(x_i)(1 - \pi_k(x_i)) \end{pmatrix}$$

une matrice diagonale contenant les éléments généraux $\pi_k(x_i)(1 - \pi_k(x_i))$.

On en déduit que :

$$\frac{\partial^2 LL}{\partial \beta_k \partial \beta_k}(\widehat{\beta}_k) = \begin{pmatrix} \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) & \dots & \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{j}}(x_i) & \dots & \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{j}}(x_i) \mathbf{X}_{\mathbf{J}}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}(x_i) & \dots & \sum_{i=1}^n \widehat{\pi}_{\mathbf{k}}(x_i)(1 - \widehat{\pi}_{\mathbf{k}}(x_i)) \mathbf{X}_{\mathbf{J}}^2(x_i) \end{pmatrix}$$

[32]

où $\widehat{\pi}_k(x_i)) = \frac{e^{\widehat{\beta}_{0,k} + \widehat{\beta}_{1,k} X_1(x_i) + \dots + \widehat{\beta}_{J,k} X_J(x_i))}}{1 + e^{-(\widehat{\beta}_{0,k} + \widehat{\beta}_{j,k} X_1(x_i) + \dots + \widehat{\beta}_{J,k} X_J(x_i))}}$ et $\widehat{\beta}_{j,k}$ est l'estimateur $\beta_{j,k}$ par la méthode de Newton-Raphson.

Enfin, la matrice suivante est la matrice hessienne de la log vraisemblance LL

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{1,1} & \dots & \mathbf{H}_{1,K-1} \\ \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \mathbf{H}_{K-1,1} & \dots & \mathbf{H}_{K-1,K-1} \end{pmatrix}$$

où

$$\mathbf{H}_{k,l} = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \dots & \dots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 0 & 0 \end{pmatrix} \text{ si } k \neq l.$$

L'estimateur du maximum de vraisemblance β_k sous certaines conditions est convergent et suit asymptotiquement une loi normale de moyenne égale à la vraie valeur β_0 des paramètres et de matrice de variance covariance. La matrice de variance covariance est égale à l'inverse de la matrice hessienne évaluée au point β_0 (HURLIN 2003). L'inverse de cette matrice notée $C = (X'VX)^{-1}$ représente l'estimation de la matrice de variance covariance de β_k . Elle sera mise en contribution dans les différents tests d'hypothèses pour évaluer la significativité des coefficients. Pour une précision élevée de l'estimation d'un paramètre ou une erreur d'estimation faible, il faut que la pente de la fonction des valeurs possibles du paramètre soit élevée. Dans ce cas, le minimum de la fonction est facilement détecté. Ainsi, plus la valeur de l'hessienne est élevée et plus la précision de l'estimation des paramètres est élevée ou plus la variance de l'estimation est faible.

Test de la significativité des coefficients du modèle logit

Après avoir trouvé les équations logits et utilisé la méthode du maximum de vraisemblance pour estimer les paramètres du modèle, il convient de déterminer la significativité des coefficients obtenus.

La méthode du maximum de vraisemblance a consisté à produire les paramètres $\beta_k = (\beta_{0,k}, \beta_{1,k}, \dots, \beta_{J,k})$ de la régression logistique exprimés dans l'équation 16.

Dans cette étude, nous utilisons les trois tests liés au maximum de vraisemblance qui sont : le test de rapport de vraisemblance, le test de Wald ou le test du score. Ceux-ci permettent de tester la nullité glo-

bale de l'hypothèse nulle et la nullité de tous les coefficients. Il est possible avec ces trois tests de tester la significativité des variables explicatives (ROUVIERE 2017). La nullité est acceptée si la probabilité critique p est grande. Pour ces tests, on obtient à chaque fois une statistique asymptotiquement χ^2 , on compare donc les valeurs obtenues aux quantiles du χ^2 . Les hypothèses s'écrivent :

$$H_0 : \beta_{j,k} = 0 \text{ contre } H_1 : \beta_{j,k} \neq 0$$

où $\beta_{j,k}$ désigne la j^{me} composante du vecteur de paramètres $\boldsymbol{\beta}_k = (\beta_{0,k}, \beta_{1,k}, \dots, \beta_{J,k})$, $0 \leq j \leq J$.

Test du rapport de vraisemblance

La statistique du test de rapport de vraisemblance est basée sur la différence entre des rapports de vraisemblances entre le modèle complet et le modèle sous l'hypothèse nulle H_0 . L'estimateur ainsi obtenu sous l'hypothèse H_0 est appelé estimateur du maximum de vraisemblance contraint noté $\widehat{\beta}_{j,k}^j$. L'estimateur du maximum de vraisemblance contraint est obtenu par suppression des premiers coefficients du modèle. C'est l'estimateur obtenu avec le modèle constant ou modèle sans aucune variable explicative (HOSMER 2000). Ainsi, la statistique s'écrit :

$$LR = 2(\ln LL(\widehat{\beta}_{j,k}) - \ln LL(\widehat{\beta}_{j,k}^j)) \sim \chi^2 p \text{ ddl} \quad [33]$$

où $\widehat{\beta}_{j,k}$ représente l'estimateur du maximum de vraisemblance sans contrainte et $\widehat{\beta}_{j,k}^j$ estimateur du maximum de vraisemblance sous la contrainte avec $\beta_{j,k} = 0$. $\ln LL(\widehat{\beta}_{j,k})$ représente le logarithme de la vraisemblance du modèle avec l'ensemble des variables explicatives ; $\ln LL(\widehat{\beta}_{j,k}^j)$ désigne le logarithme de la vraisemblance du modèle réduit à la seule constante.

Si l'on note $\chi^2_{95\%}$ le quantile à 95% de la loi χ^2 , le test du rapport des maximum de vraisemblance au seuil de 5% de l'hypothèse H_0 consiste à accepter H_0 si $LR < \chi^2_{95\%}$, et à refuser H_0 si $LR > \chi^2_{95\%}$.

Test de Wald

Le test de Wald sert spécifiquement à tester la nullité d'un ou plusieurs coefficients, en particuliers de tous sauf la constante. L'idée du test de Wald est d'accepter l'hypothèse nulle si l'estimateur non contraint $\widehat{\beta}_k$ de β_k est proche de 0. Le test est basé sur la normalité asymptotique des estimateurs.

La statistique du test de Wald sous l'hypothèse $H_0 : \beta_{j,k} = 0$ contre $H_1 : \beta_{j,k} \neq 0$:

$$T_{j,k}(W) = \sum_{i=1}^n \frac{\widehat{\beta}_{j,k}^2}{V(\widehat{\beta}_{j,k})} \sim \chi^2 p \text{ ddl} \quad [34]$$

où $V(\hat{\beta}_{j,k})$ représente la variance de l'estimateur β_k . La décision de rejeter H_0 au risque α est la suivante : si l'on note $\chi^2_{95\%}$ le quantile à 95% de la loi χ^2 , le test de Wald au seuil de 5% de l'hypothèse H_0 consiste à accepter H_0 si $T_{j,k}(W) < \chi^2_{95\%}$, et à refuser H_0 si $T_{j,k}(W) > \chi^2_{95\%}$.

Mesure de la qualité du modèle

Pour juger de la significativité du modèle, nous avons utilisé le pseudo R de MacFadden (MCFADDEN et al 1973). Menard dans ses travaux (MENARD 2002) suggère que le R de MacFadden est l'un des critères le plus adapté à la régression logistique car il est le plus proche conceptuellement du coefficient de détermination de la régression multiple. Le R de MacFadden s'interprète de la façon suivante : lorsque la régression n'est pas significative ou que les variables explicatives n'apportent pas d'explication à la variable réponse, le R vaut 0 et vaut 1 lorsque la régression est parfaite. Ainsi plus la valeur du R est proche de 1, meilleur est le modèle.

$$R^2 de Mc Faden = 1 - \frac{\ln LL(\hat{\beta}_{j,k})}{\ln LL(\hat{\beta}_{j,k}^j)} \quad [35]$$

où $\ln LL(\hat{\beta}_{j,k}^j)$ représente le maximum de la fonction de log vraisemblance avec la seule constante.

Discussion

Parmi les études examinées dans ce travail, seulement 20% ($n=2$) étaient européennes développées au Royaume Uni (HYBELS et al 2010) et en Italie (JONKMAN et al 2018), une seule était faite en Norvège (ROMOREN et al 2003). A ce jour, les connaissances disponibles décrivant la trajectoire de perte d'autonomie fonctionnelle de la personne âgée sont issues principalement d'études menées aux Etats-Unis (WAKEFIELD et HOLMAN 2007), (GILL et al 2013), (HAN et al 2013), (BUURMAN et al 2016) soit 40% ($n=4$) des études examinées et en Asie (CHEN et al 2008), (KUO et al 2017), (WEI et al 2018), (SAITO et al 2018 pour un taux de 40% ($n=4$)).

La présente étude a établi un état des lieux des trajectoires de déclin fonctionnel de l'autonomie fonctionnelle. Les nombreux défis théoriques et méthodologiques de la recherche sur les parcours de vie exigent des outils statistiques souples (GHISLETTA et al 2015). Ainsi les familles de modèles statistiques pour l'analyse de données longitudinales dont l'analyse de trajectoire de classes latentes, la modélisation linéaire hiérarchique, le modèle de mélange de croissance généralisé à effet aléatoire et beaucoup d'autres méthodes sont des modèles statistiques de changement qui ont imprégné de nombreuses disciplines scientifiques (COLLINS 2006) (DUNCAN et al 2013). A ceux là s'ajoutent les méthodes de classification ascendante hiérarchique et de partitionnement k-means. L'utilisation de ces techniques de modélisations mathématiques et statistiques pour le traitement de résultats de l'autonomie fonctionnelle chez la personne âgée en fin de vie a permis de constater des différences systématiques dans l'état fonctionnel au cours du temps. Les modèles employés dans les différents travaux ont permis d'analyser des données longitudinales et de définir des paramètres liés aux processus de stabilité et de changement de la variable d'intérêt. Ils ont aidé à la modélisation des tendances temporelles, des probabilités d'appartenance à un groupe de trajectoire, et l'hétérogénéité inter et intragroupe de trajectoire.

Cet aperçu des études de trajectoires de perte d'autonomie de la personne âgée montre que les personnes âgées suivent dans le temps différents parcours selon différentes évolutions possibles. Majoritairement, le nombre minimal de temps de suivi est égal à trois points pour toutes les études. Environ 70% (n=7) des études examinées ont identifié trois tendances : (1) une trajectoire stable dans le temps qui correspond à une moyenne plus élevée des scores d'évaluation par conséquent une autonomie stable dans le temps ; (2) une trajectoire à risque de déclin fonctionnel de l'autonomie fonctionnelle qui est constituée de personnes âgées ayant une autonomie modérée à sévère. Les moyennes des scores d'évaluation diminuent avec le temps pour ce groupe ; (3) une trajectoire constante de déclin fonctionnelle de l'autonomie fonctionnelle. Ce troisième groupe de trajectoire est composé d'individus dépendants. 20% (n=2) des études ont montré cinq tendances décrites comme suit : (1) trajectoire stable dans le temps à laquelle sont assignées les personnes indépendantes, (2) trajectoire stable regroupant des individus ayant un handicap léger, (3) une trajectoire croissante composée d'individus ayant un handicap léger à modéré, (4) trajectoire de perte d'autonomie légèrement croissante (individus présentant un handicap léger à modéré), (5) trajectoire de déclin fonctionnelle de l'autonomie fonctionnelle (personnes ayant un handicap grave). Une seule étude a identifié deux tendances : (1) trajectoire stable constituée d'individus autonomes, (2) trajectoire de déclin fonctionnel composée d'individus dépendants.

Bien qu'une partie des différents scores d'évaluation de l'autonomie fonctionnelle soit demeurée stable dans le temps (trajectoire stable) pour la majorité des études, une autre partie a augmenté de façon marquée au cours du temps (trajectoire de déclin fonctionnel) et par conséquent cette dernière se dégrade dans le temps.

Ainsi, au fil du temps, certains patients sont demeurés libres de toute mobilité et de toute limitation des activités de la vie quotidienne, des activités instrumentales de la vie courante. Ces études ont montré que chez les patients présentant initialement des limitations, deux cas de figure se présentent. Premièrement, aucune amélioration graduelle n'est observée au cours du temps pour certains patients. Deuxièmement, une amélioration est observée pour d'autres pendant une période donnée (première année de suivie par exemple) suivie d'une stabilisation ou encore d'une dégradation. Pour la plupart des travaux tels que ceux de Chen et Hybels (CHEN et al 2008), (HYBELS et al 2010), les patients du groupe ayant le plus de limitations étaient les plus à susceptibles d'avoir un âge plus élevé, un score MMS initial plus faible. Cependant, l'hétérogénéité du milieu de vie et des caractéristiques des personnes âgées engendrent une difficulté dans la comparaison des facteurs prédictifs des différentes trajectoires d'autonomie fonctionnelle. Au moyen de la régression logistique, les relations étudiées entre les variables explicatives et les trajectoires ont montré que chaque étude avait ses variables prédictives spécifiques. Si par exemple l'étude de Chen (CHEN et al 2008) intègre les variables démographiques (âge, sexe, status marital, lieu d'habitation, revenus, niveau d'éducation, profession, appartenance ethnique), les comorbidités, les médicaments, les critères d'évaluation de la fonction cognitive (MMS, GDS), le status nutritionnel (MNA), l'étude de Han (LIANG et al 2010) ajoute comme variables explicatives à son modèle les caractéristiques démographiques, la fragilité physique, les indicateurs de l'état cognitif dont les MMS, les symptômes dépressifs évalués GDS, la nutrition par l'IMC, l'activité physique habituelle selon le Physical Activity Scale et des troubles chroniques auto-évalués et diagnostiqués par des médecins à savoir : hypertension, infarctus du myocarde, insuffisance cardiaque congestive, accident vasculaire cérébral, diabète sucré, arthrite, fracture de la hanche, maladie pulmonaire chronique et cancer. A peu près toutes ces études ont révélé qu'au moins un des critères suivants : le fait d'être plus âgée, d'être une femme, d'avoir moins d'années d'études, d'avoir un score MMS de base plus élevé, d'avoir une santé auto-évaluée pas-

sable/mauvaise et un soutien social perçu moindre étaient associés à l'appartenance à une trajectoire de déclin fonctionnelle de l'autonomie fonctionnelle.

Conclusion

Dans une analyse de trajectoires utilisant les méthodes vues dans notre travail, le nombre optimal de groupes reste immuable et représente selon Nagin (NAGIN et TREMBLAY 2005) une approximation statistique pratique. Pour ce faire, les individus n'appartiennent pas réellement à des groupes de trajectoires fixes ; on leur attribue plutôt une probabilité d'appartenance à un groupe. Ainsi les groupes ne doivent pas être considérés comme concrets ou réels. Le modèle mis en place devrait alors être considéré comme un outil statistique pratique, plutôt qu'un état d'être, pour résumer les trajectoires dans des régions distinctes de la distribution (NAGIN et ODGERS 2010). Des différences dans la durée des études ont été observées : il y a eu des études de suivi à court terme (6 mois) et à long terme (12 à 13.5 ans). Selon les travaux de Palese, la durée du suivi peut modifier la description de l'état fonctionnel de la personne âgée : plus l'intervalle est grand, plus il est probable que le changement se produise (PALESE et al 2016). Cependant, toutes ces études ont respecté le nombre minimal de points nécessaire pour une étude de trajectoire pour données longitudinales (DUPERE et al 2007).

Les études examinées dans notre travail montrent que les sujets se trouvant dans les groupes de trajectoire d'autonomie modérée à sévère sont ceux qui connaissent le plus de changements importants dans les scores d'évaluation entre le début et la fin de l'étude. Par conséquent cette catégorie de patients devrait être suivie de prêt pour éviter une décompensation. Ces différences de trajectoires doivent être prises en compte lors de l'élaboration de plan d'intervention visant à prévenir ou à retarder l'apparition du déclin fonctionnel. Les modèles élaborées par les études examinées ont montré un ensemble de facteurs démographiques, environnementaux, médicaux, cliniques parfois complexes qui influencent le status d'autonomie fonctionnelle de la personne âgée. Outre les facteurs découverts, il serait nécessaire de découvrir la complexité des facteurs sous-jacents au risque de déclin fonctionnel. Notre étude conclut que les trajectoires d'autonomie fonctionnelle les plus courantes chez les personnes âgées sont de trois types : stable, modéré à sévère, et sévère dans le temps. Au regard de ces résultats, les facteurs explicatifs, prédictifs ayant un impact sur les trajectoires de déclin fonctionnel d'autonomie chez la personne âgée devrait être non seulement fortement prise en compte dans la procédure de suivi mais aussi, ils devraient faire l'object d'étude beaucoup plus approfondies. Outre les patients suivant une trajectoire de déclin fonctionnel, suivis de manière plus attentive par les professionnels de santé, il serait également nécessaire de se focaliser sur les personnes âgées qui présentent un parcours d'autonomie stable afin de leurs permettre de continuer à garder cette stabilité et à améliorer leurs capacités.

Bibliographie

- MENDES, M. F., TOMÉ, L. P., BANDEIRA, M. L., BATISTA, I., BOTELHO AZEVEDO, A., MOREIRA, M. J. G., GOMES, M. C. S., *Trajectoires de vie et transition à la vieillesse : le Portugal comme une étude de cas. In XVIIIe colloque international de l'AIDELF ?Trajectoires et âges de la vie.* Association internationale des démographes de langue française, 2016.
- CHEN, C. H., WANG, C., HUANG, G. H., *Functional trajectory 6 months posthospitalization : A cohort study of older hospitalized patients in Taiwan.* Nursing research, 57(2), 93-100, 2008.
- BOUVET, C. *Quelle prise en charge des troubles psychiques des personnes âgées ?,* 2016.
- D'EPINAY, C. L., SPINI, D., *Les années fragiles : la vie au-delà de quatre-vingts ans.* Presses Université Laval., 2007.
- KATZ, S., DOWNS, T. D., CASH, H. R., GROTH, R. C., *Progress in development of the index of ADL.* The gerontologist, 20-30, 1970.
- LAWTON, M. P., BRODY, E. M., *Assessment of older people : self-maintaining and instrumental activities of daily living.* The gerontologist, 179-186, 1969.
- HÉBERT, R., CARRIER, R., BILODEAU, A., *The Functional Autonomy Measurement System (SMAF) : description and validation of an instrument for the measurement of handicaps.* Age and ageing, 17(5), 293-302, 1988.
- MAHONEY, F. I., BARTHEL, D. W., *Functional evaluation : the Barthel Index : a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill.* Maryland state medical journal, 1965.
- GRANGER, C. V., HAMILTON, B. B., KEITH, R. A., ZIELEZNY, M., SHERWIN, F. S., *Advances in functional assessment for medical rehabilitation.* Topics in geriatric rehabilitation, 1(3), 59-74, 1986.
- HAMILTON, B. B., *A uniform national data system for medical rehabilitation.* Rehabilitation outcomes : analysis and measurement, 137-147, 1987.
- SANGHA, H., LIPSON, D., FOLEY, N., SALTER, K., BHOGAL, S., POHANI, G., TEASELL, R. W., *A comparison of the Barthel Index and the Functional Independence Measure as outcome measures in stroke rehabilitation : patterns of disability scale usage in clinical trials.* International Journal of Rehabilitation Research, 28(2), 135-139, 2005.
- KATZ, S. *Studies of illness in the aged. The index of ADL : a standardized measure of biologic and psychologic function.* JaMa, 185, 94-99, 1963.
- BUURMAN, B. M., HAN, L., MURPHY, T. E., GAHBAUER, E. A., LEO-SUMMERS, L., ALLORE, H. G., GILL, T. M., *Trajectories of disability among older persons before and after a hospitalization leading to a skilled nursing facility admission.* Journal of the American Medical Directors Association, 17(3), 225-231, 2016.
- OMS, *Classification internationale des handicaps : déficiences, incapacités et désavantages. Un manuel de classification des conséquences des maladies.* Paris : OMS/CTNERHI/Les Editions INSERM, 1988.
- KATZ, S., AKPOM, C. A., *A measure of primary sociobiological functions.* International journal of health services, vol. 6, no 3, p. 493-508, 1976.
- GRAF, C., *The Lawton instrumental activities of daily living scale.* AJN The American Journal of Nursing, 108(4), 52-62, 2008.
- SUCHY, Y., KRAYBILL, M. L., FRANCHOW, E., *Instrumental activities of daily living among community-dwelling older adults : discrepancies between self-report and performance are mediated by cognitive reserve.* Journal of Clinical and Experimental Neuropsychology, 33(1), 92-100, 2011.
- GLOSSER, G., GALLO, J., DUDA, N., DE VRIES, J. J., CLARK, C. M., GROSSMAN, M., *Visual perceptual functions predict instrumental activities of daily living in patients with dementia.* Cognitive and Behavioral Neurology, 15(3), 198-206, 2002.
- JEFFERSON, A. L., PAUL, R. H., OZONOFF, A. L., COHEN, R. A., *Evaluating elements of executive functioning as predictors of instrumental activities of daily living (IADLs).* Archives of clinical Neuropsychology, 21(4), 311-320, 2006.
- CAHN-WEINER, D. A., FARIAS, S. T., JULIAN, L., HARVEY, D. J., KRAMER, J. H., REED, B. R., CHUI, H., *Cognitive and neuroimaging predictors of instrumental activities of daily living.* Journal of the International Neuropsychological Society, 13(5), 747-757, 2007.
- BURTON, C. L., STRAUSS, E., HULTSCH, D. F., HUNTER, M. A., *Cognitive functioning and everyday problem solving in older adults.* The Clinical Neuropsychologist, 20(3), 432-452, 2006.
- GRIGSBY, J., KAYE, K., BAXTER, J., SHETTERLY, S. M., HAMMAN, R. F., *Executive cognitive abilities and functional status among community ?dwelling older persons in the San Luis Valley Health and Aging Study.* Journal of the American Geriatrics Society, 46(5), 590-596, 1998.

- JEFFERSON, A. L., CAHN ?WEINER, D., BOYLE, P., PAUL, R. H., MOSER, D. J., GORDON, N., COHEN, R. A., *Cognitive predictors of functional decline in vascular dementia*. International Journal of Geriatric Psychiatry : A journal of the psychiatry of late life and allied sciences, 21(8), 752-754, 2006.
- KIOSSES, D. N., ALEXOPOULOS, G. S., *IADL functions, cognitive deficits, and severity of depression : a preliminary study*. The American Journal of Geriatric Psychiatry, 13(3), 244-249, 2005.
- MARIANI, E., MONASTERO, R., ERCOLANI, S., RINALDI, P., MANGIALASCHE, F., COSTANZI, E., ... MECOCCI, P., *Influence of comorbidity and cognitive status on instrumental activities of daily living in amnestic mild cognitive impairment : results from the ReGAl project*. International journal of geriatric psychiatry, 23(5), 523-530, 2008.
- WARD, G., JAGGER, C., HARPER, W., *A review of instrumental ADL assessments for use with elderly people*. Reviews in Clinical Gerontology, 8(1), 65-71, 1998.
- WADE, D. T., HEWER, R. L., *Functional abilities after stroke : measurement, natural history and prognosis*. Journal of Neurology, Neurosurgery and Psychiatry, 50(2), 177-182, 1987.
- GRESHAM, G. E., PHILLIPS, T. F., LABI, M. L., *ADL status in stroke : relative merits of three standard indexes*. Archives of Physical Medicine and Rehabilitation, 61(8), 355-358, 1980.
- COLLIN, C., WADE, D. T., DAVIES, S., HORNE, V., *The Barthel ADL Index : a reliability study*. International disability studies, 10(2), 61-63, 1988.
- SAITO, J., KONDO, N., SAITO, M., TAKAGI, D., TANI, Y., HASEDA, M., KONDO, K., *Exploring 2.5-Year Trajectories of Functional Decline in Older Adults by Applying a Growth Mixture Model and Frequency of Outings as a Predictor : A 2010 ?2013 JAGES Longitudinal Study*. Journal of epidemiology, JE20170230 2018.
- KONDO, N., SUZUKI, K., MINAI, J., YAMAGATA, Z., *Positive and negative effects of finance-based social capital on incident functional disability and mortality : an 8-year prospective study of elderly Japanese*. Journal of epidemiology, 22(6), 543-550, 2012.
- MCCUTCHEON, A. L., *Latent class analysis (No. 64)*. Sage, 1987.
- NAGIN, D. S., NAGIN, D., *Group-based modeling of development*. Harvard University Press, 2005.
- HYBELS, C. F., PIEPER, C. F., BLAZER, D. G., FILLENBAUM, G. G., STEFFENS, D. C., *Trajectories of mobility and IADL function in older patients diagnosed with major depression*. International journal of geriatric psychiatry, 25(1), 74-8, 2010.
- ROMOREN, T. I., BLEKESEAUNE, M., *Trajectories of disability among the oldest old*. Journal of Aging and Health, 15(3), 548-566, 2003.
- JONKMAN, N. H., DEL PANTA, V., HOEKSTRA, T., COLPO, M., VAN SCHOOR, N. M., BANDINELLI, S., MAIER, A. B., *Predicting trajectories of functional decline in 60-to 70-year-old people*. Gerontology, 64(3), 212-221, 2018.
- WEI, M., LI, J., WANG, H., *Impact of the disability trajectory on the mortality risk of older adults in China*. Archives of gerontology and geriatrics, 74, 174-183, 2018.
- LIANG, J., WANG, C. N., XU, X., HSU, H. C., LIN, H. S., LIN, Y. H., *Trajectory of functional status among older Taiwanese : Gender and age variations*. Social science and medicine, 71(6), 1208-1217, 2010.
- MUTHÉN, B., ASPAROUHOV, T., *Latent variable analysis with categorical outcomes : Multiple-group and growth modeling in Mplus*. Mplus web notes, 4(5), 1-22, 2002.
- GILL, T. M., GAHBAUER, E. A., LIN, H., HAN, L., ALLORE, H. G., *Comparisons between older men and women in the trajectory and burden of disability over the course of nearly 14 years*. Journal of the American Medical Directors Association, 14(4), 280-286, 2013.
- HAN, L., ALLORE, H., MURPHY, T., GILL, T., PEDUZZI, P., LIN, H., *Dynamics of functional aging based on latent-class trajectories of activities of daily living*. Annals of epidemiology, 23(2), 87-92, 2013.
- JUNG, T., WICKRAMA, K. A. S., *An introduction to latent class growth analysis and growth mixture modeling*. Social and personality psychology compass, 2(1), 302-317, 2008.
- COLE, V. T., APUD, J. A., WEINBERGER, D. R., DICKINSON, D., *Using latent class growth analysis to form trajectories of premorbid adjustment in schizophrenia*. Journal of abnormal psychology, 121(2), 388, 2012.
- MUTHÉN, B., *Latent variable analysis*. The Sage handbook of quantitative methodology for the social sciences, 345(368), 106-109, 2004.
- NAGIN, D. S., LAND, K. C., *Age, criminal careers, and population heterogeneity : Specification and estimation of a nonparametric, mixed Poisson model*. Criminology, 31(3), 327-362, 1993.
- NAGIN, D., TREMBLAY, R. E., *Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency*. Child development, 70(5), 1181-1196, 1999.

- JONES, B. L., NAGIN, D. S., ROEDER, K., *A SAS procedure based on mixture models for estimating developmental trajectories*. Sociological methods and research, 29(3), 374-393, 2001.
- NAGIN, D. S., *Analyzing developmental trajectories : a semiparametric, group-based approach*. Psychological methods, 4(2), 139, 1999.
- NAGIN, D. S., OGDERS, C. L., *Group-based trajectory modeling in clinical research*. Annual review of clinical psychology, 6, 109-138, 2010.
- FAN, J., *Trajectories of Emotional Symptoms Among Survivors After Severe Traumatic Brain Injury (Doctoral dissertation, University of Pittsburgh)*, 2013.
- DENNIS, J. E., MEI, H. H. W., *Two new unconstrained optimization algorithms which use function and gradient values*. Journal of Optimization Theory and Applications, 28(4), 453-482, 1979.
- DENNIS Jr, J. E., GAY, D. M., WELSCH, R. E., *Algorithm 573 : NL2SOL an adaptive nonlinear least-squares algorithm [E4]*. ACM Transactions on Mathematical Software (TOMS), 7(3), 369-383, 1981.
- SY, O., *Identification des trajectoires développementales de fréquence de la consommation d'alcool durant l'adolescence et relation entre ces trajectoires et la consommation excessive d'alcool épisodique à l'âge jeune adulte*, 2017.
- KWON, J. Y., *Group Based Trajectories Of Blood Pressure Components From Adulthood To Elderly In Chinese Workers*, 2014.
- STEPHEN, W. R., *Hierarchical linear models and experimental design*. Applied analysis of variance in behavioral science, 137, 459, 1993.
- RAUDENBUSH, S. W., BRYK, A. S., *Hierarchical linear models : Applications and data analysis methods (Vol. 1)*. Sage, 2002.
- MASON, W. M., WONG, G. Y., ENTWISLE, B., *Contextual analysis through the multilevel linear model*. Sociological methodology, 14, 72-103, 1983.
- GOLDSTEIN, H., *Nonlinear multilevel models, with an application to discrete response data*. Biometrika, 78(1), 45-51, 1991.
- ULAGA, J. L. A. W., *Les modèles linéaires hiérarchiques : 2. ?une méthode privilégiée d'analyse des données collectées par policy capturing*, 2003.
- CNAAN, A., LAIRD, N. M., SLASOR, P., *Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data*. Statistics in medicine, 16(20), 2349-2380, 1997.
- GAUDART, J., GIORGI, R., THALABARD, J. C., THIAM, D., WHEGANG, S., *Modèles linéaires à effets mixtes*, 2010.
- GENOLINI, C., PINGAULT, J. B., DRISSE, T., CÔTÉ, S., TREMBLAY, R. E., VITARO, F., FALISSARD, B., *KmL3D : a non-parametric algorithm for clustering joint trajectories*. Computer methods and programs in biomedicine, 109(1), 104-111, 2013.
- PICKLES, A., CROUDACE, T., *Latent mixture models for multivariate and longitudinal outcomes*. Statistical Methods in Medical Research, 19(3), 271-289, 2010.
- CANOUIL, M., GENOLINI, C., ECOCHARD, R., *KmLcov : k-means for longitudinal data with covariates*.
- GENOLINI, C., FALISSARD, B., *KmL : k-means for longitudinal data*. Computational Statistics, 25(2), 317-328, 2010.
- GENOLINI, C., FALISSARD, B., *KmL : A package to cluster longitudinal data*. Computer methods and programs in biomedicine, 104(3), e112-e121, 2011.
- MACQUEEN, J., *Some methods for classification and analysis of multivariate observations*. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 14, pp. 281-297), 1967.
- CELEUX, G., GOVAERT, G., *A classification EM algorithm for clustering and two stochastic versions*. Computational statistics and Data analysis, 14(3), 315-332, 1992.
- GENOLINI, C., FALISSARD, B., GENOLINI, M. C., *Package ?kml ?, 2016*.
- KAUFMAN, L., ROUSSEEUW, P. J., *Finding groups in data : an introduction to cluster analysis (Vol. 344)*. John Wiley and Sons, 2009.
- CALI ?SKI, T., HARABASZ, J., *A dendrite method for cluster analysis*. Communications in Statistics-theory and Methods, 3(1), 1-27, 1974.
- MILLIGAN, G. W., COOPER, M. C., *An examination of procedures for determining the number of clusters in a data set*. Psychometrika, 50(2), 159-179, 1985.

- SHIM, Y., CHUNG, J., CHOI, I. C., *A comparison study of cluster validity indices using a nonhierarchical clustering algorithm*. In International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06) (Vol. 1, pp. 199-204). IEEE, 2005.
- BARBARY, O., *Una aplicacion del analisis armonico cualitativo : la tipologia de trayectorias individuales*, 1996.
- DEGENNE, A., LEBEAUX, M. O., MOUNIER, L., *Typologies d ?itinéraires comme instrument d ?analyse du marché du travail*. Troisiemes journées d ?études Céreq-Cérétim-Lasmas IdL, Rennes, 23-24, 1996.
- BARBARY, O., PINZON SARMIENTO, L. M., *L'analyse harmonique qualitative et son application à la typologie des trajectoires individuelles*. Mathématiques et sciences humaines, 144, 29-54, 1998.
- WARD Jr, J. H., *Hierarchical grouping to optimize an objective function*. Journal of the American statistical association, 58(301), 236-244, 1963.
- BAUDRY, J. P., *Sélection de modèle pour la classification non supervisée*. Choix du nombre de classes (Doctoral dissertation, Université Paris Sud-Paris XI), 2009.
- BIERNACKI, C., *Pourquoi les modèles de mélange pour la classification*. Revue de MODULAD, 40, 1-22, 2009.
- AKAIKE, H., *A NEW LOOK AT THE STATISTICAL MODEL IDENTIFICATION*. In Selected Papers of Hirotugu Akaike (pp. 215-222). Springer, New York, NY, 1974.
- BIERNACKI, C., CELEUX, G., GOVAERT, G., *Assessing a mixture model for clustering with the integrated completed likelihood*. IEEE transactions on pattern analysis and machine intelligence, 22(7), 719-725, 2000.
- YANG, Y., *Can the strengths of AIC and BIC be shared ? A conflict between model indentification and regression estimation*. Biometrika, 92(4), 937-950, 2005.
- KUHA, J., *AIC and BIC : Comparisons of assumptions and performance*. Sociological methods and research, 33(2), 188-229, 2004.
- OLIVIER, C., JOUZEL, F., EL MATOUAT, A., COURTELLEMONT, P., *Un nouveau critère pour la sélection de l'ordre d'un modèle*. In 16^e Colloque sur le traitement du signal et des images, FRA, 1997. GRETSI, Groupe d ?Etudes du Traitement du Signal et des Images, 1997.
- SCRUCCA, L., FOP, M., MURPHY, T. B., RAFTERY, A. E., *mclust 5 : clustering, classification and density estimation using Gaussian finite mixture models*. The R journal, 8(1), 289, 2016.
- RAKOTOMALALA, R., *Pratique de la Régression Logistique-Régression Logistique Binaire et Polytomique*. Université Lumière Lyon, 2, 2009.
- BONNET, C., *Le modèle logit multinomial à coefficients aléatoires*. Recherche et Applications en Marketing (French Edition), 19(3), 61-72, 2004.
- LAROSE, D. T., VALLAUD, T., *Exploration de données : méthodes et modèles du data mining*. Vuibert, 2012.
- AFSA, C., *Le modèle Logit : Theorie et application (No. m2016-01)*. Institut National de la Statistique et des Etudes Economiques, 2016.
- HURLIN, C., *Économétrie des Variables Qualitatives : Polycopié de Cours*, 2003.
- ROUVIÈRE, L., *Régression logistique avec R*, 2017.
- HOSMER, D. W., *Assessing the fit of the model*. Applied logistic regression, 143-202, 2000.
- MCFADDEN, D., *Conditional logit analysis of qualitative choice behavior*, 1973.
- MENARD, S., *Applied logistic regression analysis* (Vol. 106). Sage, 2002.
- WAKEFIELD, B. J., HOLMAN, J. E., *Functional trajectories associated with hospitalization in older adults*. Western Journal of Nursing Research, 29(2), 161-177, 2007.
- KUO, H. T., LIN, K. C., LAN, C. F., LI, I. C., *Activities of daily living trajectories among institutionalised older adults : A prospective study*. Journal of clinical nursing, 26(23-24), 4756-4767, 2017.
- GHISLETTA, P., RENAUD, O., JACOT, N., COURVOISIER, D., *Linear mixed-effects and latent curve models for longitudinal life course analyses*. In *A life course perspective on health trajectories and transitions* (pp. 155-178). Springer, Cham, 2015.
- COLLINS, L. M., *Analysis of longitudinal data : The integration of theoretical model, temporal design, and statistical model*. Annu. Rev. Psychol., 57, 505-528, 2006.
- DUNCAN, T. E., DUNCAN, S. C., STRYCKER, L. A., *An introduction to latent variable growth curve modeling : Concepts, issues, and application*. Routledge, 2013.

NAGIN, D. S., TREMBLAY, R. E., *Developmental trajectory groups : Fact or a useful statistical fiction ?*. Criminology, 43(4), 873-904, 2005.

PALESE, A., MENEGAZZI, G., TULLIO, A., FUSO, M. Z., HAYTER, M., WATSON, R., *Functional decline in residents living in nursing homes : A systematic review of the literature*. Journal of the American Medical Directors Association, 17(8), 694-705, 2016.

DUPÉRÉ, V., LACOURSE, É., VITARO, F., TREMBLAY, R. E., *Méthodes d'analyse du changement fondées sur les trajectoires de développement individuelle : Modèles de régression mixtes paramétriques et non paramétriques [1]*. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 95(1), 26-57, 2007.