ParkProg project

The 4S method for the longitudinal analysis of multidimensional questionnaires:

application to Parkinson's disease progression from patient perception

Tiphaine Saulnier *, Cécile Proust-Lima, Daniel J. Myall, Toni L. Pitcher, Leslie Livingston, Rachel Nolan, Sophie Grenfell, John C. Dalrymple-Alford, Alexandra Foubert-Samier, Wassilios G. Meissner, Tim J. Anderson, and Campbell J. Le Heron

* *postdoctoral researcher in Biostatistics, at the NZBRI*

IBS-BIBC 2025, Canberra (AUS), 24-28 November

Parkinson's Disease (PD)

Neurodegenerative disorder

+10.000.000 people with PD worldwide
(13.000 in NZ, 219.000 in AUS, 270.000 in FR, ...)
$>$1% after 65 years old

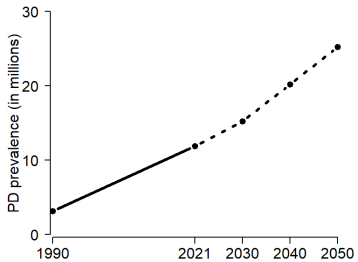
Symptomatology

Progressive loss of neurons in the brain, inducing

- motor symptoms
  rest tremor, slowness, rigidity, ..
- non-motor symptoms
  speech difficulties, sleep disturbances, cognitive issues, anxiety/depression, ...

Potential factors ageing, pesticide exposure, genetic mutation, ...

Treatment no cure



[Su *et al.*, BMJ 2025]

☞ **Real need of research**

Forecast: PD prevalence to double by 2050
↪ Better understand the origin, progression, and factors
↪ Assess risks, provide better support to patients, and explore therapeutic avenues

Health-related Quality of Life (Hr-QoL)

What?

Health-related Quality of Life = [Testa *et al.*, 1996]
patient's perception of their own health condition, combining

- physical health (pain, fatigue, mobility, ...)
- psychological well-being (self-esteem, +/- affects, ...)
- social relationships (support, role, work, ...)

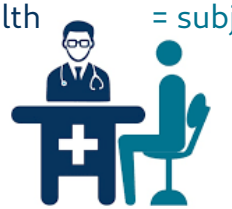
Why?

Subjective health, growing interest in studies with

- another perspective on disease progression
- deliver better support based on patient's feelings
- FDA requirement for the development of new treatments

clinician's perception = objective health

patient's perception = subjective health



How?

Hr-QoL measure instruments

patient-reported questionnaires (PRO data)

Hr-QoL in PD

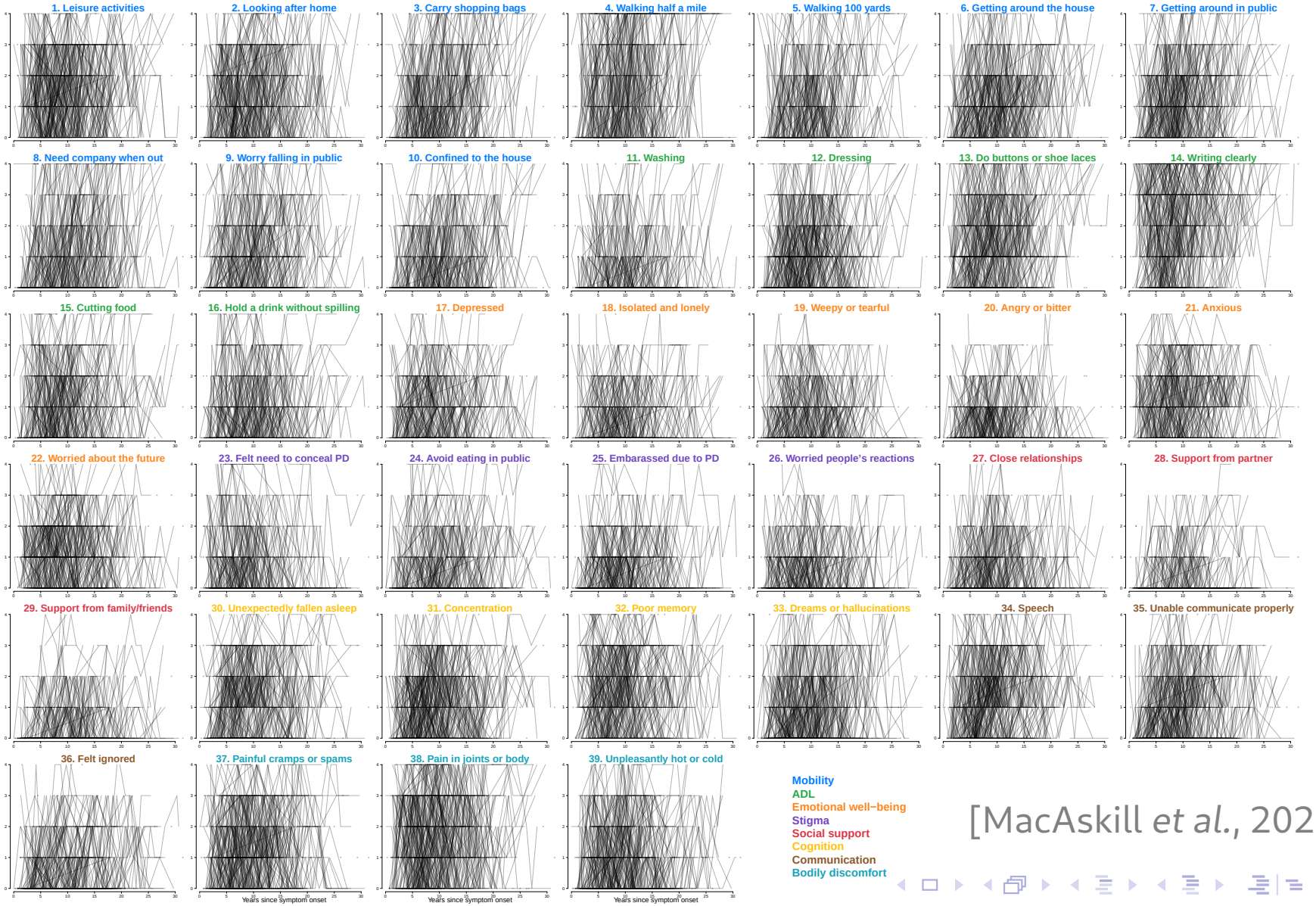
		Due to having PD, how often during the last month have you...
ORIGINAL PDQ-39 SCALE	MOBILITY	1. Had difficulty doing the leisure activities which you would like to do?
		2. Had difficulty looking after your home, e.g. DIY, housework, cooking?
		3. Had difficulty carrying bags of shopping?
		4. Had problems walking half a mile?
		5. Had problems walking 100 yards?
		6. Had problems getting around the house as easily as you would like?
		7. Had difficulty getting around in public?
		8. Needed someone else to accompany you when you went out?
		9. Felt frightened or worried about falling over in public?
		10. Been confined to the house more than you would like?
	ADL	11. Had difficulty washing yourself?
		12. Had difficulty dressing yourself?
		13. Had problems doing up your shoe laces?
		14. Had problems writing clearly?
		15. Had difficulty cutting up your food?
		16. Had difficulty holding a drink without spilling it?
	EMOTIONAL WELL-BEING	17. Felt depressed?
		18. Felt isolated and lonely?
		19. Felt weepy or tearful?
		20. Felt angry or bitter?
		21. Felt anxious?
		22. Felt worried about your future?
	STIGMA	23. Felt you had to conceal your Parkinson's from people?
		24. Avoided situations which involve eating or drinking in public?
		25. Felt embarrassed in public due to having Parkinson's disease?
		26. Felt worried by other people's reaction to you?
	SOCIAL SUPPORT	27. Had problems with your close personal relationships?
		28. Lacked support in the ways you need from your spouse or partner?
		29. Lacked support in the ways you need from your family or close friends?
	COGNITION	30. Unexpectedly fallen asleep during the day?
		31. Had problems with your concentration, e.g. when reading or watching TV?
		32. Felt your memory was bad?
		33. Had distressing dreams or hallucinations?
	COMMUNICATION	34. Had difficulty with your speech?
		35. Felt unable to communicate with people properly?
		36. Felt ignored by people?
	BODILY DISCOMFORT	37. Had painful muscle cramps or spasms?
		38. Had aches and pains in your joints or body?
		39. Felt unpleasantly hot or cold?

the PDQ-39 questionnaire

developed by Peto *et al.*, 1995
specific to PD
self-reported

39 questions/*items*
8 dimensions
5 ordinal levels of response
- 0 : never
- 1 : occasionally
- 2 : sometimes
- 3 : often
- 4 : always

PDQ-39 item observations on NZP[3] (N=410, n=1436)



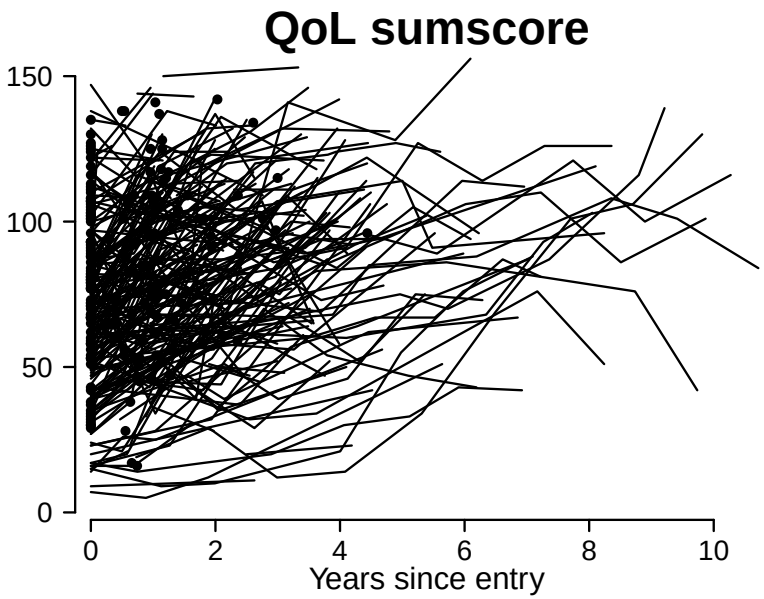
[MacAskill *et al.*, 2023]

Summarize multivariate ordinal data

Sumscore

$\longrightarrow$ loss of information

- no distinction between items
- no missing item allowed

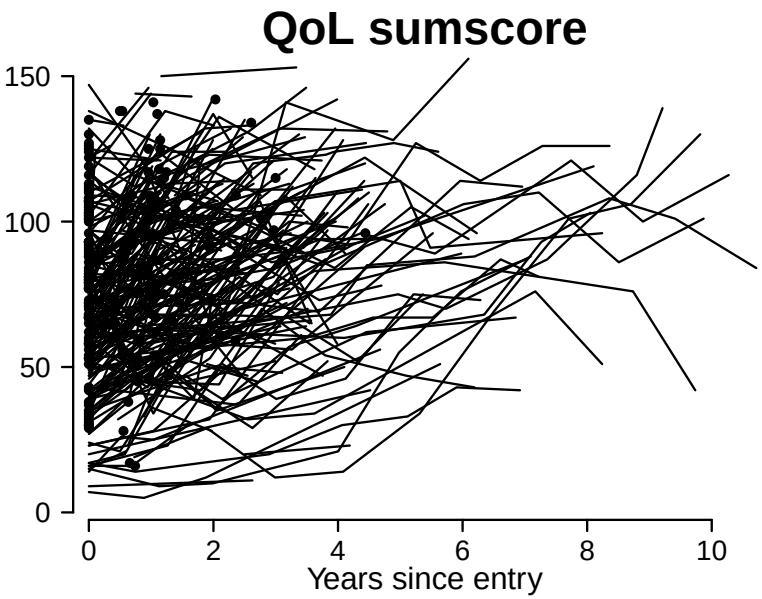


Summarize multivariate ordinal data

Sumscore

$\longrightarrow$ loss of information

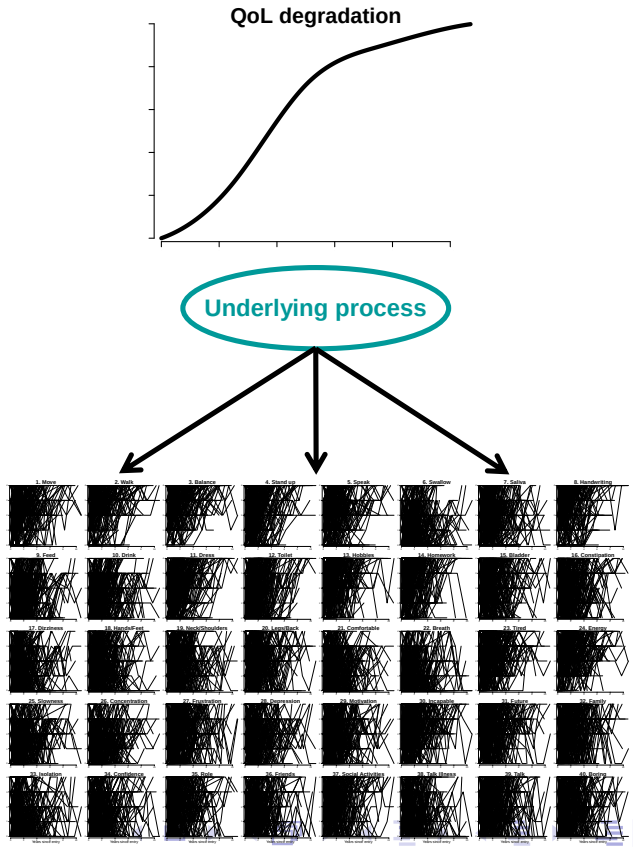
- no distinction between items
- no missing item allowed



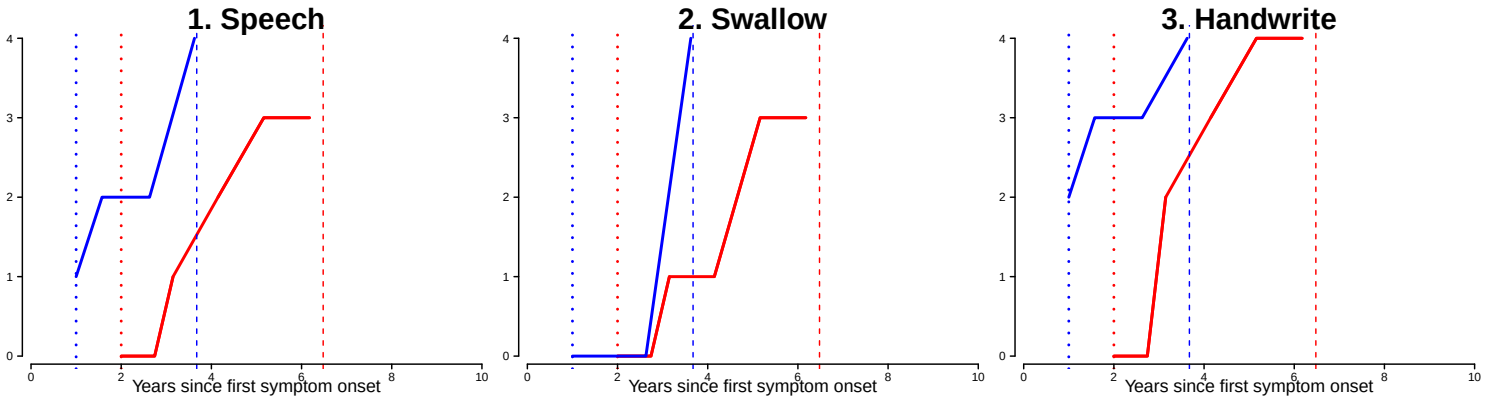
Latent variable

real quantity of interest

- not directly measurable/observable
- latent trajectory underlying items



Other challenges

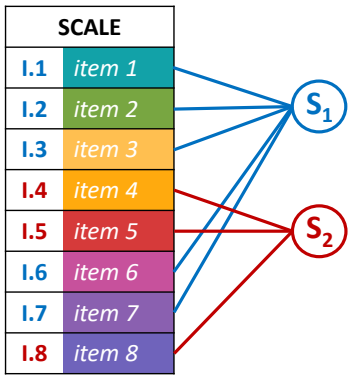


...

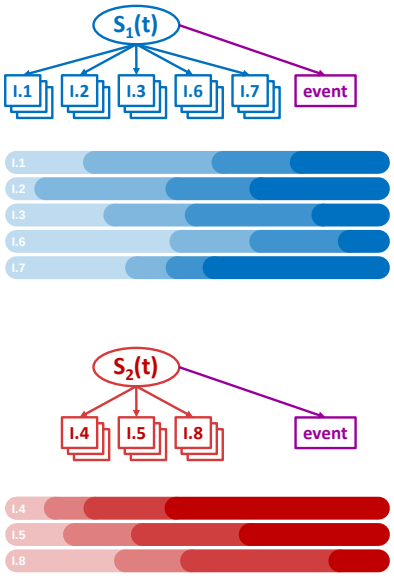
- ordinal items → *nature-adapted model*
- multiple items → *multivariate model*
- repeated measures → *mixed model*
- delayed entry → *left truncation model*
- informative dropout / MNAR data (eg. due to death) → *joint model*
- questionnaire multidimensionality → *strategy of analysis*

The 4S method

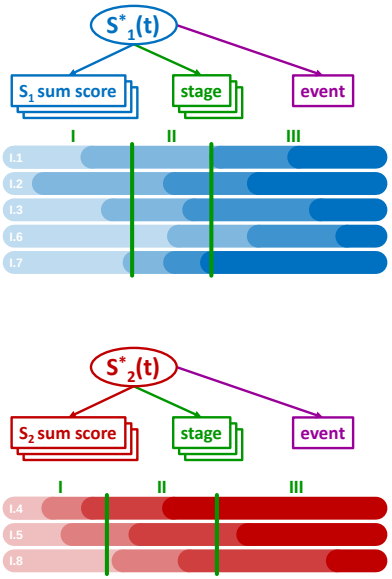
Structuring



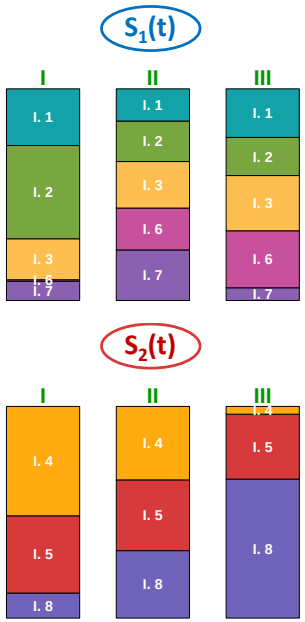
Sequencing



Staging



Selecting

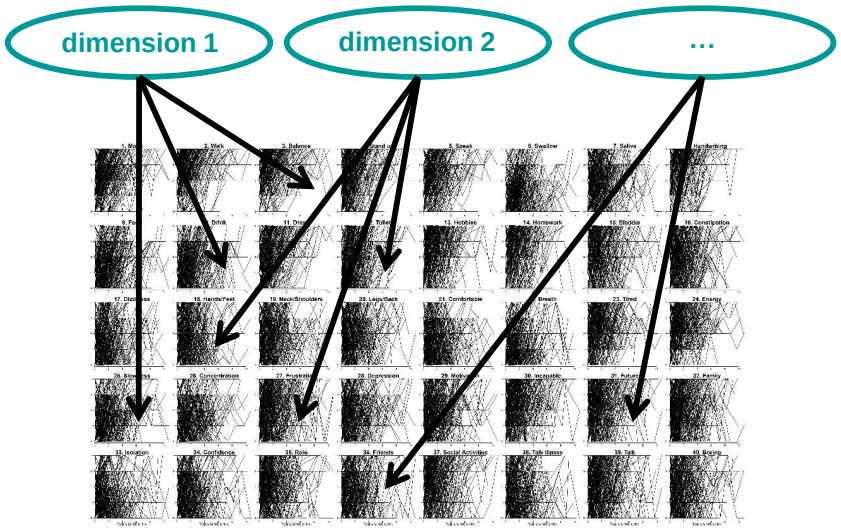


Step 1 - Structuring: Identify questionnaire dimensions

Hypotheses of scale calibration

PROMIS method from *[B.B. Reeve et al., 2007]* for cross-sectional data
+ replication approach to adapt to longitudinal data

- **Unidimensionality**
  items from a same dimension
  measure the same trait
  → *factor analyses*
- **Conditional independance**
  no redundant items
  within the same dimension
  → *residual correlations*
- **Increasing monotonicity**
  the higher the item level,
  the higher the dimension level
  → *item response probabilities*

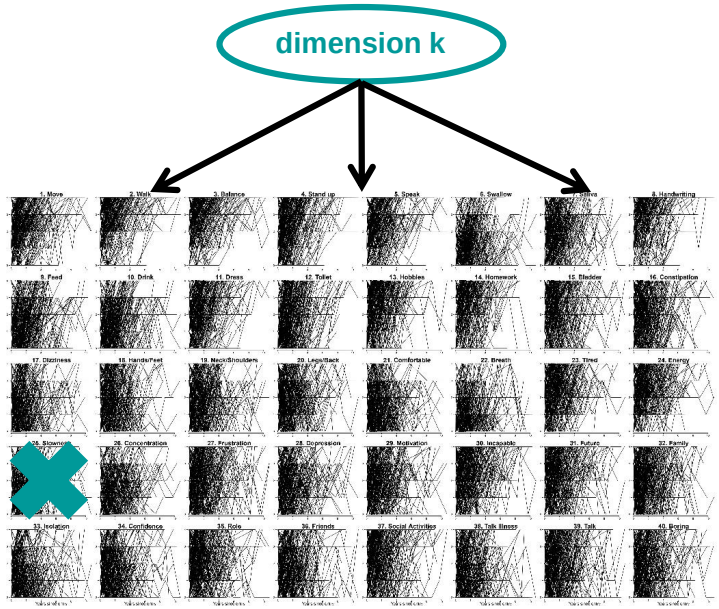


Step 1 - Structuring: Identify questionnaire dimensions

Hypotheses of scale calibration

PROMIS method from *[B.B. Reeve et al., 2007]* for cross-sectional data
+ replication approach to adapt to longitudinal data

- **Unidimensionality**
  items from a same dimension
  measure the same trait
  → *factor analyses*
- **Conditional independance**
  no redundant items
  within the same dimension
  → *residual correlations*
- **Increasing monotonicity**
  the higher the item level,
  the higher the dimension level
  → *item response probabilities*

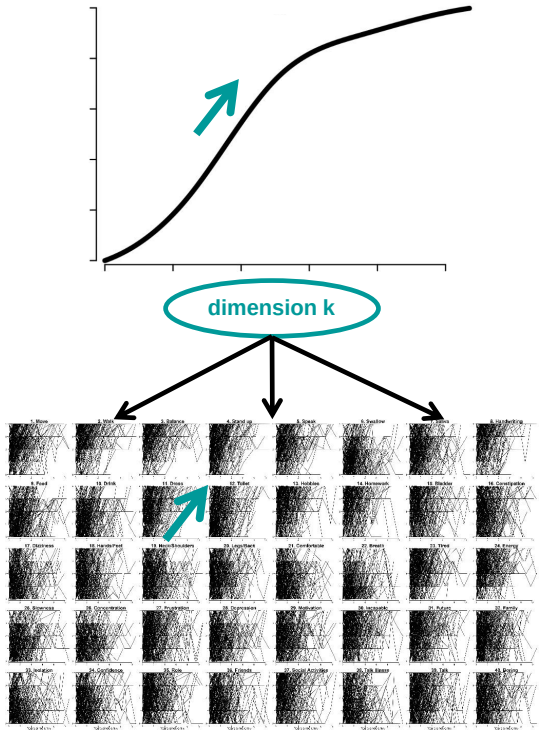


Step 1 - Structuring: Identify questionnaire dimensions

Hypotheses of scale calibration

PROMIS method from *[B.B. Reeve et al., 2007]* for cross-sectional data
+ replication approach to adapt to longitudinal data

- **Unidimensionality**
  items from a same dimension
  measure the same trait
  → *factor analyses*
- **Conditional independance**
  no redundant items
  within the same dimension
  → *residual correlations*
- **Increasing monotonicity**
  the higher the item level,
  the higher the dimension level
  → *item response probabilities*



Step 1 - Structuring: PDQ-39 dimensions

Original PDQ-39 questionnaire	Mobility	1	Leisure activities
		2	Looking after home
		3	Carry shopping bags
		4	Walking half a mile
		5	Walking 100 yards
		6	Getting around the house
		7	Getting around in public
		8	Need company when going out
		9	Worry falling in public
		10	Confined to the house
	Activities in Daily Living	11	Washing
		12	Dressing
		13	Do buttons or shoe laces
		14	Writing clearly
		15	Cutting food
		16	Hold a drink without spilling
	Emotional well-being	17	Depressed
		18	Isolated and lonely
		19	Weepy or tearful
		20	Angry or bitter
		21	Anxious
		22	Worried about the future
	Stigma	23	Felt need to conceal PD
		24	Avoid eating/drinking in public
		25	Embarassed due to PD
		26	Worried people's reactions
	Social support	27	Close relationships
		28	Support from partner
		29	Support from family or friends
	Cognition	30	Unexpectedly fallen asleep
		31	Concentration
		32	Poor memory
		33	Dreams or hallucinations
	Communication	34	Speech
		35	Unable communicate properly
		36	Felt ignored
	Bodily discomfort	37	Painful cramps or spams
		38	Pain in joints or body
		39	Unpleasantly hot or cold

Step 1 - Structuring: PDQ-39 dimensions

Original PDQ-39 questionnaire	Mobility	1 Leisure activities	Mobility	PROMIS-revised PDQ-39 questionnaire
		2 Looking after home		
		3 Carry shopping bags		
		4 Walking half a mile		
		5 Walking 100 yards		
		6 Getting around the house		
		7 Getting around in public		
		8 Need company when going out		
		9 Worry falling in public		
		10 Confined to the house		
	Activities in Daily Living	11 Washing	ADL	
		12 Dressing		
		13 Do buttons or shoe laces		
		14 Writing clearly		
		15 Cutting food		
		16 Hold a drink without spilling		
	Emotional well-being	17 Depressed	Psycho-social	
		18 Isolated and lonely		
		19 Weepy or tearful		
		20 Angry or bitter		
		21 Anxious		
		22 Worried about the future		
	Stigma	23 Felt need to conceal PD	Stigma	
		24 Avoid eating/drinking in public		
		25 Embarassed due to PD		
		26 Worried people's reactions		
	Social support	27 Close relationships	Psycho-social	
		28 Support from partner		
		29 Support from family or friends		
	Cognition	30 Unexpectedly fallen asleep	Cognition/ Communication/ Bodily discomfort	
		31 Concentration		
		32 Poor memory		
		33 Dreams or hallucinations		
	Communication	34 Speech		
		35 Unable communicate properly		
		36 Felt ignored		
	Bodily discomfort	37 Painful cramps or spams		
		38 Pain in joints or body		
		39 Unpleasantly hot or cold		

Results

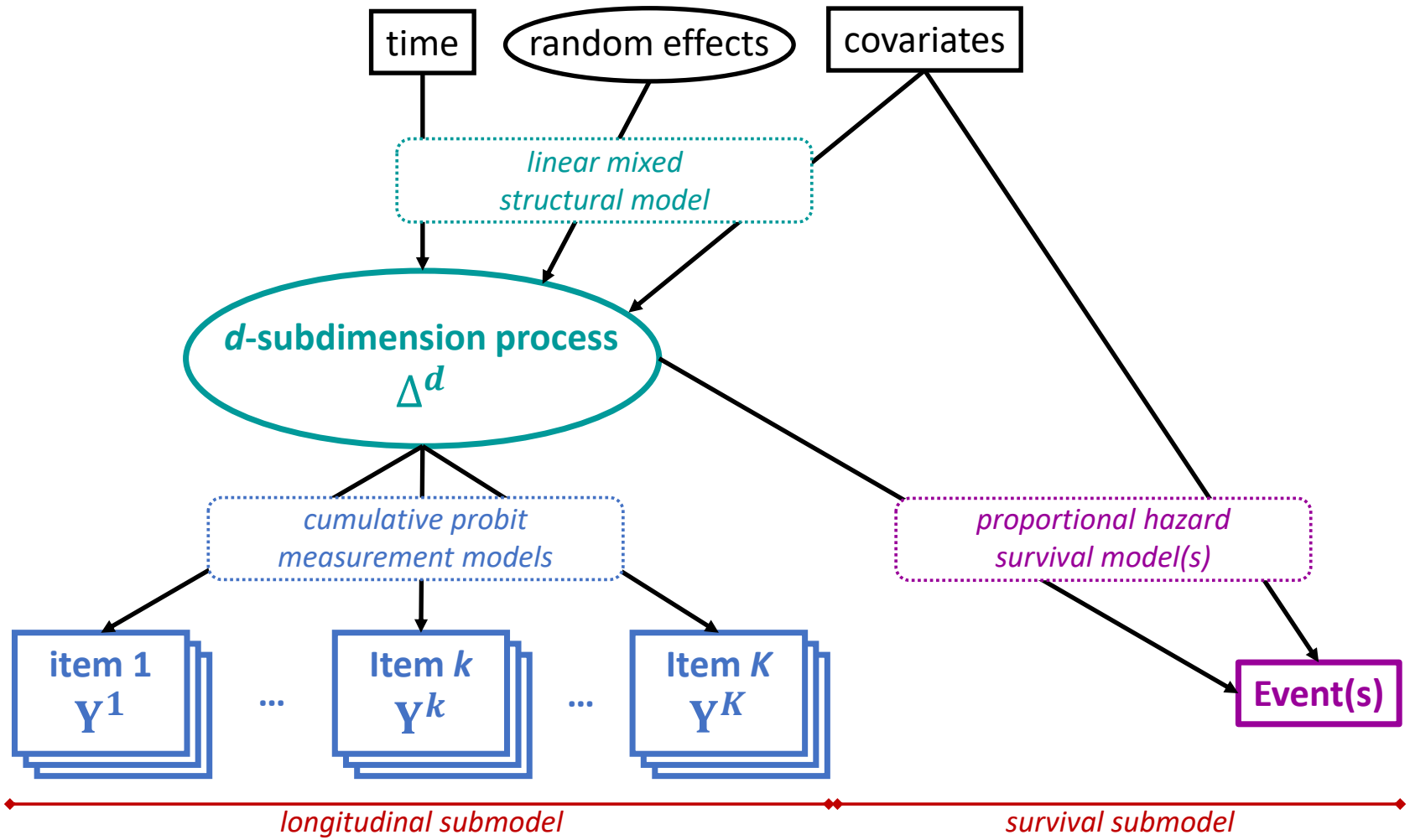
Unidimensionnality
- highlight **5 dimensions** (instead of 8)
- similarities with the original scale

Conditional independance
- satisfying ✓

Unidimensionnality
- satisfying ✓

Step 2 - Sequencing : Model each dimension progression

R-package **JLPM** (for *Joint Latent Process Model*)



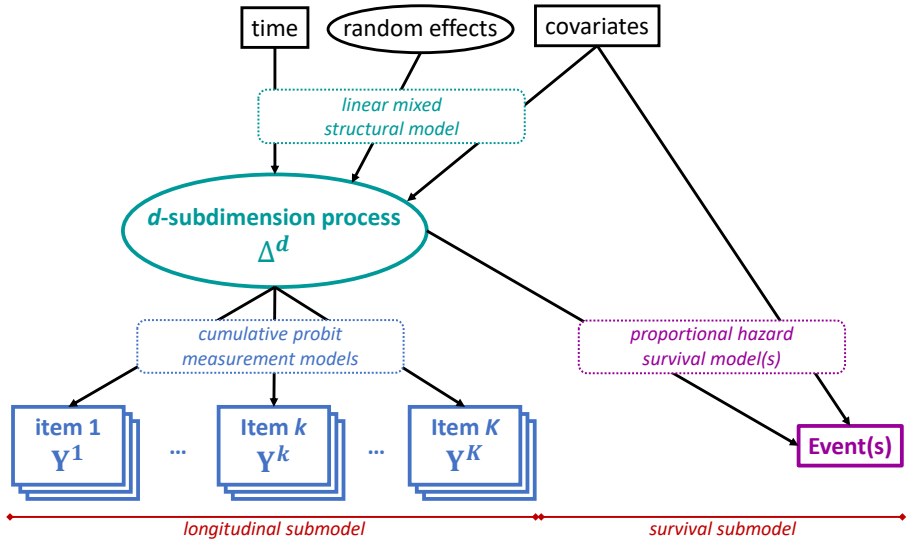
Step 2 - Sequencing : Structure of the **JLPM** model

Dimension process

linear mixed model, dimension d

$$\Delta_i^d(t) = X_i^\top(t)\beta^d + Z_i^\top(t)b_i^d$$

with $b_i^d \sim \mathcal{N}(0, B^d)$



Items

cumulative probit models, item $k \in \mathcal{K}^d$

$$Y_i^k(t_{ij}) = m \quad \Leftrightarrow \quad \delta_{k,m} < \Delta_i^d(t_{ij}) + \epsilon_{ij}^k \leq \delta_{k,m+1}$$

with $\epsilon_i^k \sim \mathcal{N}(0, \sigma_k^2)$; $m \in \{1, ..., M_k\}$;

$$-\infty = \eta_{k,1} \leq ... \leq \eta_{k,m} \leq \eta_{k,m+1} \leq ... \leq \eta_{k,M_k+1} = +\infty$$

Survival part

proportional hazard model(s), cause p

$$\lambda_i^{dp}(t) = \lambda_0^{dp}(t; \xi^{dp})\ exp(W_i^{dp\top}\gamma^{dp} + g^{dp}(b_i^d, t)^\top \alpha^{dp})$$

Step 2 - Sequencing : Estimation of the **JLPM** model

Vector of parameters $\Theta^d = (\underbrace{\beta^d, B^d}_{\text{structural model}}, \underbrace{(\delta_{k,m})_{k,m}, (\sigma_k)_k}_{\text{measurement models}}, \underbrace{(\xi^{dp})_p, (\gamma^{dp})_p, (\alpha^{dp})_p}_{\text{survival model(s)}})^\top$

Maximum log-likelihood estimation

$$\hat{\Theta} = \underset{\Theta}{\text{argmax}} \ \log \mathrm{L}(\Theta)$$

with the model **likelihood**

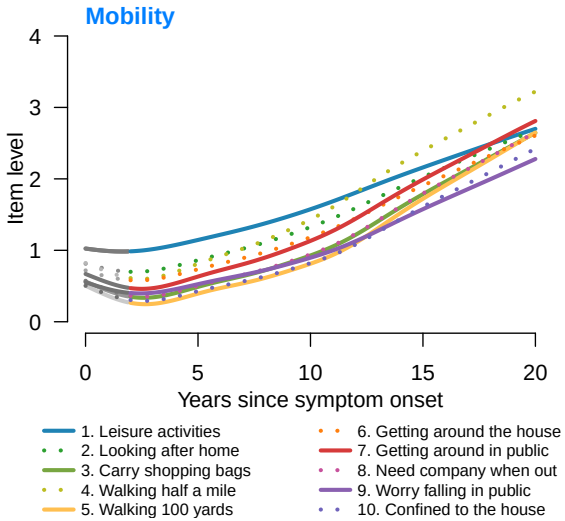
$$\mathrm{L}(\Theta) = \prod_{i=1}^{N} \mathrm{L}_i(\Theta)$$

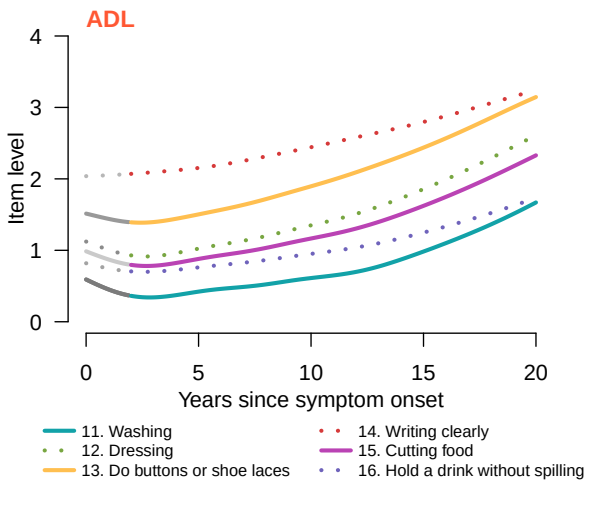
and the **individual contribution**

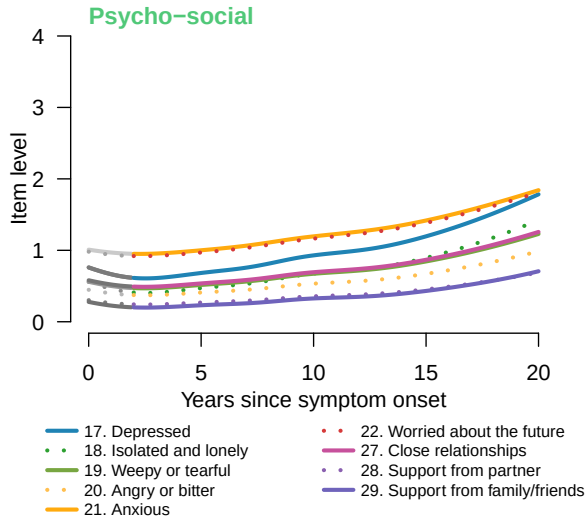
$$\begin{aligned} \mathrm{L}_i(\Theta) &= f_{Y_i,T_i}(Y_i, T_i\,; \Theta) \\ &= \int_{\mathbb{R}^q} f_{Y_i|b_i}(Y_i|b\,; \Theta)\, f_{T_i|b_i}(T_i|b\,; \Theta)\, f_{b_i}(b\,; \Theta)\, db \end{aligned}$$

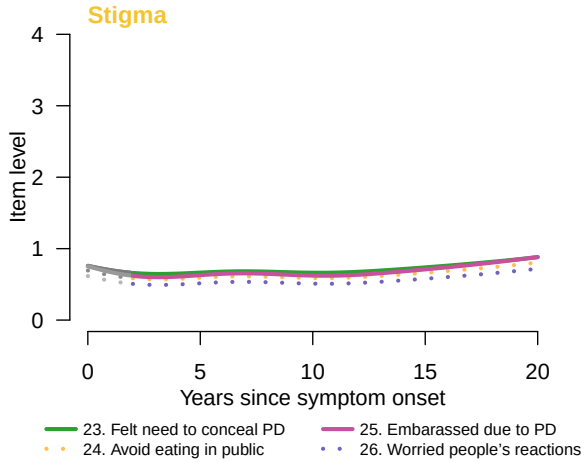
using Marquart-Levenberg **algorithm for optimization**, and Quasi-Monte-Carlo and Gauss-Kronrod **algorithms for integral approximation**

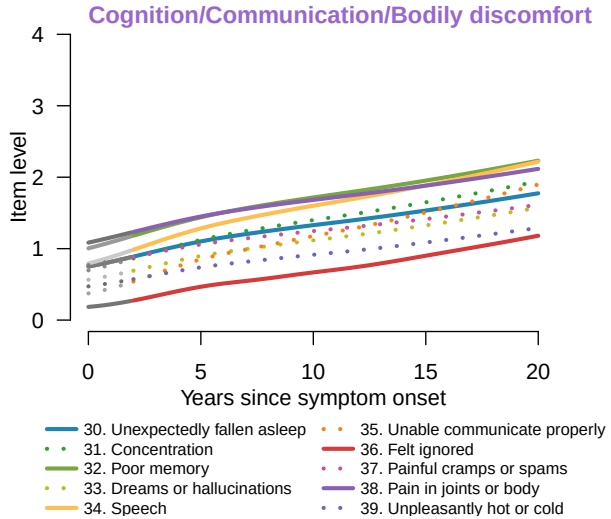
Step 2 - Sequencing : Progression of PDQ-39 dimensions over time



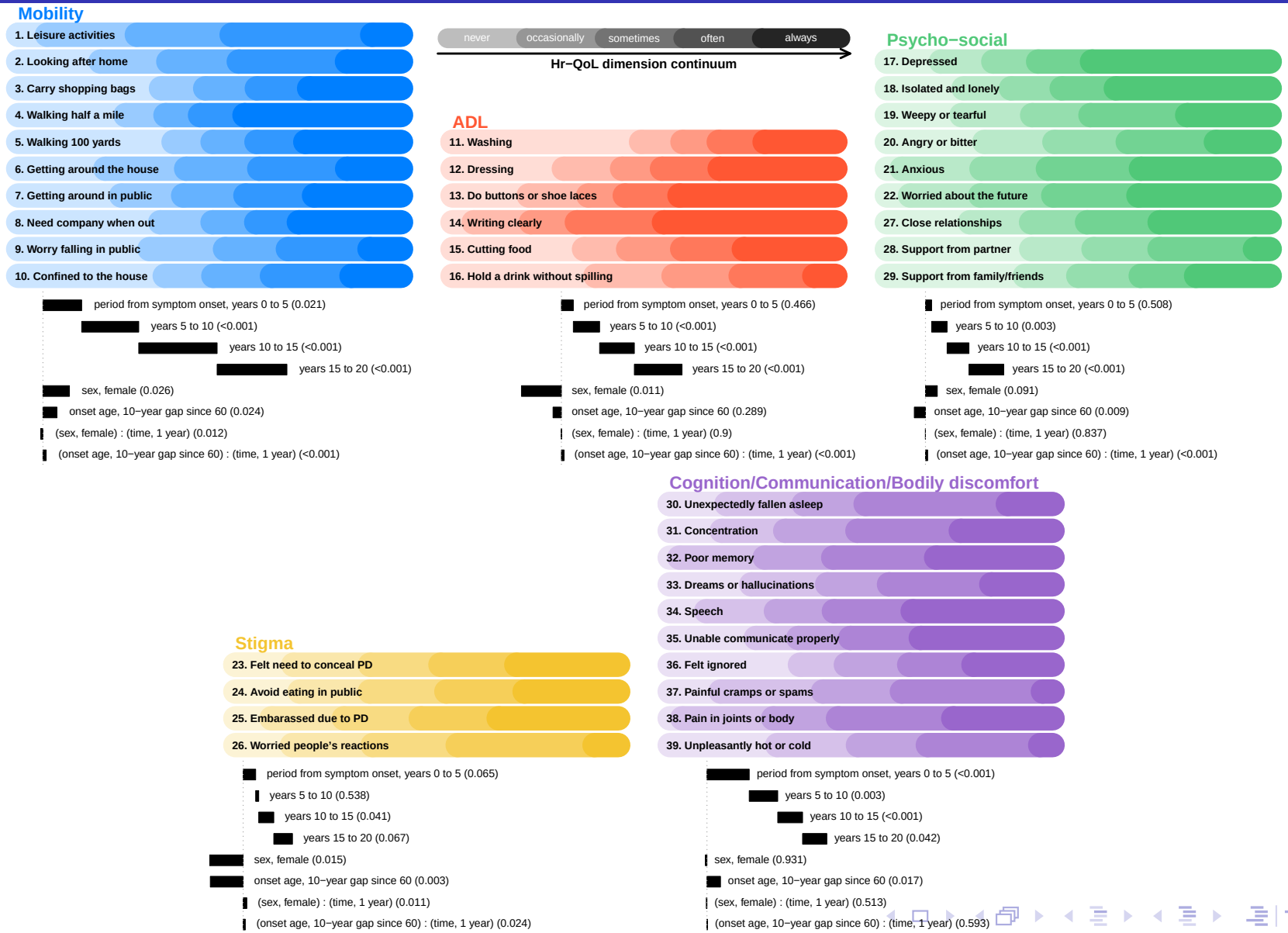






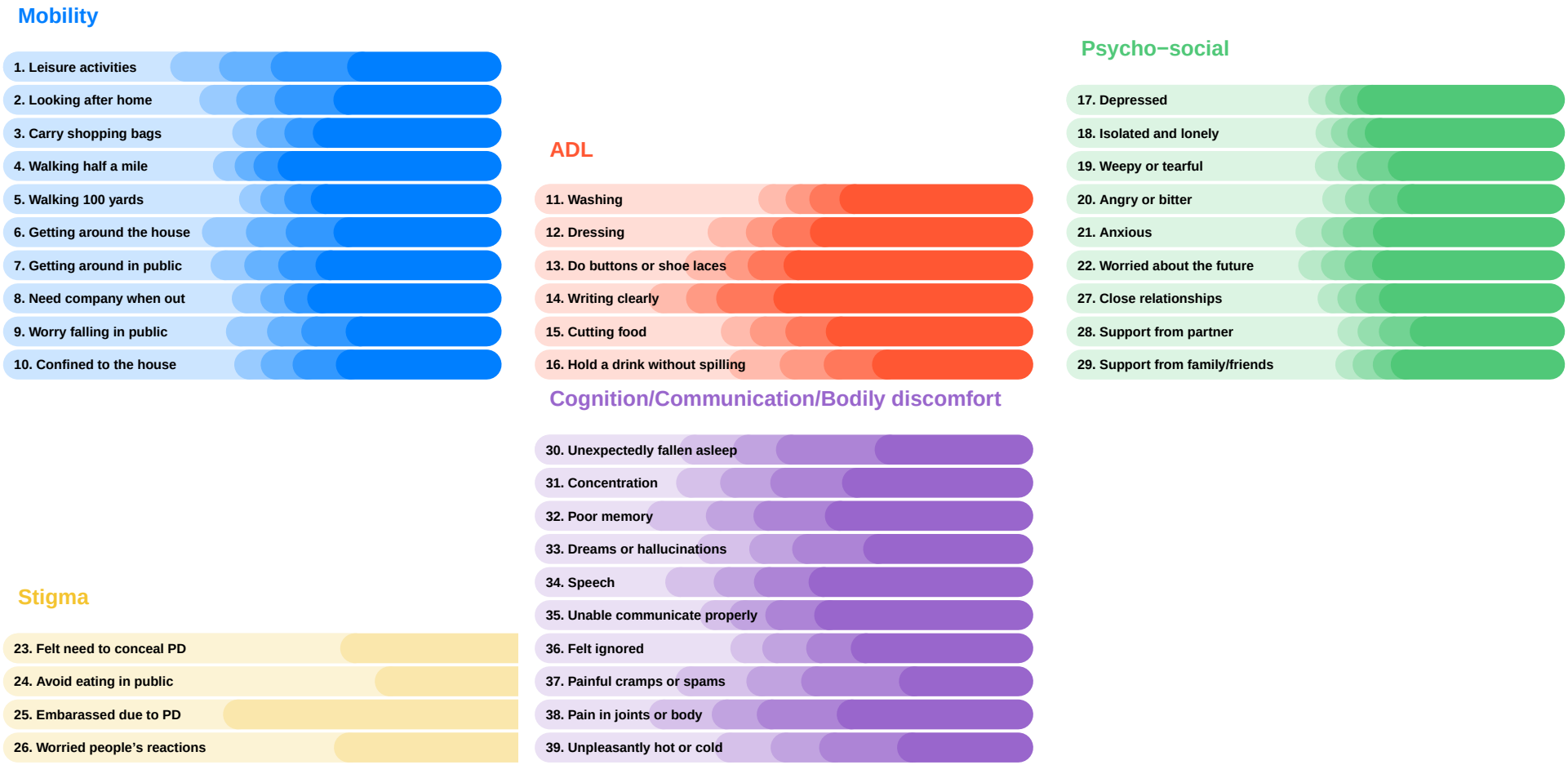


Step 2 - Sequencing : PDQ-39 sequences of impairment

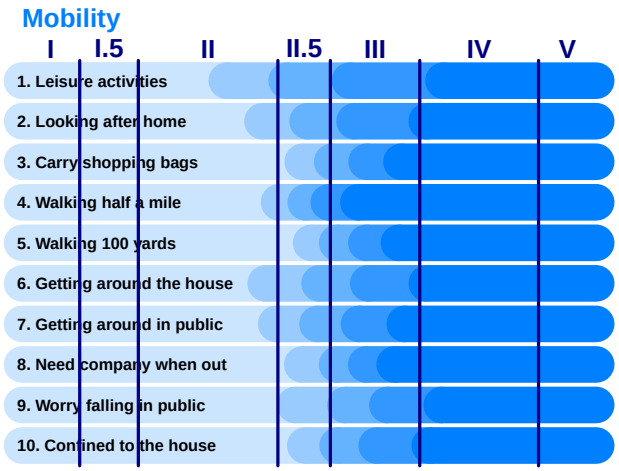


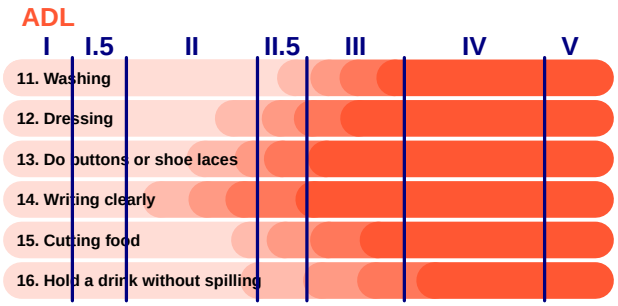
Step 2 - Sequencing : PDQ-39 sequences of impairment

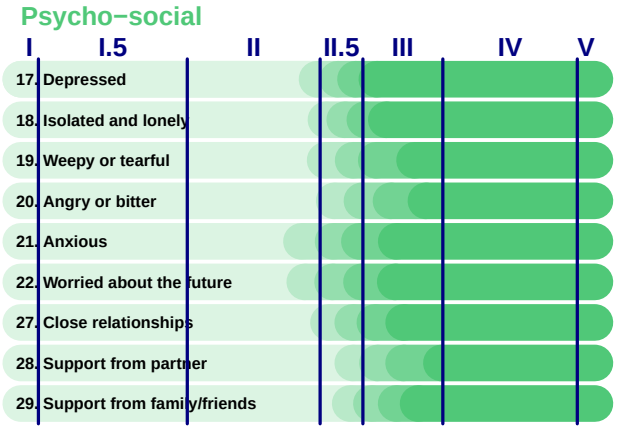
Step 3 - Staging: PDQ-39 progression

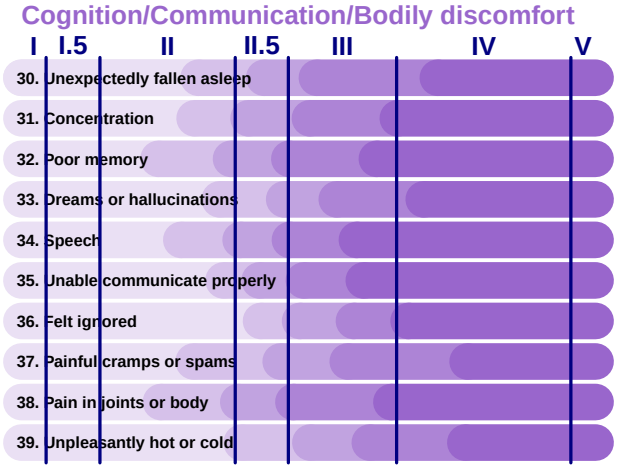


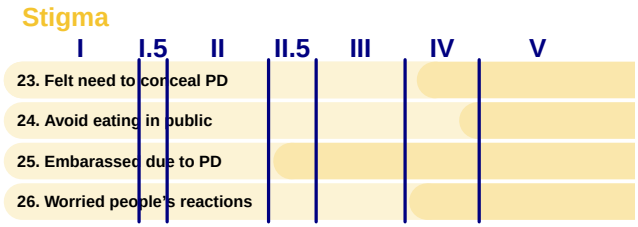
Step 3 - Staging: PDQ-39 progression according to PD stages











Hoehn & Yahr scores

I	Unilateral involvement only
I.5	Unilateral and axial involvement
II	Bilateral involvement without balance impairment
II.5	Mild bilateral disease with recovery on pull test
III	Mild to moderate bilateral disease ; some postural instability ; physically independent
IV	Severe disability ; still able to walk or stand unassisted
V	Wheelchair bound or bedridden unless aided

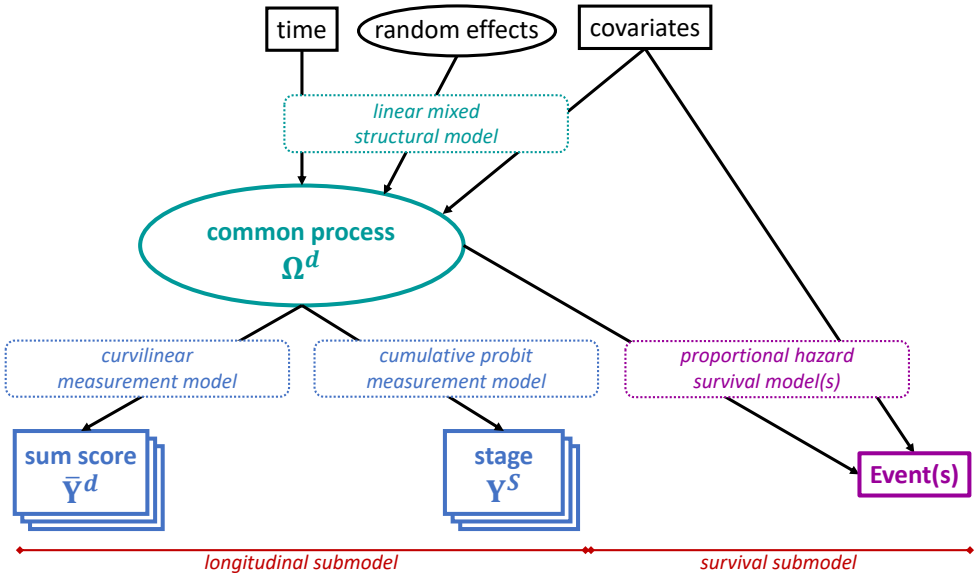
Step 3 - Staging: Link dimension progressions to disease stages

Projection approch

determine thresholds δ_s^d in Δ^d latent scale corresponding to stage s transition

1. Bivariate JLPM

btw dimension score $\bar{Y}^d = \sum_{k \in \mathcal{K}^d} Y^k$ and disease stage Y^S

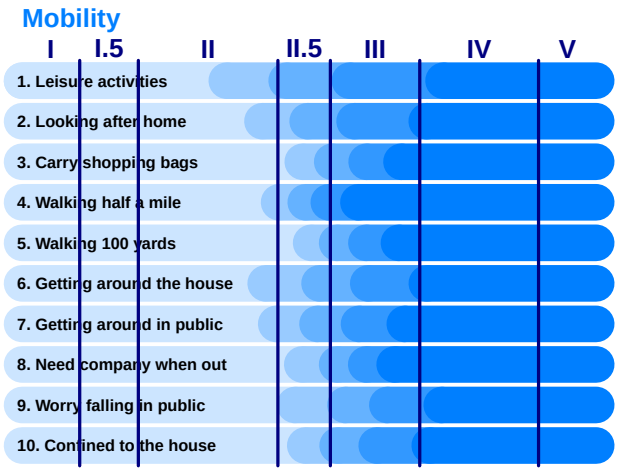


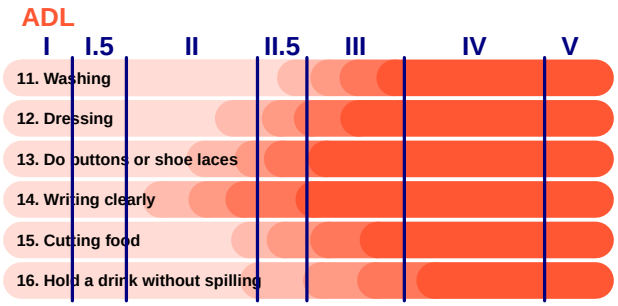
$$\begin{cases} \Omega_i^d(t) = X_i(t)^\top \mu^d + Z_i(t)^\top v_i^d \\ \text{linear mixed model} \\ \bar{Y}_i^d(t_{ij}) = H_d^{-1}\left(\Omega_i^d(t_{ij}) + \varepsilon_{ij}^{\bar{Y}d}\right) \\ \text{curvilinear model} \\ Y_i^S(t_{ij}) = s \Leftrightarrow \omega_s^d < \Omega_i^d(t_{ij}) + \varepsilon_{ij}^{Y^S d} \leq \omega_{s+1}^d \\ \text{cumulative probit model} \\ \zeta_i^{dp}(t) = \zeta_0^{dp}(t;\psi^d)\exp\left(W_i^{dp\top}\kappa^{dp} + g^{dp}(v_i^d,t)^\top \tau^{dp}\right) \\ \text{risk model for cause } p \end{cases}$$

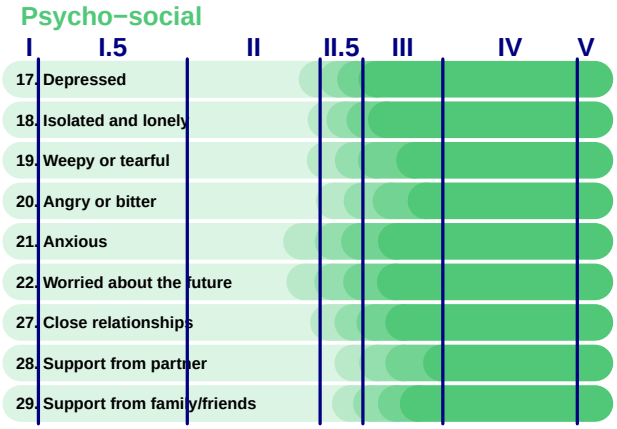
2. Matching the sum of the items $\sum_{k \in \mathcal{K}^d} Y^k$ and the score $\bar{Y}^d$

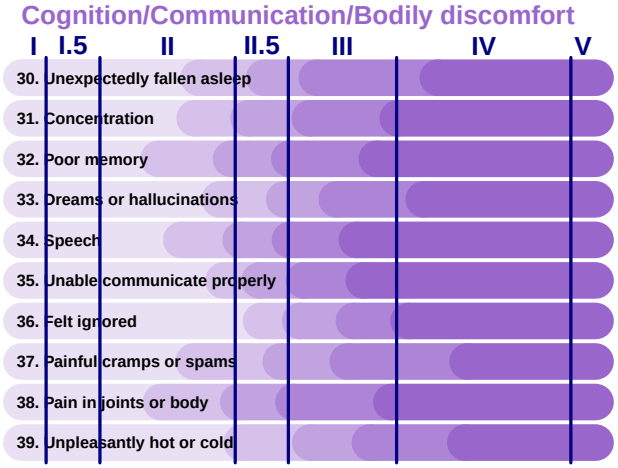
$$\mathbb{E}\left[\sum_{k \in \mathcal{K}^d} Y^k \mid \Delta^d = \hat{\delta}_s^d\right] = \mathbb{E}\left[\bar{Y}^d \mid \Omega^d = \hat{\omega}_s^d\right]$$

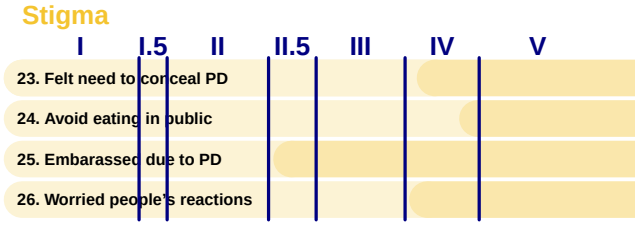
Step 3 - Staging: PDQ-39 progression according to PD stages











Hoehn & Yahr scores

I	Unilateral involvement only
I.5	Unilateral and axial involvement
II	Bilateral involvement without balance impairment
II.5	Mild bilateral disease with recovery on pull test
III	Mild to moderate bilateral disease ; some postural instability ; physically independent
IV	Severe disability ; still able to walk or stand unassisted
V	Wheelchair bound or bedridden unless aided

Step 4 - Selecting : Quantify each item information during stages

Fisher information function of item k for dimension level Δ

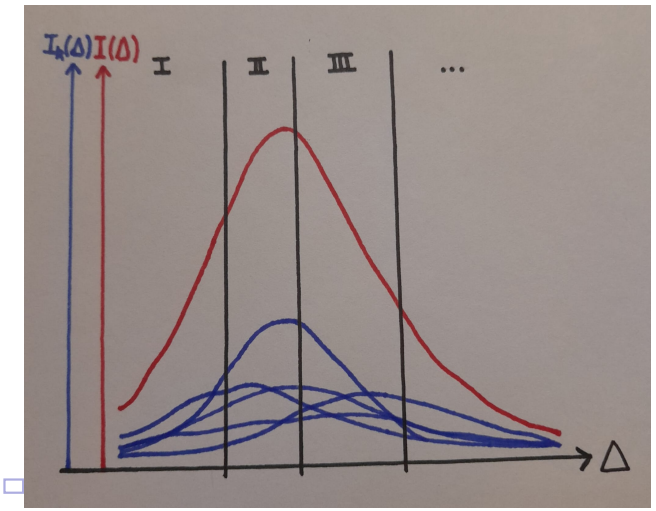
$$\begin{aligned} I_k(\Delta) &= -\mathbb{E}[\frac{\partial^2}{\partial \Delta^2} \log \mathrm{P}_k(\Delta)] \\ &= -\sum_{m=0}^{M_k} \mathbb{P}(Y_k = m \mid \Delta) \frac{\partial^2}{\partial \Delta^2} \log \mathbb{P}(Y_k = m \mid \Delta) \end{aligned}$$

Information carried by item k during stage s

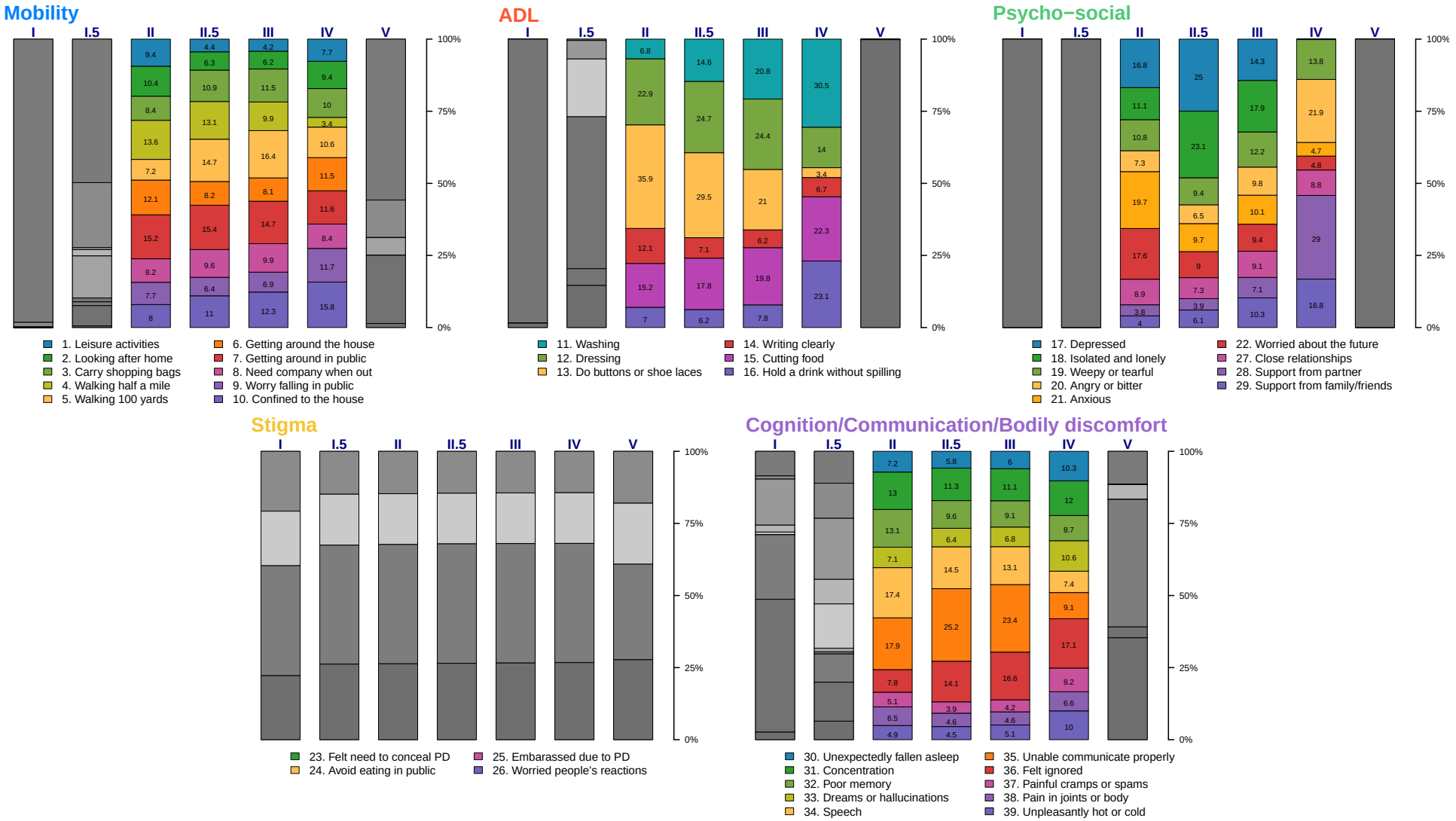
$$I_{k,s} = \int_{\hat{\delta}_s^d}^{\hat{\delta}_{s+1}^d} I_k(\Delta) d\Delta$$

Pourcentage of information carried by item k during stage s

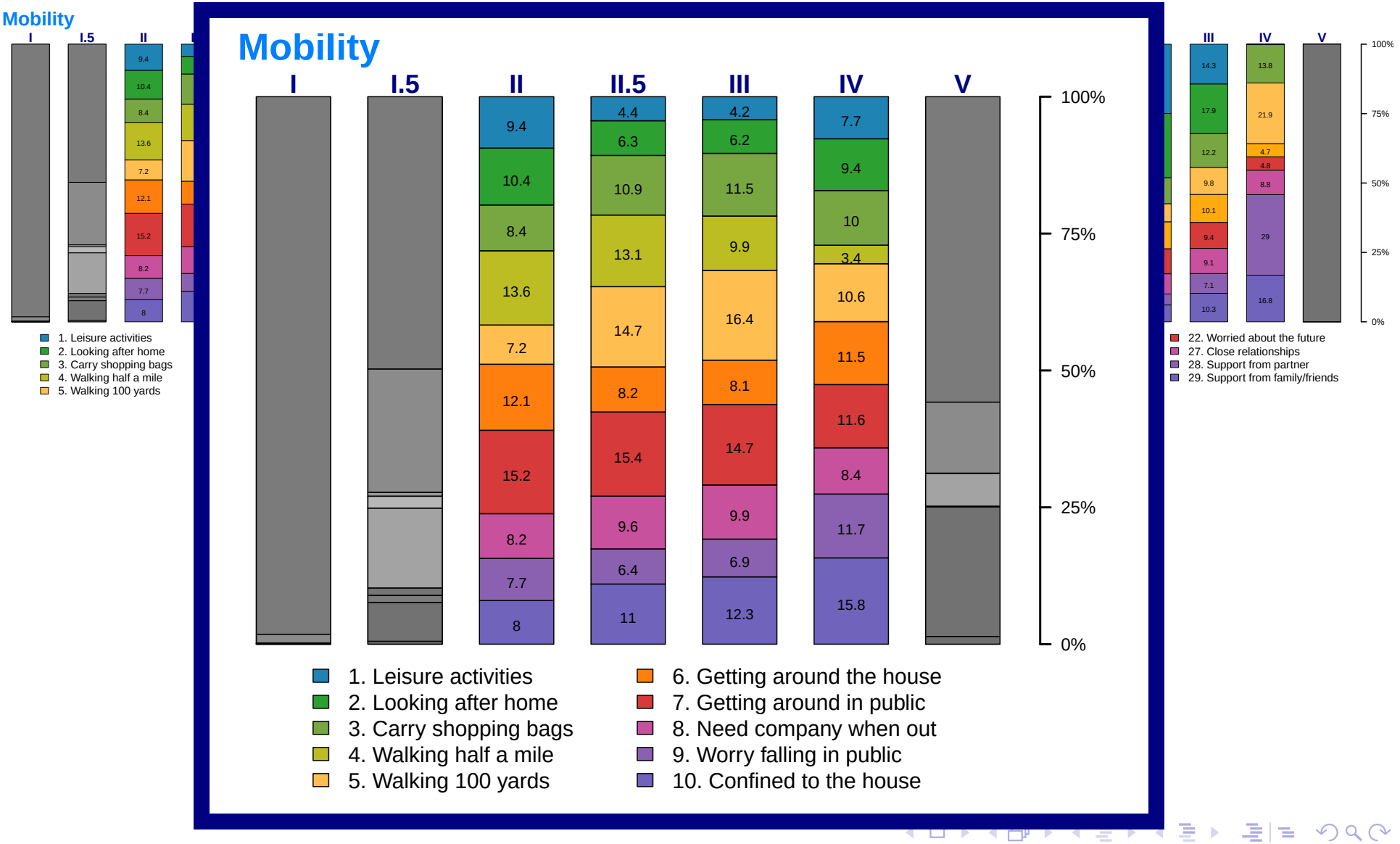
$$I_{k,s}^{\%} = \frac{I_{k,s}}{\sum_{l \in \mathcal{K}^d} I_{l,s}}$$



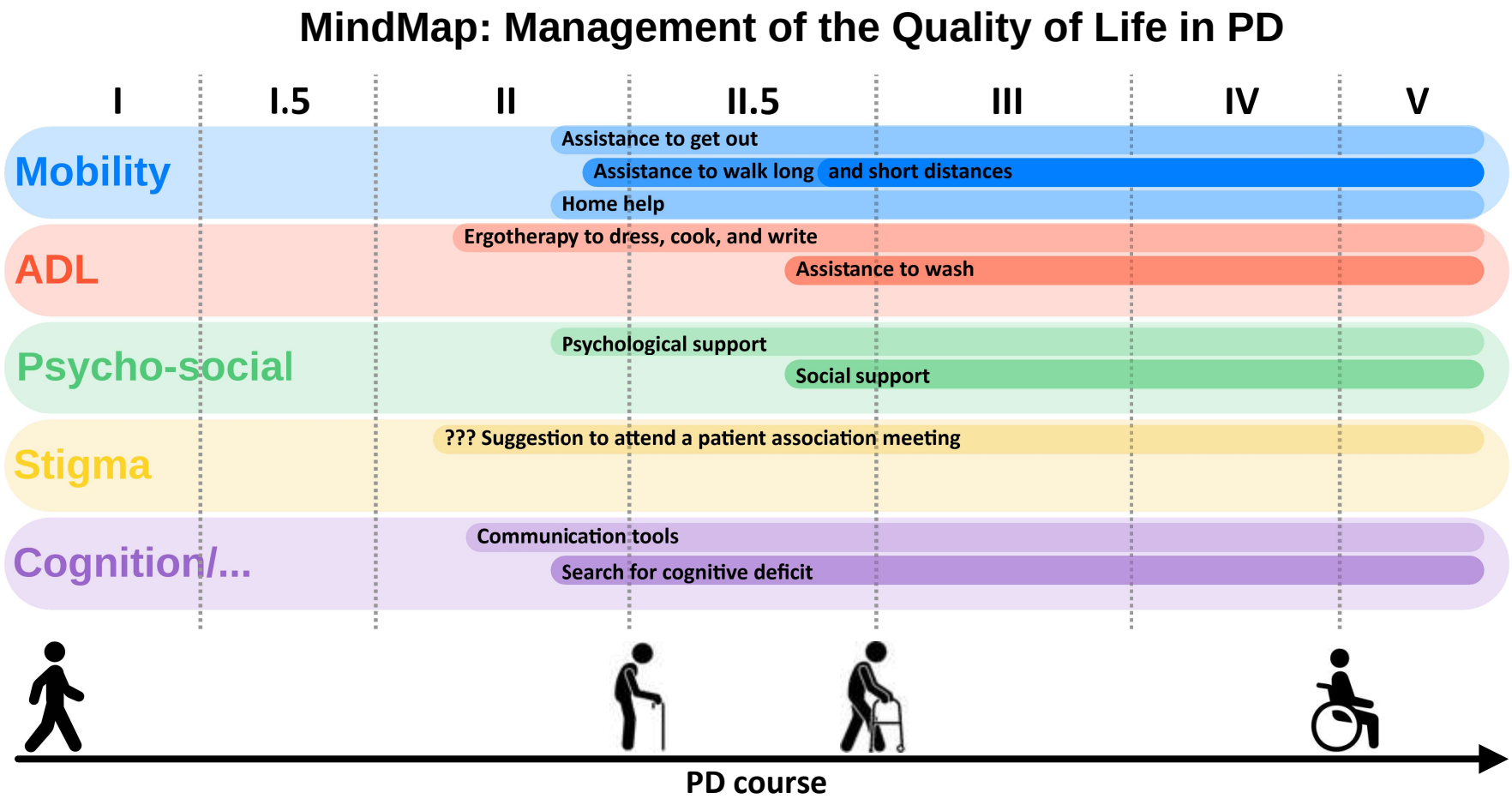
Step 4 - Selecting : PDQ-39 most informative items during PD course



Step 4 - Selecting : PDQ-39 most informative items during PD course



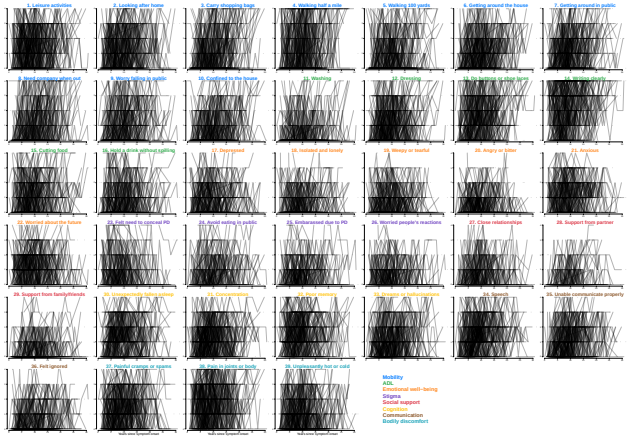
4S method - final output : Monitoring mind map draft

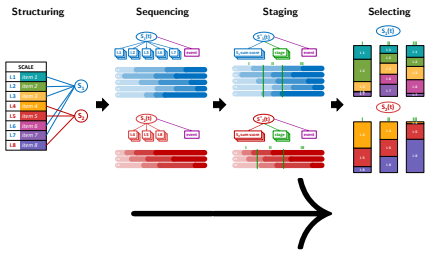


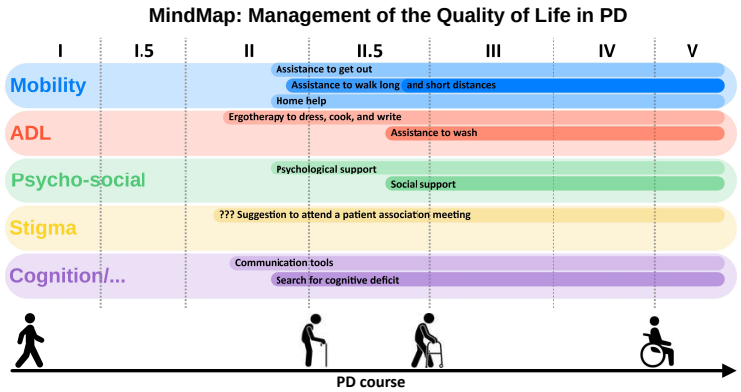
Conclusion

4S method = complete strategy to analyze questionnaire data
- based on a joint model adapted to **ordinal items**
- accounting for potential **informative dropout**
- treating **multidimensionality**

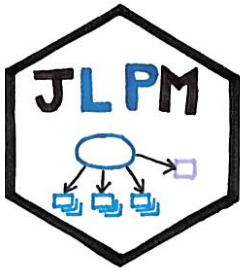
Study of Parkinson's progression from patient perspective
Useful to **other pathologies and questionnaires' studies**







Scientific contributions



T. Saulnier et al., *Joint models for the longitudinal analysis of measurement scales in the presence of informative dropout*, **Methods** (2022)

package `JLPM`, available on CRAN *in collaboration with Viviane Philipps*

T. Saulnier et al., *Structuring, Sequencing, Staging, Selecting: the 4S method for the longitudinal analysis of multidimensional questionnaires in chronic diseases*, **Biometrics Practice** (2025) (*preprint arXiv:2407.08278*)

T. Saulnier et al., *Patient-perceived progression in multiple system atrophy: natural history of quality of life*, **Journal of Neurology, Neurosurgery & Psychiatry** (2024)

replication script, available on GitHub
https://github.com/VivianePhilipps/JLPM/blob/main/vignettes/script_4Smethod.Rmd

Thank you for your attention!

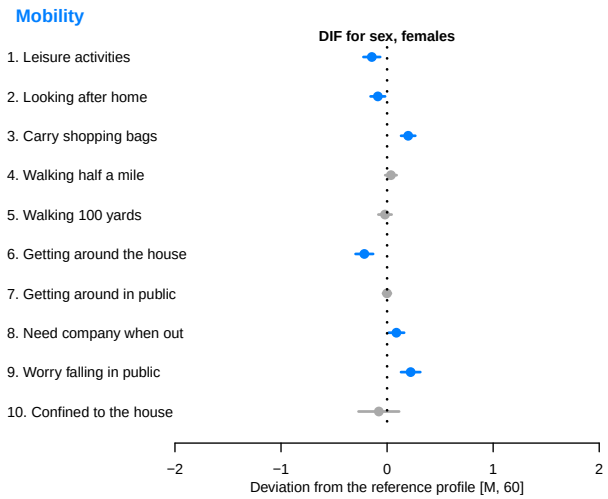
Other references

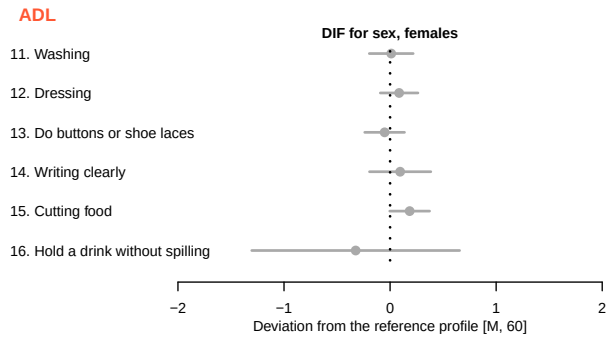
- Su *et al.* (2025) Projections for prevalence of Parkinson's disease and its driving factors in 195 countries and territories to 2050: modelling study of Global Burden of Disease Study 2021. BMJ, 388.
- Testa *et al.* (1996) Assessment of quality-of-life outcomes. New England journal of medicine, 334(13), 835-840.
- Peto *et al.* (1998) PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. Journal of neurology, 245(Suppl 1), S10-S14.
- MacAskill *et al.* (2023) The New Zealand Parkinson's progression programme. Journal of the Royal Society of New Zealand, 53(4), 466-488.
- Reeve *et al.* (2007) Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical care, 45(5), S22-S31.
- Philipps *et al.* (2020) Robust and efficient optimization using a Marquardt-Levenberg algorithm with R package marqLevAlg. R Journal.

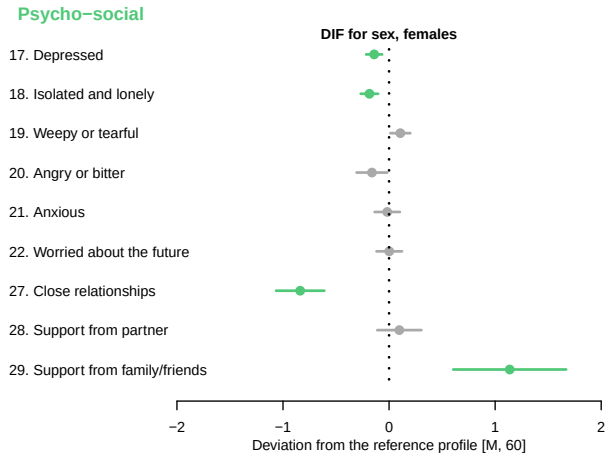
Acknowledgments

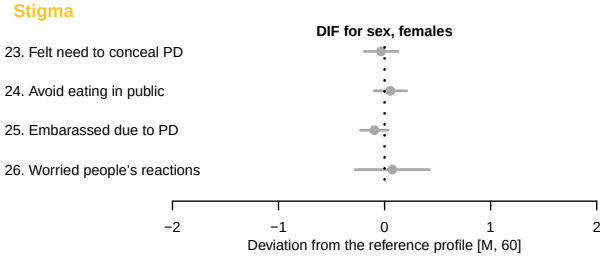
New Zealand Brain Research Institute
université de BORDEAUX
IMN BORDEAUX neurocampus
FRANCE PARKINSON
IBS

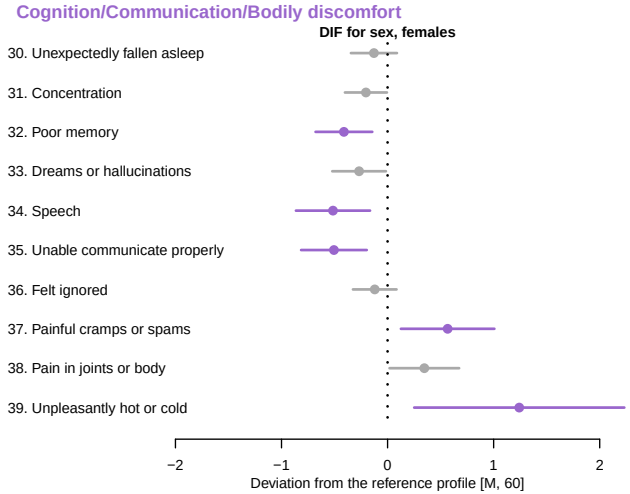
Step 2 - Sequencing : PDQ-39 detected DIF for sex











Step 2 - Sequencing : PDQ-39 detected DIF for onset age