

No 11-522-XIF au catalogue

**La série des symposiums internationaux
de Statistique Canada - Recueil**

**Symposium 2005 : Défis
méthodologiques reliés aux
besoins futurs d'information**



2005



Statistique
Canada

Statistics
Canada

Canada

LES MODÈLES DE CROISSANCE : UNE APPROCHE INTÉRESSANTE POUR L'ÉTUDE DE L'ÉVOLUTION DE L'ÉTAT DE SANTÉ DES PERSONNES ÂGÉES

Marie-France Dubois, Réjean Hébert, et Michel Raïche¹

RÉSUMÉ

L'évaluation de l'impact des changements de services sur l'état de santé des personnes âgées fragiles nécessite des études longitudinales. Beaucoup de sujets sont toutefois perdus lors du suivi vu l'incidence élevée de décès dans cette population. Les méthodes traditionnelles d'analyses de mesures répétées deviennent inappropriées puisque la nécessité d'ignorer les sujets ayant des suivis incomplets mène à une diminution considérable de la taille d'échantillon ainsi qu'à des biais. Parce qu'ils sont valides pour tous les types de non-réponse ignorable, les modèles de croissance permettent entre autres une réduction considérable des biais. La non-réponse aléatoire est plausible ici puisque le décès est souvent relié aux mesures précédentes de l'état de santé. Un exemple concret est présenté.

MOTS CLÉS: Modèles de croissance; études longitudinales; perte d'autonomie; personnes âgées fragiles; non-réponse ignorable.

1. INTRODUCTION

Le vieillissement de la population au Québec et au Canada commande des changements majeurs dans l'organisation et la prestation des services sociosanitaires. Des études longitudinales sont alors nécessaires pour évaluer l'impact de ces changements sur l'état de santé des personnes âgées, notamment les plus fragiles. L'incidence élevée de décès est toutefois une particularité de la population âgée qui rend inappropriée l'utilisation des méthodes traditionnelles d'analyses de mesures répétées. Tout d'abord, la dilution des cohortes dans le temps entraîne une diminution de la puissance des tests statistiques. Mais principalement, les estimés sont biaisés puisque les sujets perdus sont certainement différents des survivants qui participent à tous les temps de mesure.

2. LES MODÈLES DE CROISSANCE

Les modèles de croissance constituent une approche intéressante pour l'étude de l'évolution de l'état de santé des individus. Aussi appelés modèles multi-niveaux du changement, ils tiennent compte simultanément du changement intra et inter-individuel à l'aide d'une paire de sous-modèles. Le sous-modèle de niveau 1 décrit comment les personnes changent dans le temps et le sous-modèle de niveau 2 décrit comment ces changements diffèrent entre les individus (Rogosa & Willett, 1985; Bryk & Raudenbush, 1987; Singer & Willett, 2003).

2.1 Modélisation d'une trajectoire linéaire dans le temps

¹M.-F. Dubois (Marie-France.Dubois@USherbrooke.ca) et R. Hébert (Rejean.Hebert@USherbrooke.ca), Faculté de médecine et des sciences de la santé, Université de Sherbrooke, 3001, 12^e Avenue Nord, Sherbrooke, Qc, Canada, J1H 5N4; M.-F. Dubois, R. Hébert et M. Raïche (Michel.Raiche@USherbrooke.ca), Centre de recherche sur le vieillissement du CSSS-IUGS, 1036, rue Belvédère Sud, Sherbrooke, Qc, Canada, J1H 4C4.

Commençons par supposer que le changement est une fonction linéaire du temps. Le modèle de niveau 1 s'écrit alors:

$$Y_{ij} = \pi_{0i} + \pi_{1i} T_{ij} + \varepsilon_{ij} \quad \text{où } \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2) \quad (1)$$

où Y_{ij} , T_{ij} et ε_{ij} sont respectivement l'état de santé, la mesure du temps et l'erreur de mesure pour le sujet i à la mesure j .

L'équation (1) utilise deux indices, i et j , pour identifier les individus et les temps de mesure, respectivement. Ainsi, i varie de 1 à n (la taille de l'échantillon) et j de 1 à t_i (le nombre de temps de mesure pour le sujet i). On voit bien que le nombre de temps de mesure n'a pas besoin d'être identique pour chaque sujet de l'échantillon.

Deux paramètres, appelés les paramètres de croissance, doivent être estimés dans l'équation (1) : l'ordonnée à l'origine, π_{0i} et la pente, π_{1i} . Cette dernière représente la vitesse du changement de l'individu i dans le temps. La partie stochastique de l'équation (1) représente pour sa part l'erreur de mesure pour le sujet i à la mesure j . La variance de ε_{ij} , σ_ε^2 capte la dispersion des résidus de niveau 1 autour de la trajectoire réelle de changement de chaque individu.

Les indices i associés aux paramètres de croissance indiquent que même si le modèle suppose que les trajectoires de changement des individus ont une forme algébrique commune, il ne présume pas que tous ont exactement la même trajectoire. Le sous-modèle de niveau 2 codifie la relation entre les différences inter-individuelles des trajectoires de changement et des caractéristiques fixes (qui ne bougent pas dans le temps) des individus. Ce sous-modèle est écrit en parties séparées, une partie pour chaque paramètre de croissance du niveau 1. Le modèle de niveau 2 accompagnant l'équation (1) est donc :

$$\begin{aligned} \pi_{0i} &= \gamma_{00} + \gamma_{01} X_{1i} + \gamma_{02} X_{2i} + \dots + \gamma_{0p} X_{pi} + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11} X_{1i} + \gamma_{12} X_{2i} + \dots + \gamma_{1p} X_{pi} + \zeta_{1i} \end{aligned} \quad (2)$$

où $X_{1i}, X_{2i}, \dots, X_{pi}$ sont p variables explicatives mesurées chez le sujet i

et $\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix}\right)$

Les paramètres γ_{00} , γ_{0k} , γ_{10} et γ_{1k} sont connus collectivement comme étant les effets fixes. Ils captent les différences systématiques inter-individuelles dans les trajectoires de changement selon les valeurs des variables explicatives de niveau 2. γ_{00} et γ_{10} représentent respectivement l'ordonnée à l'origine moyenne et la pente moyenne des individus dont les variables explicatives de niveau 2 sont toutes égales à 0. γ_{0k} et γ_{1k} représentent respectivement la différence moyenne de l'ordonnée à l'origine et la différence moyenne de la pente pour une différence d'une unité dans la variable explicative X_k . Enfin, les résidus, ζ_{0i} et ζ_{1i} représentent la portion des paramètres de croissance qui demeure inexplicée par les variables explicatives de niveau 2. Leurs variances et covariance sont notées σ_0^2 , σ_1^2 et σ_{01} .

Les paramètres sont estimés par la technique du maximum de vraisemblance et l'estimation peut se faire avec PROC MIXED de SAS. Les estimés sont non biaisés pour tous les types de non-réponse ignorable : non-réponse parfaitement aléatoire (NRPA), non-réponse dépendante d'une covariable (NRDC) et non-réponse aléatoire (NRA). Dans le cas de mesures répétées, on peut parler de NRA lorsque la non-réponse dépend des mesures précédentes faites après du sujet non remesuré. Ce type de non-réponse invalide les résultats de plusieurs méthodes d'analyses de mesures répétées alors qu'il est admissible dans les modèles de croissance.

2.2 Extensions du modèle linéaire

Le modèle peut s'étendre à toute trajectoire polynomiale d'ordre n . Il s'écrit alors :

$$Y_{ij} = \pi_{0i} + \sum_{l=1}^n \pi_{li} T_{ij}^l + \varepsilon_{ij} \quad (3)$$

$$\pi_{li} = \gamma_{l0} + \sum_{k=1}^p \gamma_{lk} X_{ki} + \zeta_{li}$$

Une contrainte toutefois : la construction d'un polynôme de degré n nécessite au moins $n+2$ vagues de collecte, c'est-à-dire au moins une vague de collecte de plus qu'il y a de paramètres dans le modèle de niveau 1. Et ceci n'est qu'une exigence minimale. Une plus grande précision et une meilleure puissance nécessitent plus de vagues de collecte. Des trajectoires discontinues ou continues mais non linéaires dans les paramètres (logistique, hyperbolique, inverse d'un polynôme, exponentielle, ...) peuvent également être modélisées.

2.3 La mesure du temps

Le choix d'une échelle de mesure du temps est une étape importante dans l'élaboration du modèle. Plusieurs possibilités s'offrent à nous. Premièrement, le temps peut être représenté par le numéro de la vague de collecte. Par exemple, si des mesures annuelles sont effectuées pendant trois ans, la variable prendrait les valeurs 0 (mesure initiale), 1, 2 et 3. Par contre, si les mesures ont lieu au congé puis trois et douze mois après la sortie d'un hôpital, la variable prendrait les valeurs 0, 3 et 12. Ceci illustre le fait que les vagues de collectes peuvent être espacées irrégulièrement, un avantage du modèle de croissance lorsqu'on le compare aux analyses de variance pour mesures répétées.

Le nombre de jours depuis la première évaluation est une mesure plus fine que le numéro de vague permettant non seulement aux vagues d'être espacées irrégulièrement, mais également à l'espacement des vagues de collecte de varier d'un participant à l'autre. Enfin, l'âge du sujet peut également servir de métrique du temps pour examiner l'évolution d'un phénomène avec l'âge.

3. EXEMPLE

3.1 Source des données

Les données proviennent de l'étude PRISMA (Programme de Recherche sur l'Intégration des Services de Maintien de l'Autonomie). Dans cette étude, le déclin de l'autonomie fonctionnelle est étudié en relation avec l'implantation graduelle ou non d'un réseau intégré de services aux personnes âgées (RISPA) en perte d'autonomie. Pour constituer un RISPA, des mécanismes et outils sont mis en place pour accroître l'intégration et la continuité des soins et des services aux personnes âgées.

3.2 Méthodologie de l'étude PRISMA

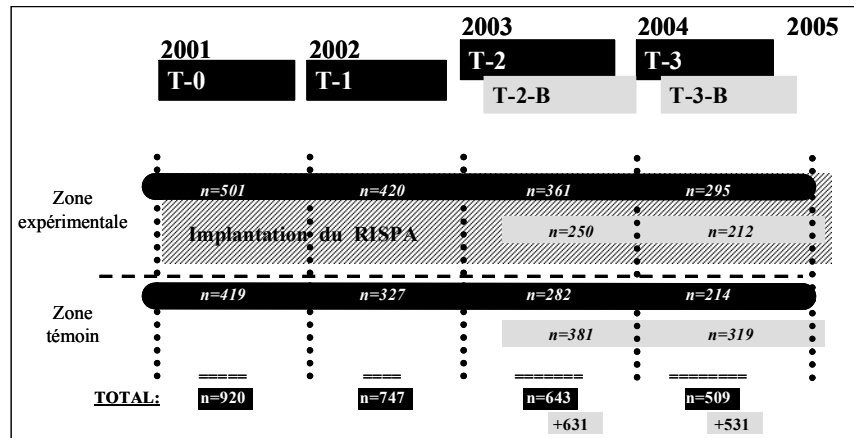
Un dispositif quasi-expérimental avec mesures répétées permet de comparer trois territoires de l'Estrie (zone expérimentale dans laquelle s'implante un RISPA) jumelés à trois territoires comparables de Chaudière-Appalaches (zone témoin) : Sherbrooke avec Lévis, Granit avec Montmagny, Coaticook avec L'Islet.

L'échantillon est constitué de personnes répondant aux critères d'inclusion suivants : (i) être âgé de 75 ans ou plus, (ii) vivre à domicile, (iii) habiter dans l'un des six territoires d'étude et (iv) être à risque de perte d'autonomie au cours de la prochaine année. Une liste des personnes répondant aux trois premiers critères d'inclusion a été tirée selon un plan aléatoire simple parmi les bénéficiaires du régime d'assurance-maladie du Québec. Cette liste constitue une excellente base de sondage puisque le régime d'assurance-maladie est universel au Québec et donc tous les québécois y sont inscrits. La vérification du quatrième critère d'inclusion nécessitait l'envoi d'un questionnaire postal à six questions (Hébert et al, 1996) à tous ces sujets potentiels. Les personnes âgées présentant trois réponses à risque ou plus au questionnaire de même que celles ne le retournant pas composaient l'échantillon et étaient approchées pour le recrutement. Le fait de ne pas retourner le questionnaire a été démontré comme constituant un facteur de risque de perte d'autonomie. Deux échantillons ont ainsi été sélectionnés, l'un en 2001 et l'autre en 2003. Pour cette deuxième phase de recrutement, l'âge minimal était maintenant de 77 ans et le fait

d'avoir été choisi en 2001 devenait un critère d'exclusion. Lors des deux vagues de recrutement, le refus de participer s'est avéré non différentiel entre les groupes expérimental et témoin.

L'autonomie fonctionnelle est mesurée au recrutement puis annuellement jusqu'en 2004 pour tous les sujets vivant encore à domicile. Ainsi, les sujets de l'échantillon 2001 sont évalués quatre fois alors que ceux de l'échantillon 2003 sont évalués deux fois. La Figure 1 schématise la méthodologie de l'étude et rapporte les tailles d'échantillon aux différents temps de mesure.

Figure 1 : Schéma de l'étude PRISMA



L'autonomie fonctionnelle est mesurée par le Système de mesure de l'autonomie fonctionnelle (SMAF). Le SMAF est un instrument d'évaluation développé selon la classification internationale des déficiences, incapacités et handicaps de l'Organisation Mondiale de la Santé (Hébert, 1982). Il évalue la capacité de la personne à accomplir 29 fonctions couvrant cinq secteurs : les activités de la vie quotidienne (7 énoncés), la mobilité (6 énoncés), les communications (3 énoncés), les fonctions mentales (5 énoncés) et les tâches domestiques (8 énoncés). Pour chaque énoncé, l'incapacité est cotée sur une échelle à cinq niveaux : 0 (autonome), -0,5 (autonome mais avec difficulté), -1 (besoin de stimulation ou de surveillance), -2 (besoin d'aide), -3 (dépendant). La cotation est réalisée par un professionnel de la santé selon des critères précis à partir de renseignements obtenus par questionnement du sujet ou d'un tiers, d'observation ou même de réalisation d'épreuves. Un score total, variant de 0 (autonomie totale) à -87 (dépendance totale) est obtenu en additionnant les cotes obtenues à chacun des énoncés. Les qualités métrologiques (fidélité, validité et sensibilité au changement) de l'instrument sont largement démontrées (Hébert et al, 1988; Desrosiers et al, 1995; Hébert et al, 2001).

3.3 Le modèle

Un modèle quadratique est choisi pour l'étude de l'évolution de l'autonomie fonctionnelle. Rappelons que nous sommes intéressés par l'impact de l'implantation des réseaux intégrés sur la trajectoire. L'implantation des différentes composantes du réseau intégré ayant lieu à des dates fixes, il est plus pertinent de regarder l'évolution des sujets par rapport à ces dates plutôt que par rapport à des numéros de vagues de mesures (0, 1, 2, 3, 4) artificiels qui ont lieu à des dates très différentes pour chacun des sujets. Ainsi, dans notre modèle, nous choisissons d'utiliser le nombre de jours depuis le 1^{er} janvier 2001 comme mesure du temps. Au deuxième niveau du modèle, nous regardons si la zone d'appartenance a un effet sur l'ordonnée à l'origine, la pente instantanée ou la courbure de notre trajectoire. Le modèle s'écrit donc comme suit :

$$\begin{aligned} \text{SMAF}_{ij} &= \pi_{0i} + \pi_{1i}(\text{NB JOURS DEPUIS } 01/01/01)_{ij} + \pi_{2i}(\text{NB JOURS DEPUIS } 01/01/01)_{ij}^2 + \varepsilon_{ij} & (4) \\ \pi_{0i} &= \gamma_{00} + \gamma_{01}\text{ZONE}_i + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11}\text{ZONE}_i + \zeta_{1i} \\ \pi_{2i} &= \gamma_{20} + \gamma_{21}\text{ZONE}_i + \zeta_{2i} \end{aligned}$$

3.4 Résultats

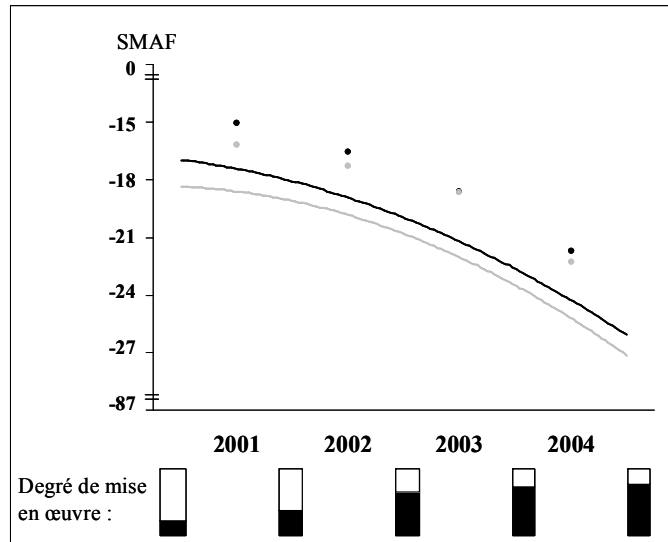
Le Tableau 1 présente les estimés des paramètres spécifiés au modèle (4) alors que la Figure 2 illustre ces résultats. Les trajectoires obtenues avec le modèle de croissance sont confrontées à celles obtenues par la méthode classique d'analyse de variance pour mesures répétées. Sous l'abscisse, le degré de mise en œuvre du RISPA (Hébert & Veil, 2004) est rapporté, indiquant le pourcentage d'implantation du RISPA dans la zone expérimentale au milieu de chaque année. Ainsi, jusqu'à maintenant, la perte d'autonomie suit la même trajectoire dans les deux zones, peut-être parce que l'implantation du RISPA n'est pas encore complétée en Estrie. Fait intéressant à noter, quoique les analyses classiques de mesures répétées basées sur les survivants fassent également ressortir une perte d'autonomie équivalente entre les zones, elles surestiment systématiquement le niveau d'autonomie fonctionnelle. Ceci n'est pas surprenant puisque les sujets exclus de telles analyses sont ceux qui décèdent ou sont institutionnalisés au cours du suivi, ces sujets étant parmi les plus fragiles au départ.

Tableau 1 : Résultats du modèle de croissance appliqué sur les données PRISMA

		Paramètre	Modèle initial	En fixant $\zeta_{2i}=0$	
<i>Effets fixes</i>					
État initial, π_{0i}	{	Ordonnée à l'origine	γ_{00}	16,95***	16,96***
		Zone témoin	γ_{01}	1,38~	1,38~
Pente initiale, π_{1i}	{	Ordonnée à l'origine	γ_{10}	0,79*	0,73*
		Zone témoin	γ_{11}	-0,46	-0,47
Courbure, π_{2i}	{	Ordonnée à l'origine	γ_{20}	0,36***	0,39***
		Zone témoin	γ_{21}	0,10	0,10
<i>Composantes de variance</i>					
Niveau 1	Intra-individu	σ^2_ε	14,64***	15,85***	
Niveau 2	Dans l'état initial	Terme linéaire	σ^2_0	124,26***	127,84***
		variance	σ^2_1	11,72**	6,35***
		covar avec état initial	σ_{01}	4,42	0,95
	Terme quadratique	variance	σ^2_2	0,39~	
		covar avec état initial	σ_{02}	-1,13	
		covar avec terme linéaire	σ_{12}	-1,40	

~ $p < 0,10$; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Figure 2 : Perte d'autonomie au cours du suivi selon la zone expérimentale (en noir) ou témoin (en gris) et le type d'analyse : modèles de croissance (lignes pleines) ou ANOVA pour mesures répétées (points)



4. DISCUSSION

Les modèles de croissance ont de nombreux avantages pour l'étude de l'évolution de l'état de santé des personnes âgées. Ils permettent une réduction considérable des biais pour tous les types de non-réponse ignorable. L'hypothèse de non-réponse aléatoire (NRA) est plausible ici puisque le décès et l'institutionnalisation sont souvent reliés aux mesures précédentes de l'état de santé. En outre, les modèles de croissance permettent le maintien de la puissance statistique désirée puisque le modèle permet l'ajout de sujets aux différentes vagues d'évaluations. Enfin d'autres caractéristiques du modèle le rendent intéressant, notamment le fait que les vagues de collectes peuvent être espacées irrégulièrement, que le nombre et l'espacement des vagues de collecte peuvent varier d'un participant à l'autre et que le modèle soit flexible dans la définition du temps (numéro de vague, date d'évaluation, âge du sujet). De plus, il y a plusieurs possibilités pour modéliser le patron de changement d'une variable dans le temps (linéaire, quadratique, logistique, ...). Enfin, un avantage additionnel non discuté dans cet article, les variables explicatives du changement peuvent bouger dans le temps; on les incorpore alors dans le modèle de niveau 1.

RÉFÉRENCES

- Bryk, A.S. et Raudenbush, S.W. (2002), *Hierarchical linear models: Applications and data analysis methods* (2nd ed.), Thousand Oaks, CA: Sage.
- Desrosiers, J., Bravo, G., Hébert, R. et Dubuc, N. (1995), "Reliability of the revised functional autonomy measurement system (SMAF) for epidemiological research", *Age and Ageing*, 24, pp. 402-406.
- Hébert, R. (1982), "L'évaluation de l'autonomie fonctionnelle des personnes âgées", *Canadian Family Physician*, 28, pp. 754-762.
- Hébert, R., Bravo, G., Korner-Bitensky, N. et Voyer, L. (1996), "Predictive validity of a postal questionnaire for screening community dwelling elderly individuals at risk of functional decline", *Age and Ageing*, 25, pp. 159-167.
- Hébert, R., Carrier, R. et Bilobeau, A. (1988), "The functional autonomy measurement system (SMAF): Description and validation of an instrument for measurement of handicaps", *Age and Ageing*, 17, pp. 293-302.

- Hébert, R., Guilbault, J., Desrosiers, J. et Dubuc, N. (2001), "The functional autonomy measurement system (SMAF): A clinical-based instrument for measuring disabilities and handicaps in older people", *Geriatrics Today: Journal of the Canadian Geriatrics Society*, 4, pp. 141-147.
- Hébert, R. et Veil, A. (2004), "Monitoring the degree of implementation of an Integrated Delivery System", *International Journal of Integrated Care*, 4, pp. 1-11.
- Rogosa, D.R. et Willett, J.B. (1985), "Understanding correlates of change by modeling individual differences in growth", *Psychometrika*, 50, pp. 203-228.
- Singer, J.D. et Willett, J.B. (2003), *Applied longitudinal data analysis: modeling change and event occurrence*, New York: Oxford University Press.