



Sciences Economiques et Sociales de la Santé
& Traitement de l'Information Médicale

sesstim.univ-amu.fr

Thibault MURA

*Unité de recherche Clinique et Epidémiologique, Université de Montpellier, INSERM U1061,
CHRU Montpellier, France*

**Effet plafond, effet plancher, curvilinearité et analyses de données longitudinales :
Biais lié à l'ajustement sur les valeurs initiales, l'exemple de l'analyse du déclin cognitif**

avril 2018



Cliquez ici pour voir l'intégralité des ressources associées à ce document

Effet plafond, effet plancher, curvilinearité et analyses de données longitudinales : Biais lié à l'ajustement sur les valeurs initiales, L'exemple de l'analyse du déclin cognitif

Dr Thibault MURA
MCU-PH
Unité de Recherche Clinique et Epidémiologique
CHRU Montpellier

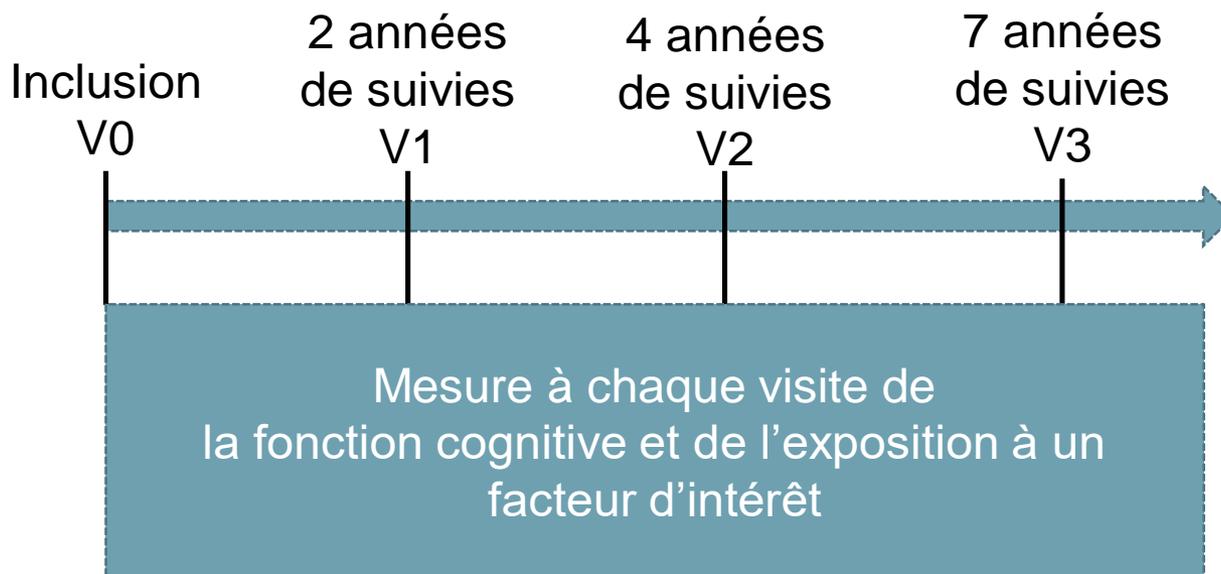
Plan

1. Analyse longitudinale d'un critère de jugement quantitatif :
exemple du déclin cognitif
2. Effet plafond, effet plancher, Curvilinearité
3. Problème posé par l'ajustement sur les valeurs initiales
4. Méthodes de prise en compte des effets plafond et plancher
et de la curvilinearité

Analyse longitudinale d'un critère
de jugement quantitatif : exemple du déclin cognitif

Analyses longitudinales

- On s'intéresse à l'analyse d'une *variable quantitative* dont la *mesure est répétée* dans le temps et dont on souhaite étudier *l'évolution* en fonction d'une exposition d'intérêt
- Exemple l'étude du déclin cognitif dans l'étude des « 3 Cités »



Plusieurs stratégies d'analyse possibles

1. Régression linéaire sur la base d'une différence entre deux points de suivi (score V3 - score V0)
 - Méthode fréquemment utilisée si l'on dispose uniquement de deux points de mesure
 - *Inconvénients : choix des visites d'intéret, données manquantes, perdus de vue, perte d'information pour les données entres les visites d'intérêt.*
2. Régression logistique après dichotomisation du critère de jugement absence/présence de déclin
 - Dichotomisation après calcul de différence entre deux points de suivi
 - Utilisation d'un seuil pour définir le déclin
 - *Même limite que précédemment + Pb du choix du seuil*

Plusieurs stratégie d'analyses possibles

3. Méthode de régression linéaire mixte (avec intercept +/- pente aléatoire propre au sujet)

$$Y_{ij} = (\beta_0 + \gamma_{0i}) + (\beta_1 + \gamma_{1i})t_{ij} + \beta_2 X_i + \beta_3 X_i t_{ij} + \varepsilon_{ij}$$

avec $\gamma_{0i} \sim \mathcal{N}(0, \sigma_0^2)$, $\gamma_{1i} \sim \mathcal{N}(0, \sigma_1^2)$ et $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

La validité du modèle linéaire mixte repose sur un certain nombre d'hypothèses :

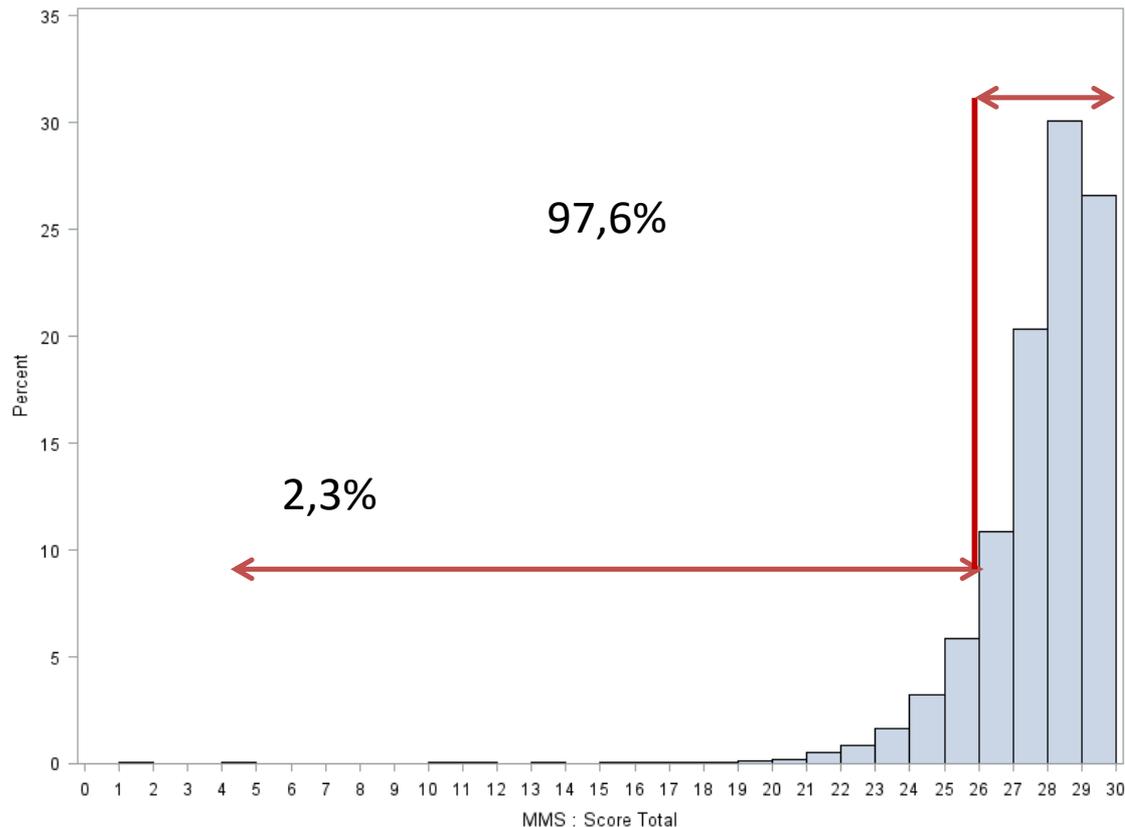
1. variable à expliquer doit être quantitative continue
2. Les effets aléatoires du modèle doivent être Gaussiens
3. Une variation d'une unité sur une variable explicative (prédicteur) doit être associée à un changement fixe, constant sur la variable à expliquer.

Modèle robuste pour l'hypothèse 2 à condition que l'hypothèse N°3 soit respectée

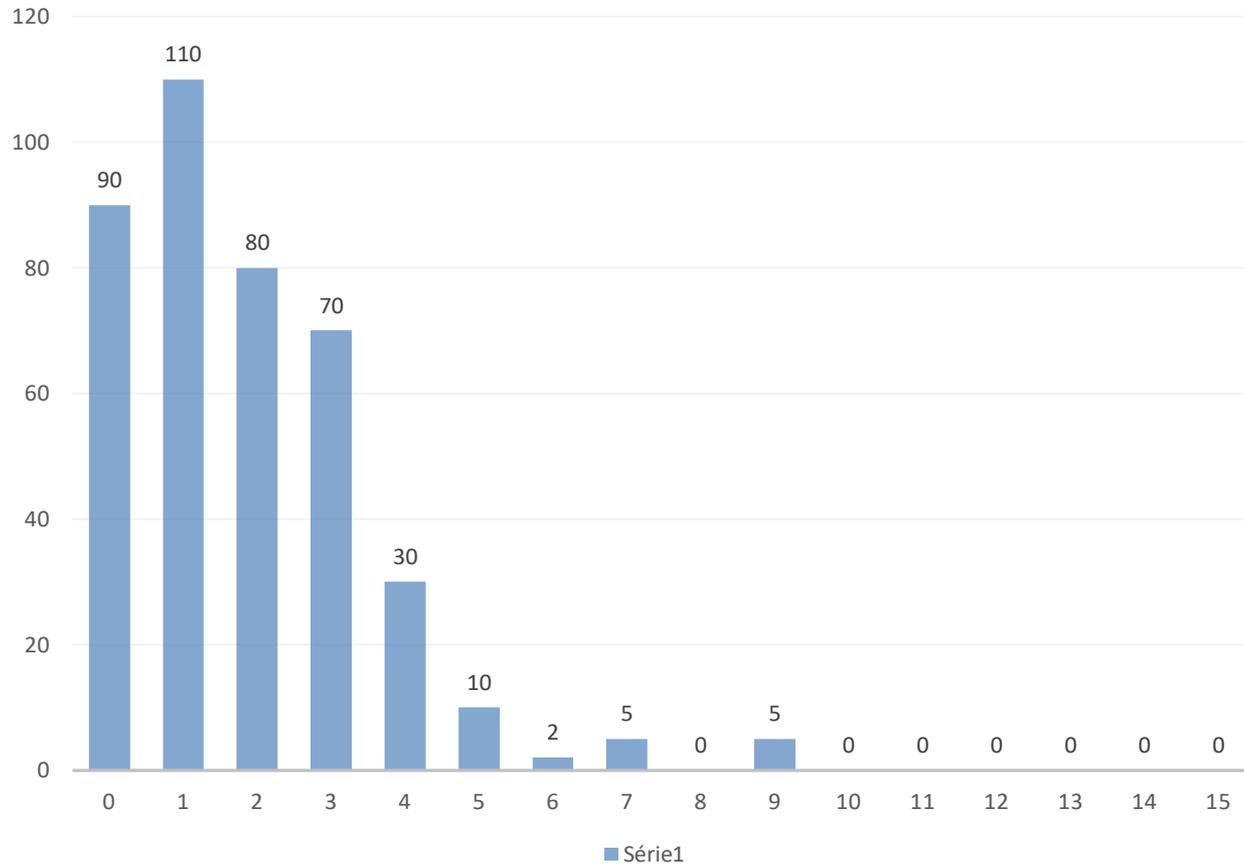
Effet plafond, effet plancher, Curvilinearité

Effet Plafond : le MMSE en population générale

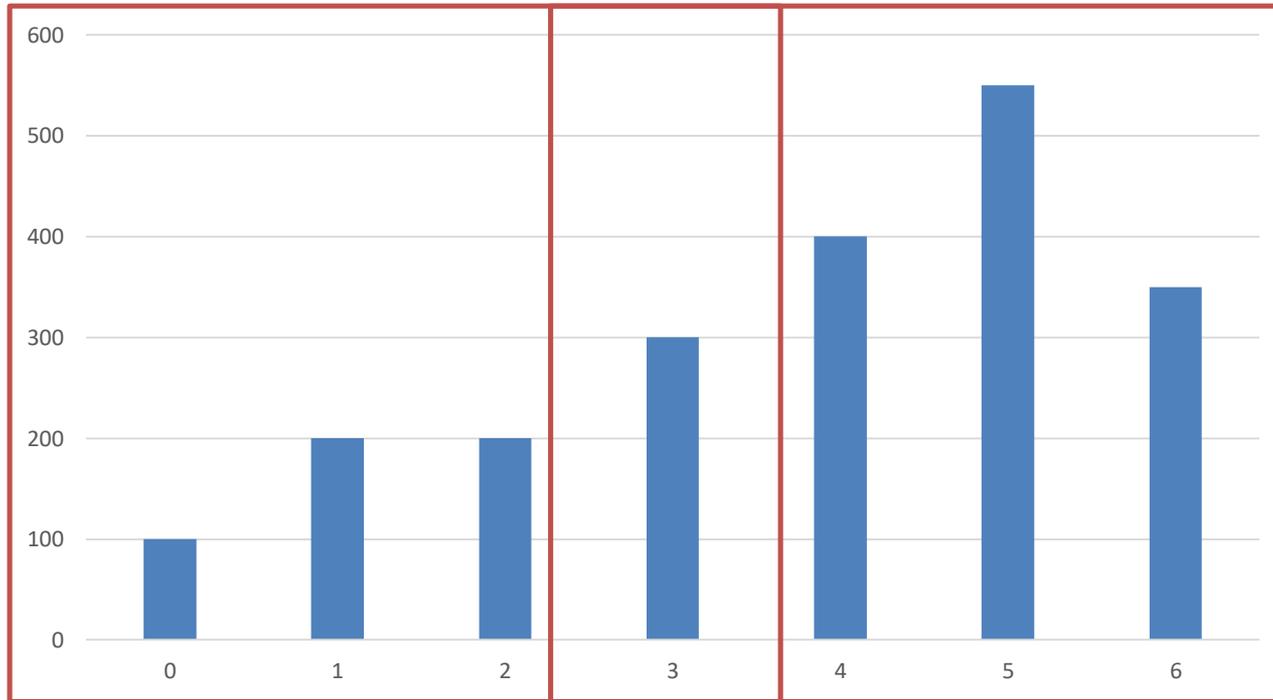
- Un effet de plafond apparait sur un variable lorsque la variance d'une variable dépendante n'est pas mesurée ou estimée au-dessus d'un certain niveau. Quand un effet de plafond se produit il y a un regroupement des scores au niveau supérieur rapporté par un instrument



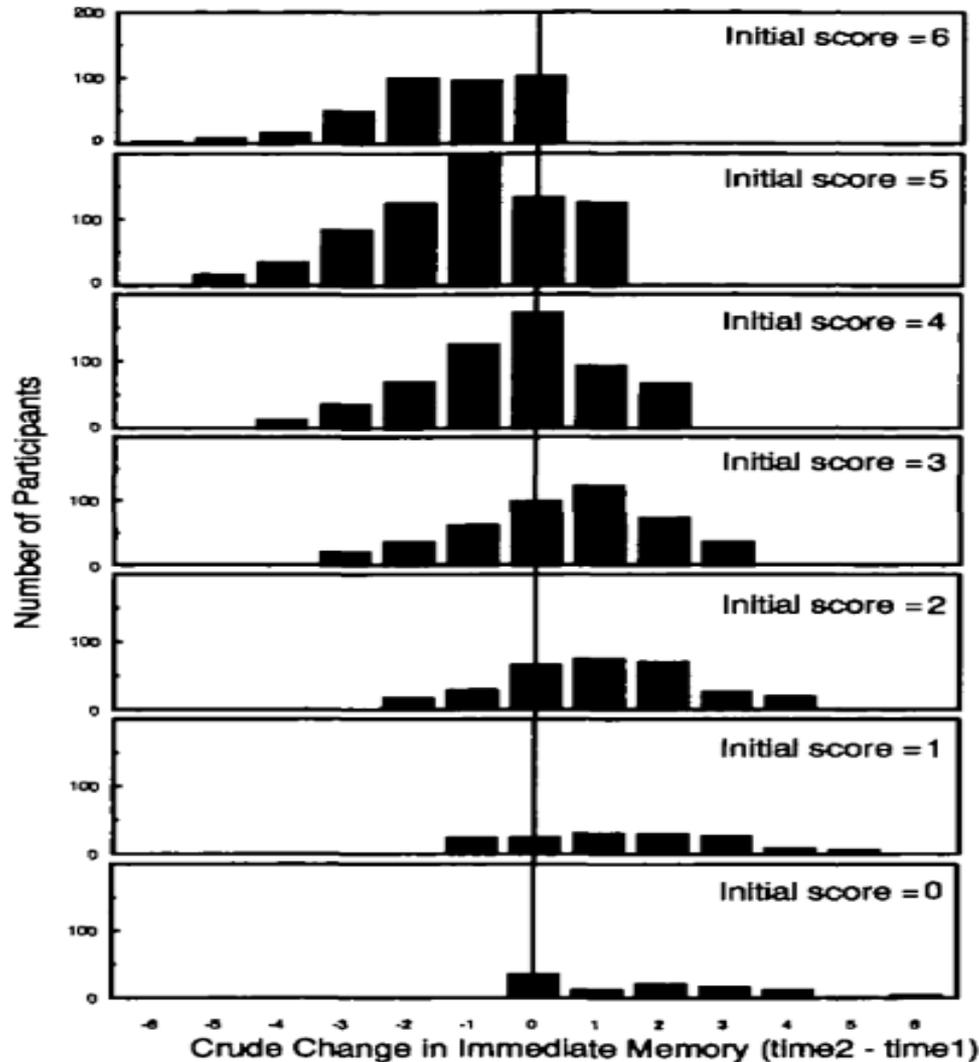
Effet plancher



Effet plafond et effet plancher



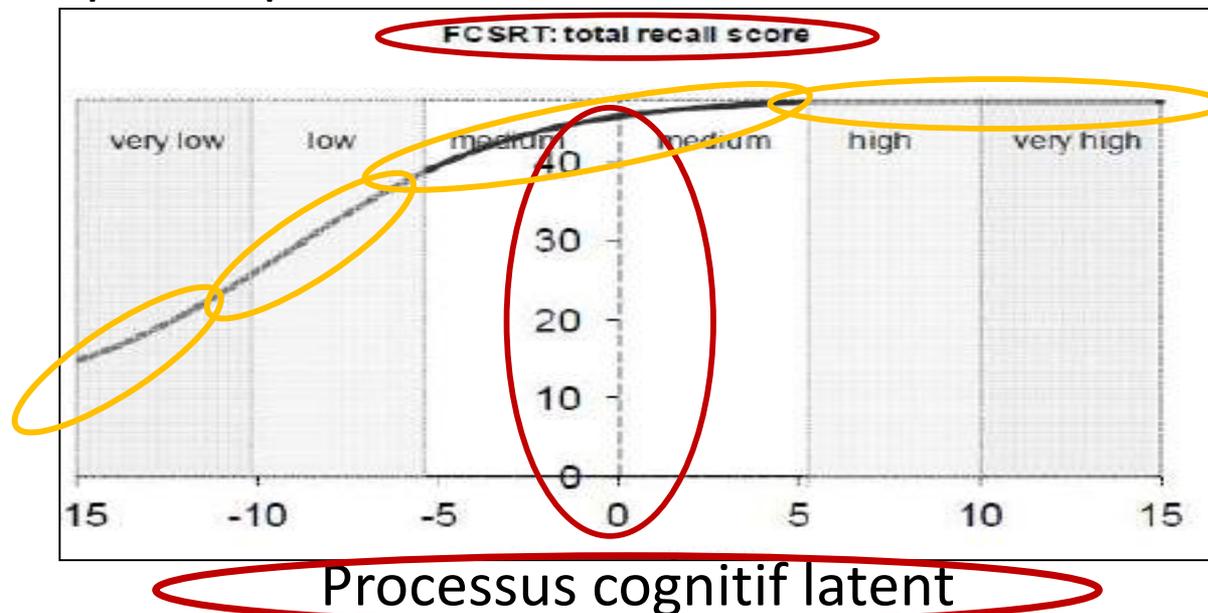
Conséquence des effets plafonds et planchers sur l'analyse de l'évolution du score cognitif



*rapporté par (Morris, Evans et al. 1999)
d'après (Evans, Beckett et al. 1993)*

Curvilinearité : Sensibilité variable au changement cognitif

- Une propriété importante demandée aux tests neuropsychologiques est de présenter des intervalles de mesure "égaux" tout au long du score.
- Exemple: une personne qui perd 5 points à un test en passant de 30 à 25 manifeste-t-elle la même évolution sur la fonction cognitive qu'une personne évoluant de 15 à 10



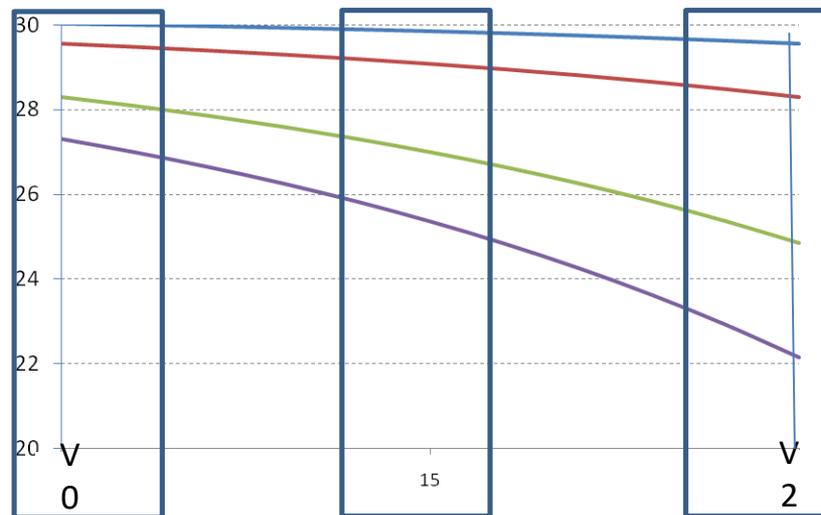
Curvilinearité : Sensibilité variable au changement cognitif :

- Une propriété importante demandée aux tests neuropsychologiques est de présenter des intervalles de mesure "égaux" tout au long du score.
- propriété est en réalité rarement démontrée, et fautive pour beaucoup des tests neuropsychologiques utilisés (*Galasko, Gould et al. 2000; Proust-Lima, Amieva et al. 2007*).
- non-prise en compte → estimations biaisées et mise en évidence de fausses associations, surtout si association entre la variable d'intérêt et le niveau cognitif initial (*Proust-Lima, Dartigues et al. 2011*).

Ajustement sur les valeurs initiales lors de l'analyse du déclin cognitif

Illustration des variations d'un test neuropsychologique d'un même déclin cognitif latent sur 3 visites chez 4 sujets en fonction du score initial

Test neuropsychologique avec effet plafond et faible sensibilité au changement pour les niveaux cognitifs élevés



Ajustement sur les valeurs initiales lors de l'analyse du déclin cognitif

TABLE 5. Relationship* between chronic use of benzodiazepine and cognitive decline. Results based on the exposed-non exposed subsample†

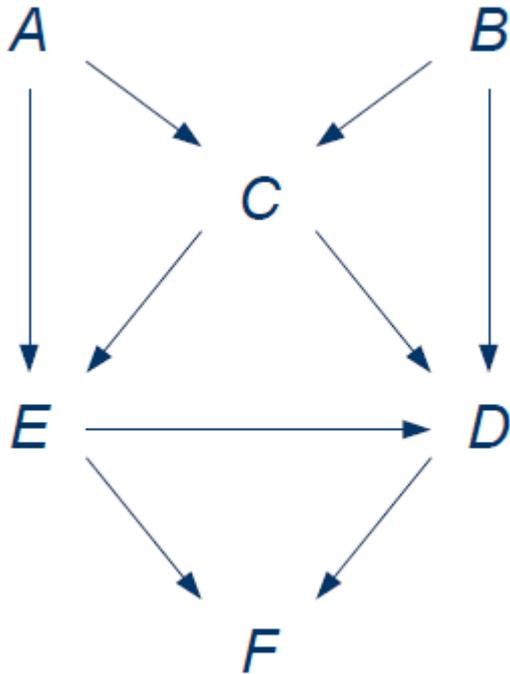
	Nonusers (N = 126)	Chronic users (n = 63)	
Cognitive decline	<i>n</i> (%)	<i>n</i> (%)	OR (95% CI)
Mini-Mental State Examination	19 (15.3)	14 (22.2)	2.0 (0.9–4.4)
Wechsler Digit Symbol Substitution Test	31 (25.0)	23 (36.5)	2.2 (1.1–4.5)
Trail Making Test	35 (29.7)	26 (43.3)	2.2 (1.1–4.6)
Finger Tapping Test	19 (15.5)	8 (12.7)	0.9 (0.3–2.2)
Auditory Verbal Learning Test	9 (8.0)	7 (12.3)	1.7 (0.5–5.1)

*ORs computed from logistic regression adjusting on baseline cognitive score.

†CI, confidence interval; OR, odds ratio.

Biais lié à l'ajustement sur les valeurs initiales

Causal Directed Acyclic Graph (DAG)



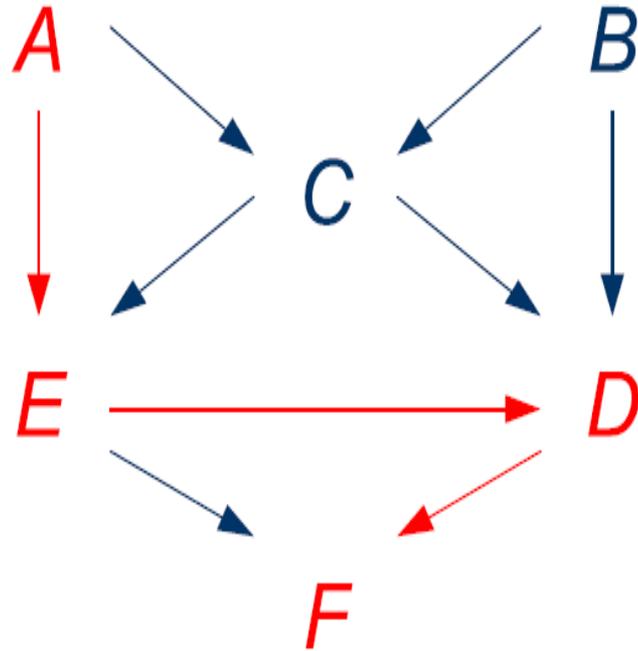
DAG pour :

- D: Directed, dirigé puisque toutes les flèches vont dans un seul sens
- A: Acyclique, puisqu'il ne présente pas de boucle, aucune variable ne se cause elle-même
- G: Graph

Composé de nœuds (variables) et de flèches

- Permet de représenter des chemins d'associations causales directes ou indirectes, les associations structurelles et les associations conditionnelles
- ➔ Permet de choisir une stratégie d'ajustement et d'éviter les biais de « collision » (association conditionnelles)

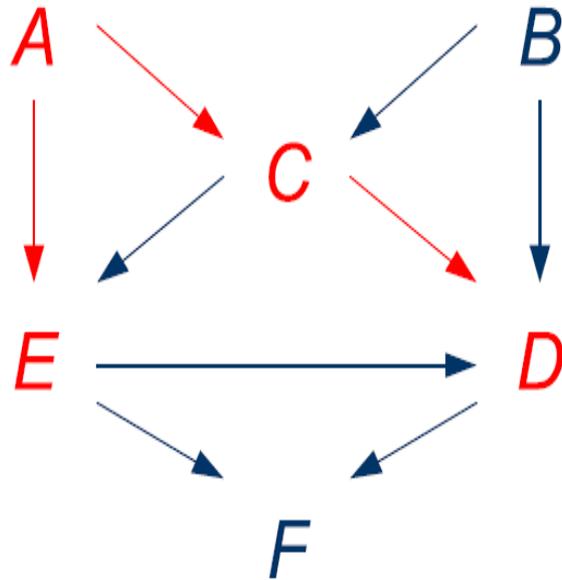
Chemin d'association causale directe et indirecte



Chemin dirigé (directed path): suite de flèches qui pointent dans la même direction. ($A \rightarrow E \rightarrow D \rightarrow F$)

Il s'agit d'un **chemin d'association causal**

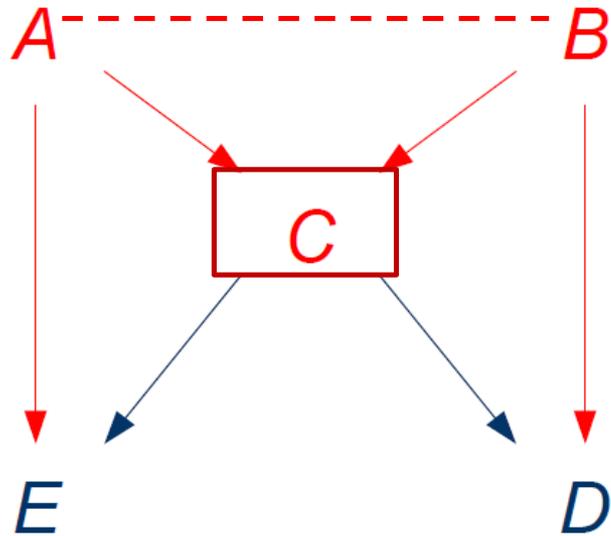
un chemin d'association structurelle (facteur de confusion)



Chemin dérobé (back-door path) : les flèches pointent sur les variables initiales (E ← A → C → D)

Il s'agit d'un chemin d'association structurelle et non d'un chemin causal

chemin contenant un « collider » si deux flèches pointent sur un même nœud

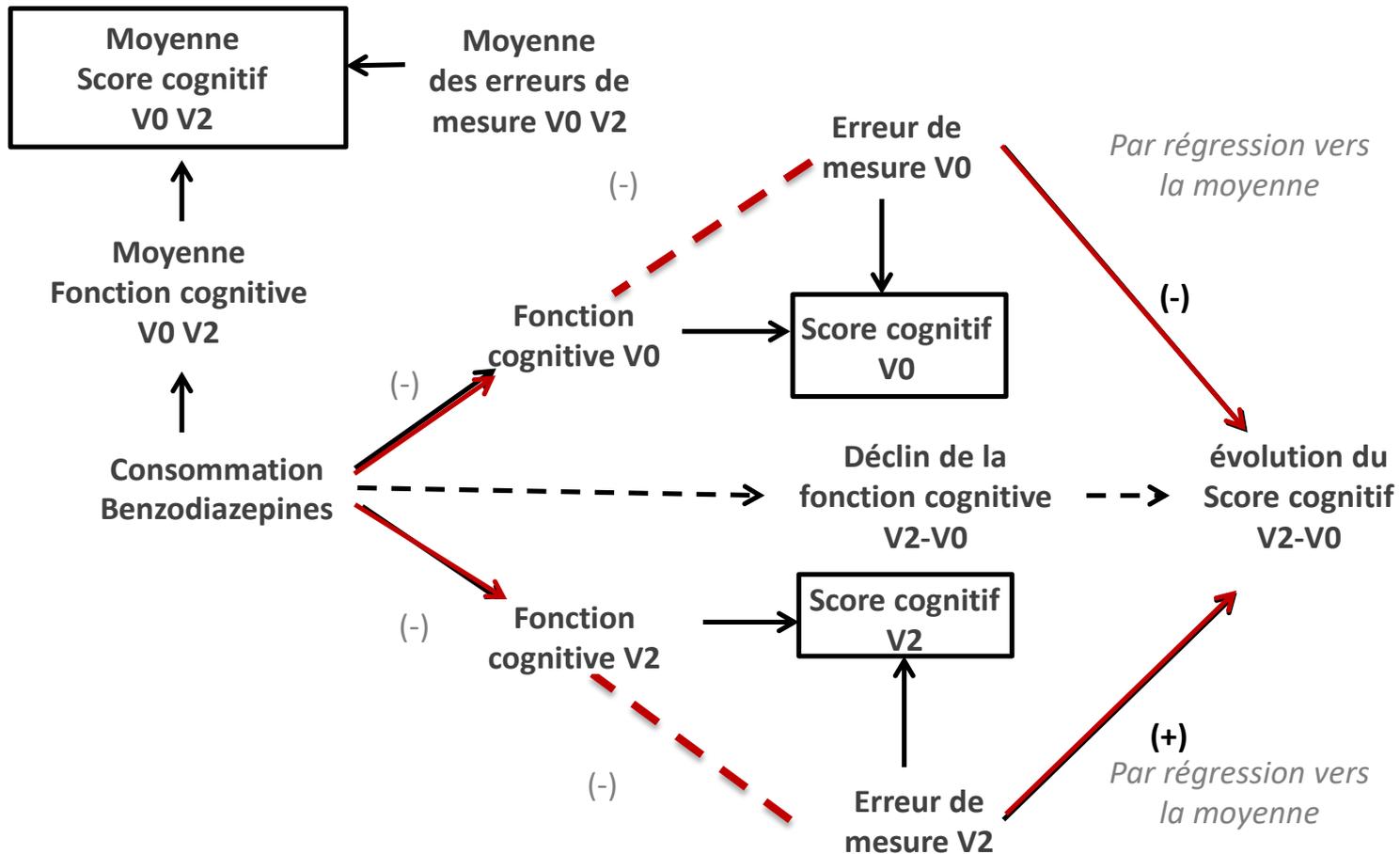


- Si C est une conséquence à la fois de A et de B
- Alors A et B seront associées dans les strates de C, même si A et B sont indépendantes dans la population

Un chemin contenant un « collider » : ne transmet pas d'association sauf si on conditionne (ajuste) sur le « collider » (ou sur un descendant d'un « collider »)

En cas d'ajustement sur le « collider » on parle **d'association conditionnelle (biais de collision)**

Exemple: étude des relations entre consommation de benzodiazépines et déclin cognitif



Ajustement lors de l'analyse du déclin cognitif

Etude du déclin entre 2 visites (V0 et V2) d'un score de fluence verbale (Isaac Set Test (IST)) selon la consommation de benzodiazépines

Modélisation du delta (V2-V0) en fonction de la consommation de benzodiazépines (4200 sujets):

Modèle linéaire avec différents ajustements :

- M1: sur facteurs de confusion classiques
- M2: sur facteurs de confusion classiques + score initial V0
- M3: sur facteurs de confusion classiques + score V2
- M4: sur facteurs de confusion classiques + score moyen

Beta pour benzo=1

M1		M2 (ajust V0)		M3 (ajust V2)		M4	
Beta	pvalue	Beta	pvalue	Beta	pvalue	Beta	pvalue
+0,07	0,77	-0,52	0,02	+0,59	0,01	+0,11	0,64

Ajustement lors de l'analyse du déclin cognitif

Etude du déclin sur 4 visites (V0 à V3) d'un score de fluence verbale (Isaac Set Test (IST)) selon la consommation de benzodiazépines

Modélisation de l'évolution IST entre V0 et V3 en fonction de la consommation de benzodiazépines (4200 sujets) (Interaction temps*benzo)

Modèle linéaire mixte avec différents ajustements :

- M1: benzo âge temps benzo*temps
- M2: M1+interaction Isaac_V0*temps
- M3: M1+interaction Isaac_V3*temps
- M4: M1+interaction Isaac_moyen*temps

Beta pour Interaction benzo*temps

M1		M2 (ajust V0)		M3 (ajust V3)		M4	
Beta	pvalue	Beta	pvalue	Beta	pvalue	Beta	pvalue
-0,10	0,29	-0,27	0,004	+0,09	0,30	-0,10	0,34

Simulations

- Un niveau cognitif latent initial (de moyenne différente entre les groupes, de variance fixée, identique) + une erreur de mesure aléatoire à chaque visite (de moyenne 0, de variance fixée)
- Évolution dans le temps identique dans les deux groupes à partir du niveau cognitif latent initiale

Conclusion :

- Biais systématique sur l'interaction entre l'évolution et les valeurs initiales (Tu, Y. K. and M. S. Gilthorpe (2007)) qui se répercute sur l'interaction avec le facteur d'intérêt.
- Biais proportionnel à l'importance de la différence entre les groupes à l'inclusion
- $Biais = difference\ initiale * \left(\frac{var(latent) + var(erreur) - var(latent)}{var(latent) + var(erreur)} \right)$

Méthodes de prise en compte des effets plancher, plafond et de la curvilinearité

prise en compte des effets plancher, plafond et de la curvilinearité

1. Transformation de la variable quantitative

– Avantage

- Susceptible de corriger le phénomène de curvilinearité
- Possible si on ne dispose que de 2 mesures

– Difficulté

- Identifier la transformation adéquat
- incapacité à corriger les biais liés aux effet plafond et plancher

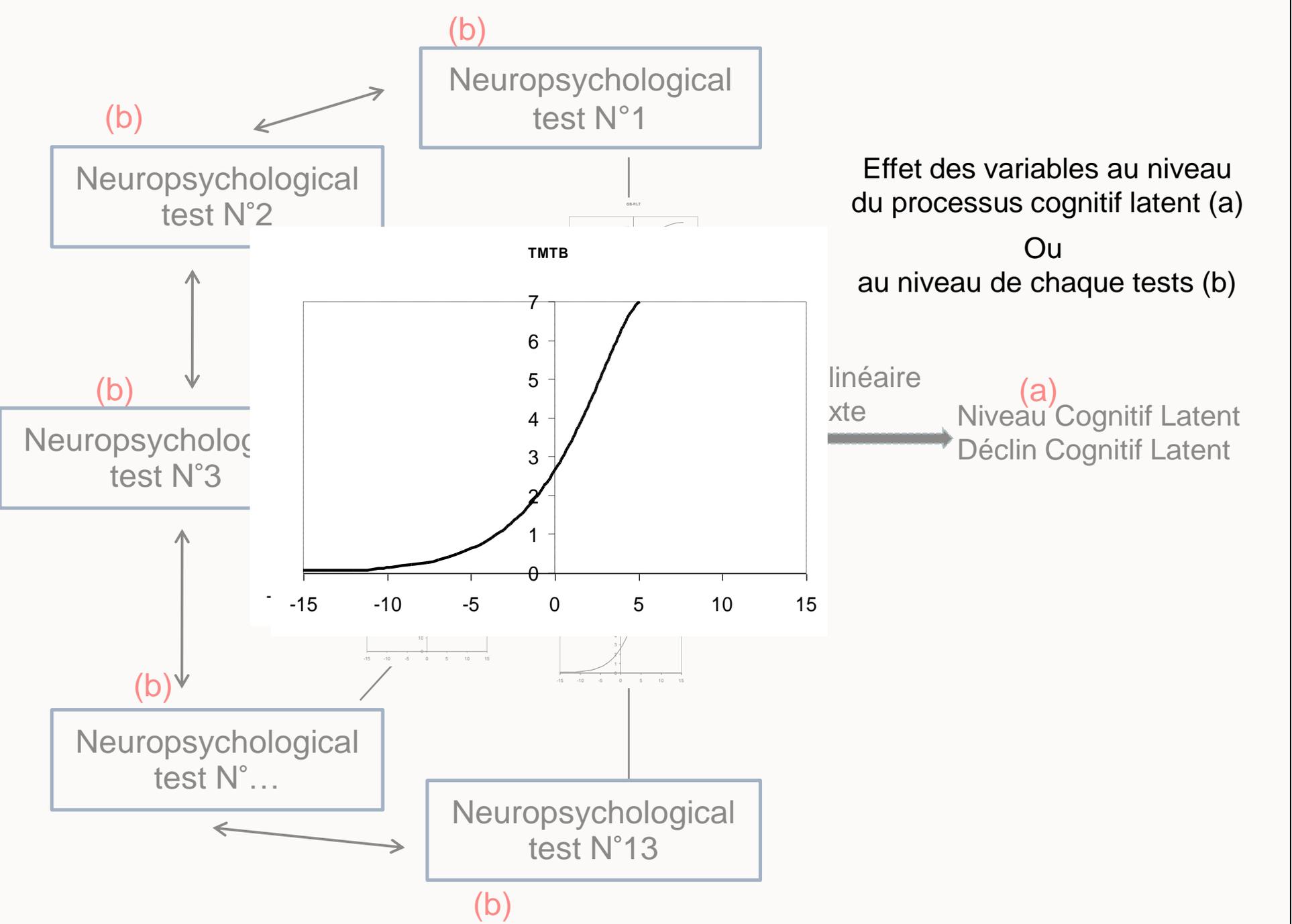
2. Ajustement sur la valeur moyenne des observations

- Méthode de Oldham's (Tu, Y. K. and M. S. Gilthorpe (2007))
- Efficacité à valider pour prendre en compte l'ensemble des biais liés aux effets plafonds et planchers

Modèle non linéaire à processus latent pour données longitudinales

3. Modèle non linéaire à processus latent pour données longitudinales (Proust, C., (2006)) :

- Lorsqu'il existe plusieurs mesure d'un même concept latent
- Package R : LCMM
- Exemple du déclin cognitif :
 - *Le modèle assume que la corrélation entre les tests neuropsychologiques est induite par l'existence d'un processus cognitif latent non mesurable directement. Les scores des tests sont ainsi considérés comme des mesures indirectes d'un processus cognitif commun, latent, transformé par une relation non linéaire spécifique à chacun de ces tests.*



Modèle en deux parties : 1/ modèle structurel

- **Modèle structurel** : décrit l'évolution du processus cognitif latent au fil du temps et évalue les effets communs de co-variables sur cette trajectoire à l'aide d'un modèle linéaire mixte

$$\Lambda_i(t) = Z_i(t)^T U_i + X1_i(t)^T \beta + \sigma_\omega \omega_i(t)$$

Avec :

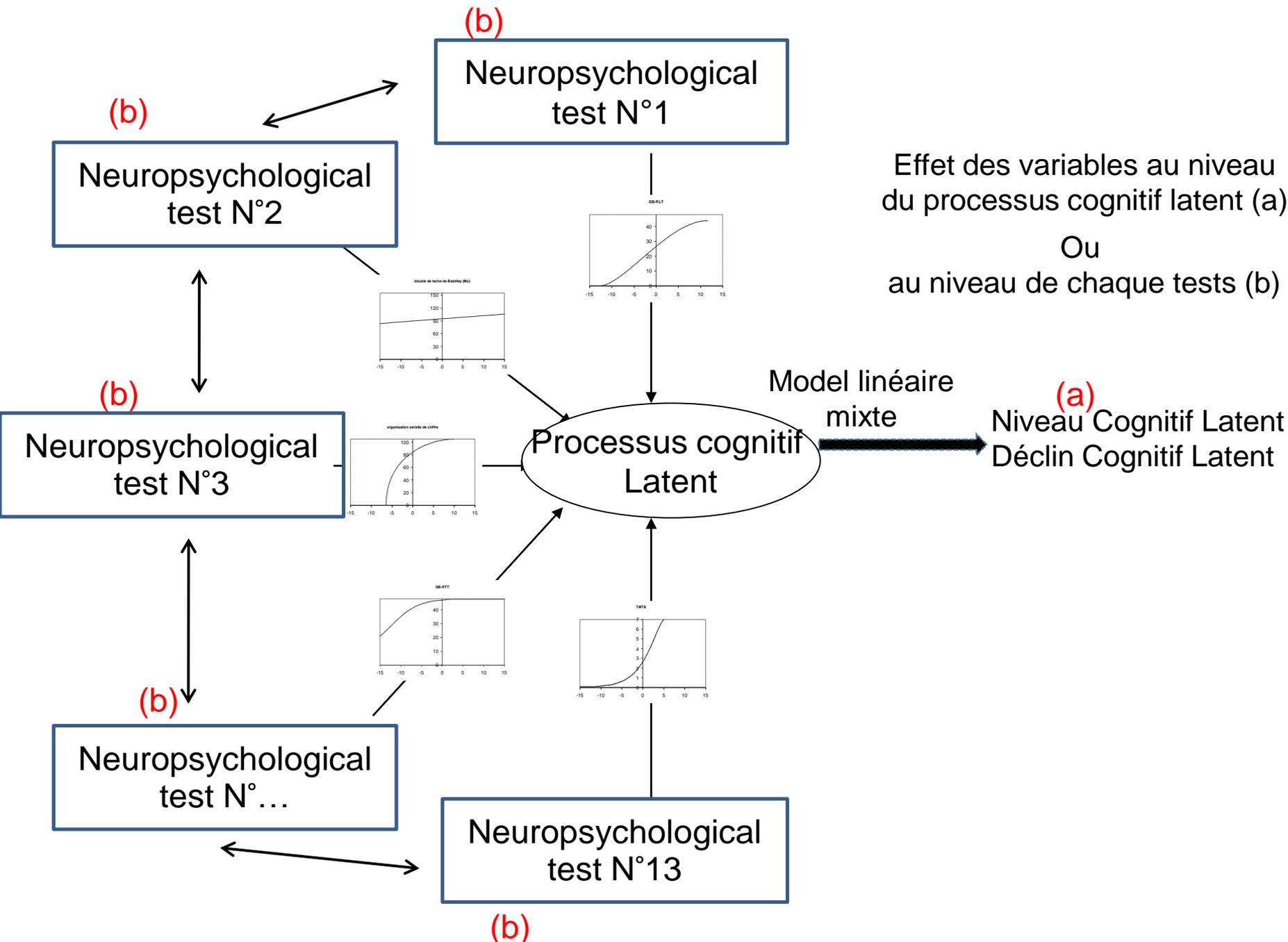
- $\Lambda_i(t)$: Valeur du processus latent pour le sujet i au temps t
- $Z_i(t)$: contient 1 pour l'intercept aléatoire et t pour la pente aléatoire +/- d'autres fonctions du temps (fréquemment une fonction polynomiale du temps)
- $U_i(t)$: vecteur des effets aléatoires au niveau du sujet i (intercept aléatoire, pente aléatoire selon fonctions du temps)
- $X1_i$: vecteur des variables explicatives (dépendant ou non du temps)
- β : vecteur des effets fixes du modèle
- $\sigma_\omega \omega_i(t)$: est un mouvement Brownien standard, modélisant les variations locales et les écarts à la trajectoire du processus latent.

Modèle en deux parties : 2/modèle d'observation ou de mesure

- **Modèle d'observation** : relie les tests neuropsychologiques avec le processus cognitif latent selon une relation non linéaire spécifique au test permet de décrire les propriétés métrologiques des tests neuropsychologiques et d'évaluer l'effet spécifique des co-variables sur les tests neuropsychologiques.

$$((g(Y_{ijk}; \eta_{1k}, \eta_{2k}) - \eta_{3k}) / \eta_{4k} = \check{y}_{ijk} = \Lambda_i(t_{ijk}) + \alpha_{ik} + X_{2i}(t_{ijk})^T \gamma_k + \varepsilon_{ijk}$$

- \check{y}_{ijk} : est la valeur transformée, estimée par le modèle, de la mesure du test k pour le sujet i, lors de l'occasion j.
- $\Lambda_i(t)$: Valeur du processus latent pour le sujet i au temps t
- α_{ik} : effet aléatoire au niveau du sujet i pour le test k. Permet de prendre en compte le fait que pour une même valeur de processus latent, deux sujets peuvent présenter un score différent dans le domaine cognitif associé au test neuropsychologique k
- $X_{2i}(t)$: vecteur des variables explicatives (dépendant ou non du temps)
- γ_k : Vecteur de contrastes associés aux variables explicatives pour le test k. Ces contrastes correspondent à l'effet des variables explicatives spécifiquement sur le test k en plus de leur effet sur le processus cognitif latent. (Remarque : la somme de ces contrastes pour une variable donnée sur les k tests est égale à 0).
- ε_{ijk} : erreur de mesure (distribution gaussienne de moyenne 0 et de variance $\sigma_{\varepsilon k}^2$)



Références

Références

- **Galasko, D. R., R. L. Gould, I. S. Abramson and D. P. Salmon (2000).** "Measuring cognitive change in a cohort of patients with Alzheimer's disease." Stat Med **19**(11-12): 1421-32.
- **Greenland, S., Pearl, J., Robins, J.M., (1999).** Causal diagrams for epidemiologic research. Epidemiology. 10, 37-48.
- **Morris, M. C., D. A. Evans, L. E. Hebert and J. L. Bienias (1999).** "Methodological issues in the study of cognitive decline." Am J Epidemiol **149**(9): 789-93.
- **Glymour, M. M., J. Weuve, L. F. Berkman, I. Kawachi and J. M. Robins (2005).** "When is baseline adjustment useful in analyses of change? An example with education and cognitive change." Am J Epidemiol **162**(3): 267-78.
- **Tu, Y. K. and M. S. Gilthorpe (2007).** "Revisiting the relation between change and initial value: a review and evaluation." Stat Med **26**(2): 443-457.
- **Proust-Lima, C., J. F. Dartigues and H. Jacqmin-Gadda (2011).** "Misuse of the linear mixed model when evaluating risk factors of cognitive decline." Am J Epidemiol **174**(9): 1077-88.
- **Proust, C., H. Jacqmin-Gadda, J. M. Taylor, J. Ganiayre and D. Commenges (2006).** "A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data." Biometrics **62**(4): 1014-24.

Merci