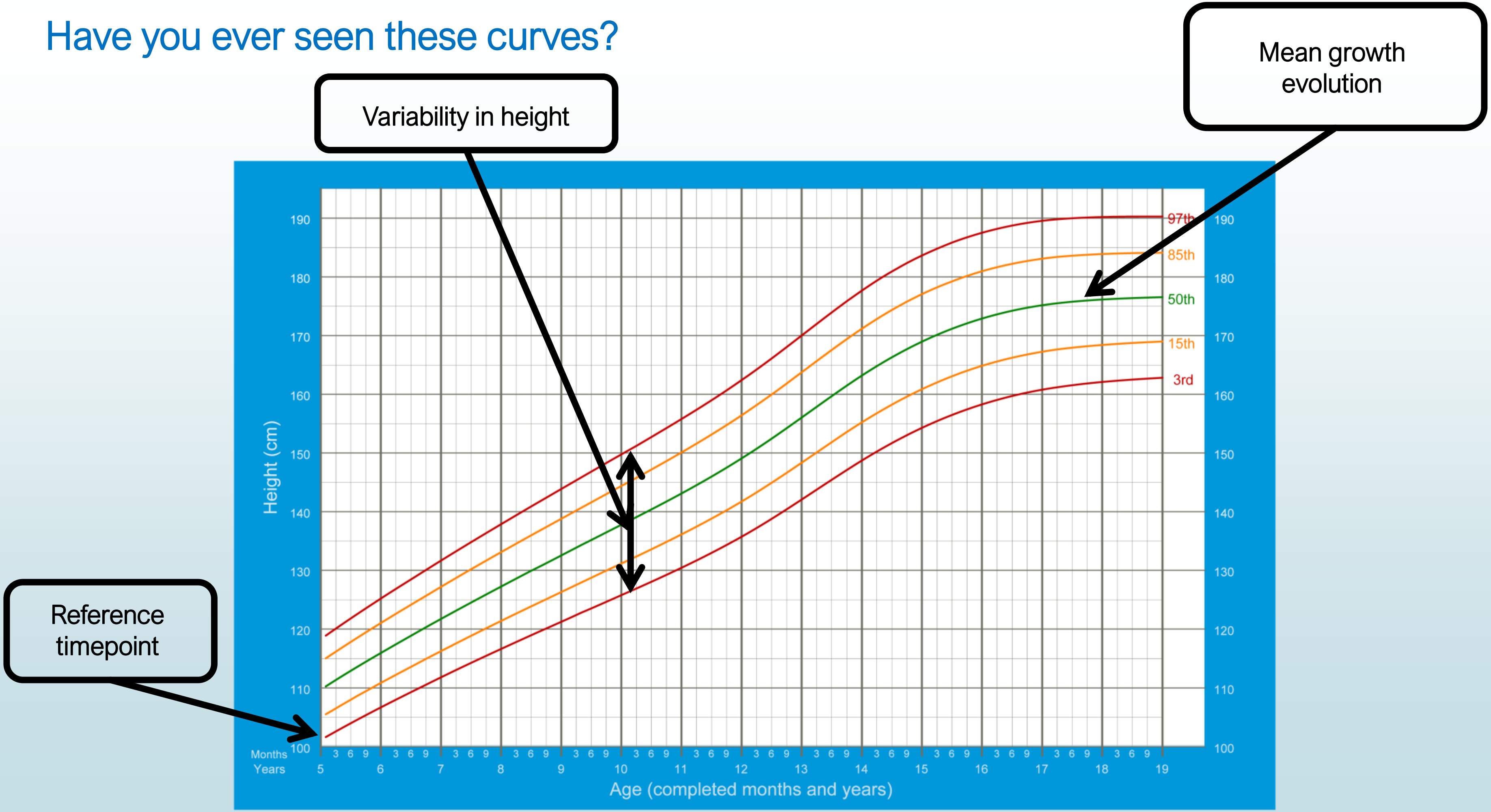
Mixed effect models for the spatio-temporal analysis of manifold valued data: application to Alzheimer's disease screening

Stéphanie Allassonnière

Professor and Vice-President, Paris Cité University

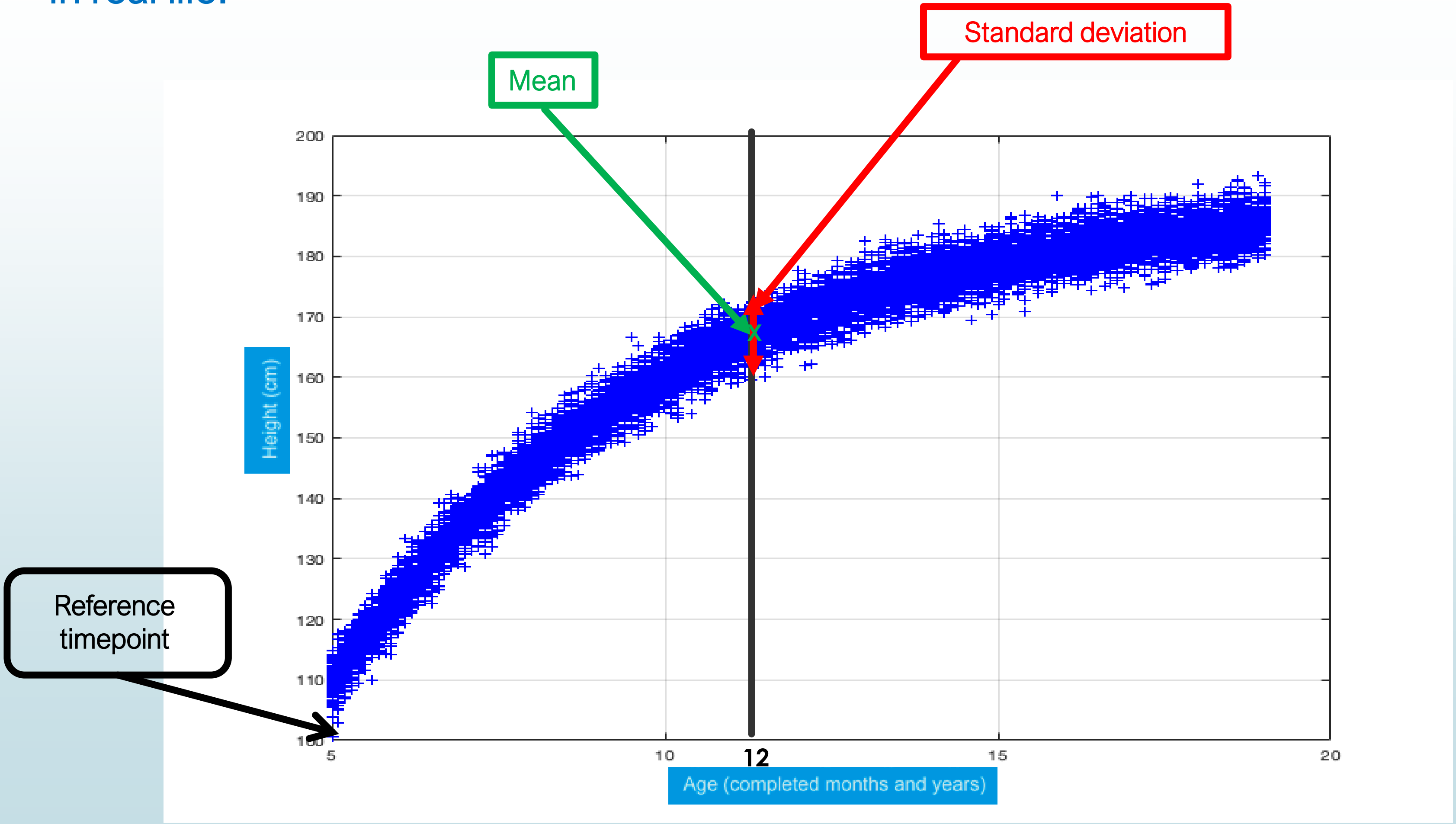
Fellow and Deputy Director of the PR[AI]RIE Institute

Have you ever seen these curves?



2007 WHO Reference

In real life:

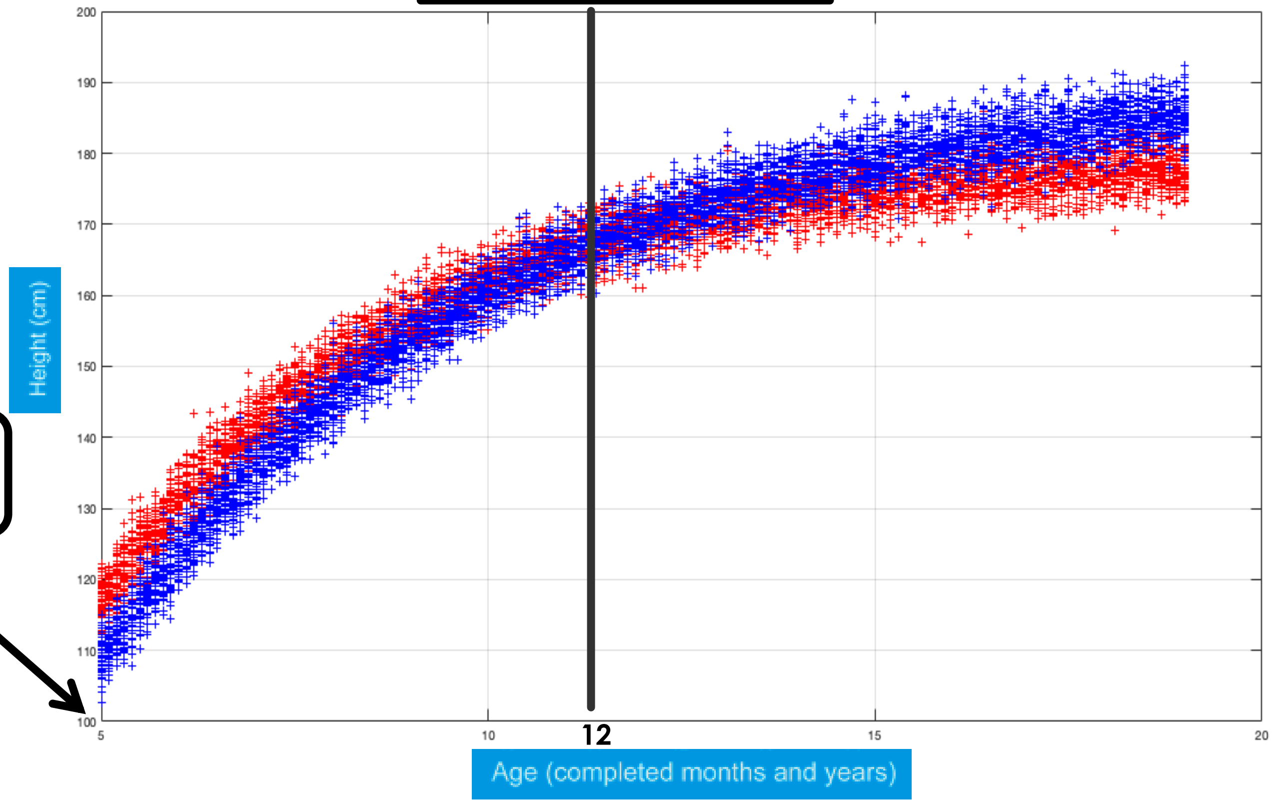

Standard deviation
Mean
Reference
timepoint
200
190
180
170
160
150
140
130
120
110
5
10
12
15
20
Height (cm)
Age (completed months and years)

And even more:

Mean

Standard deviation

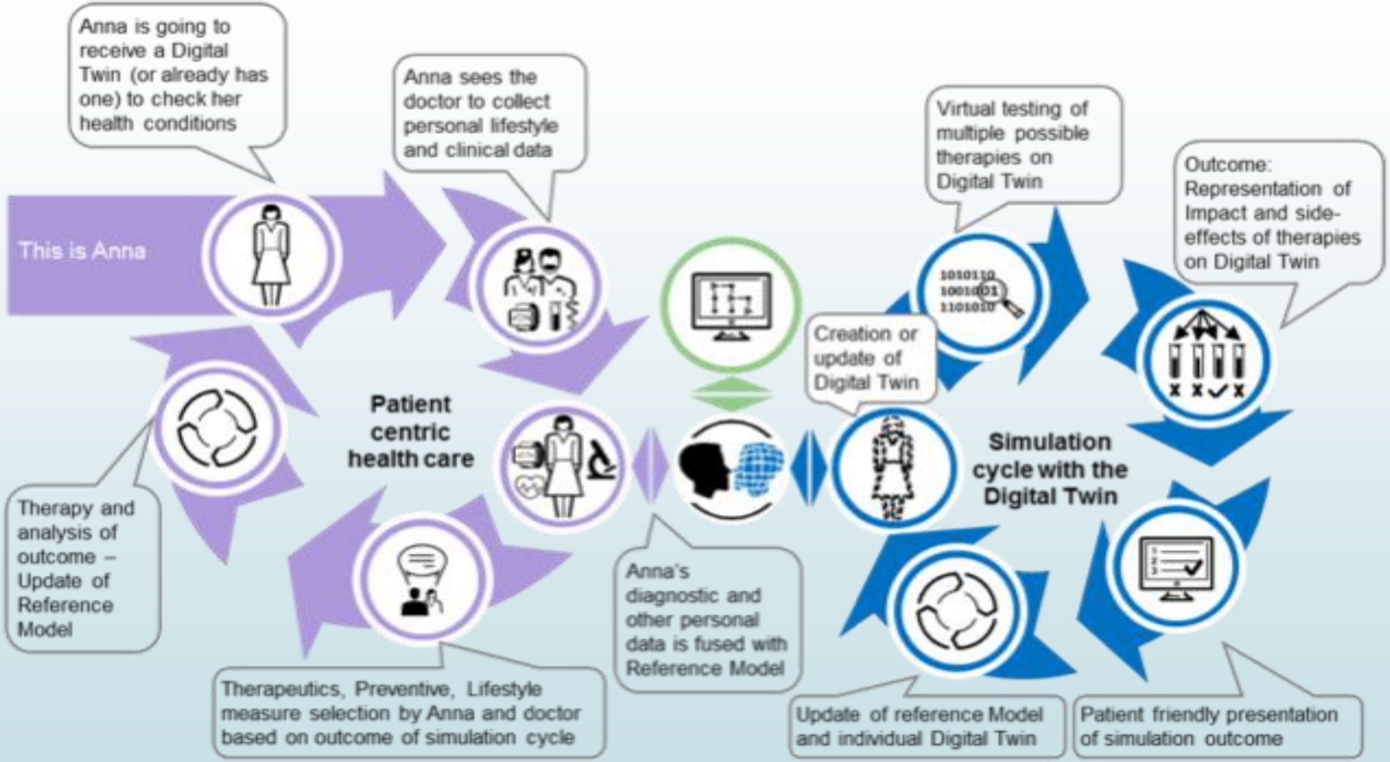
Not that well defined



Reference timepoint

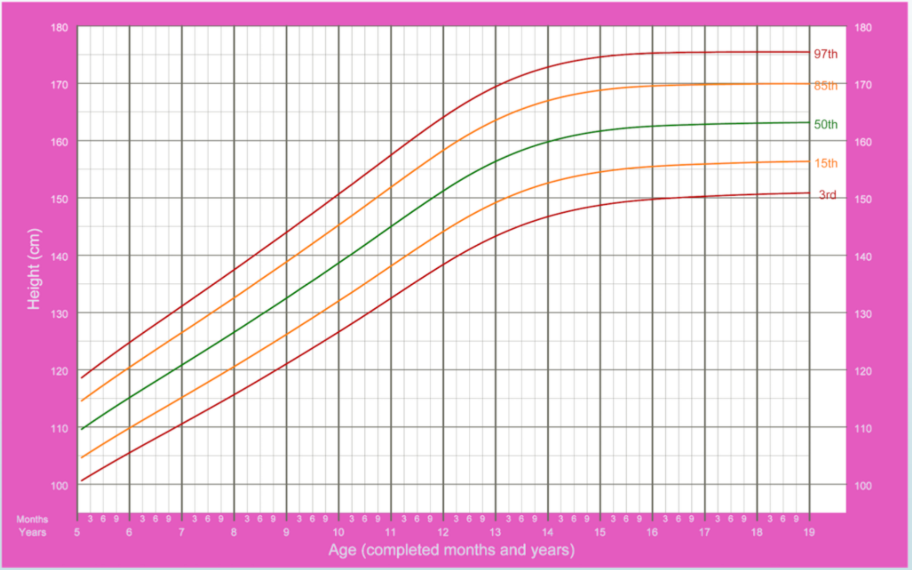
Le jumeau numérique

LE JUMEAU NUMÉRIQUE :



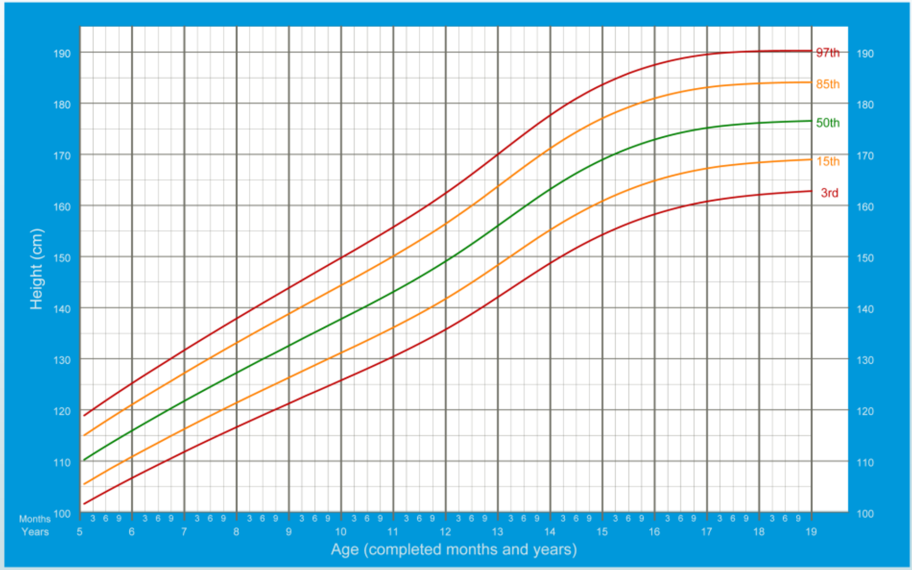
Existance of clusters:

Height-for-age GIRLS
5 to 19 years (percentiles)



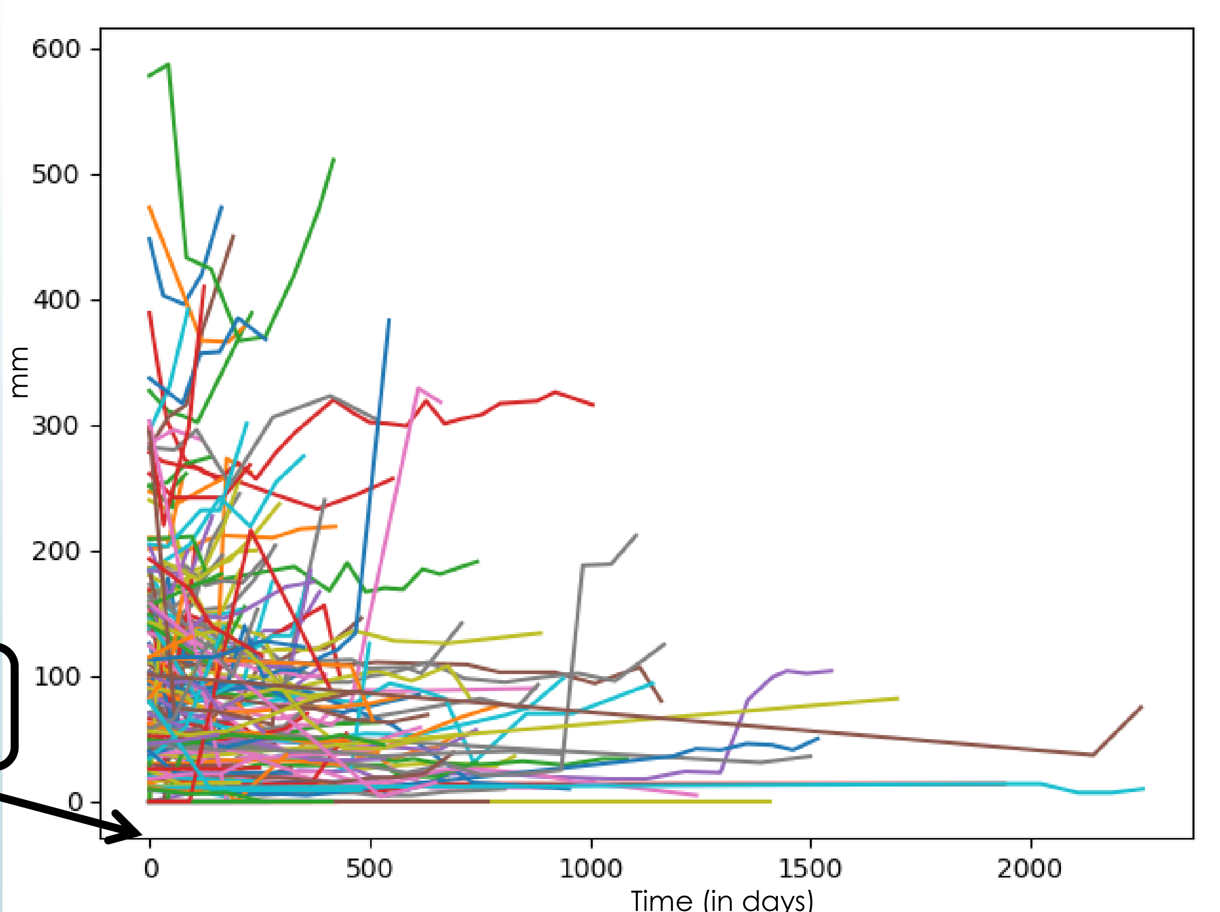
2007 WHO Reference

Height-for-age BOYS
5 to 19 years (percentiles)



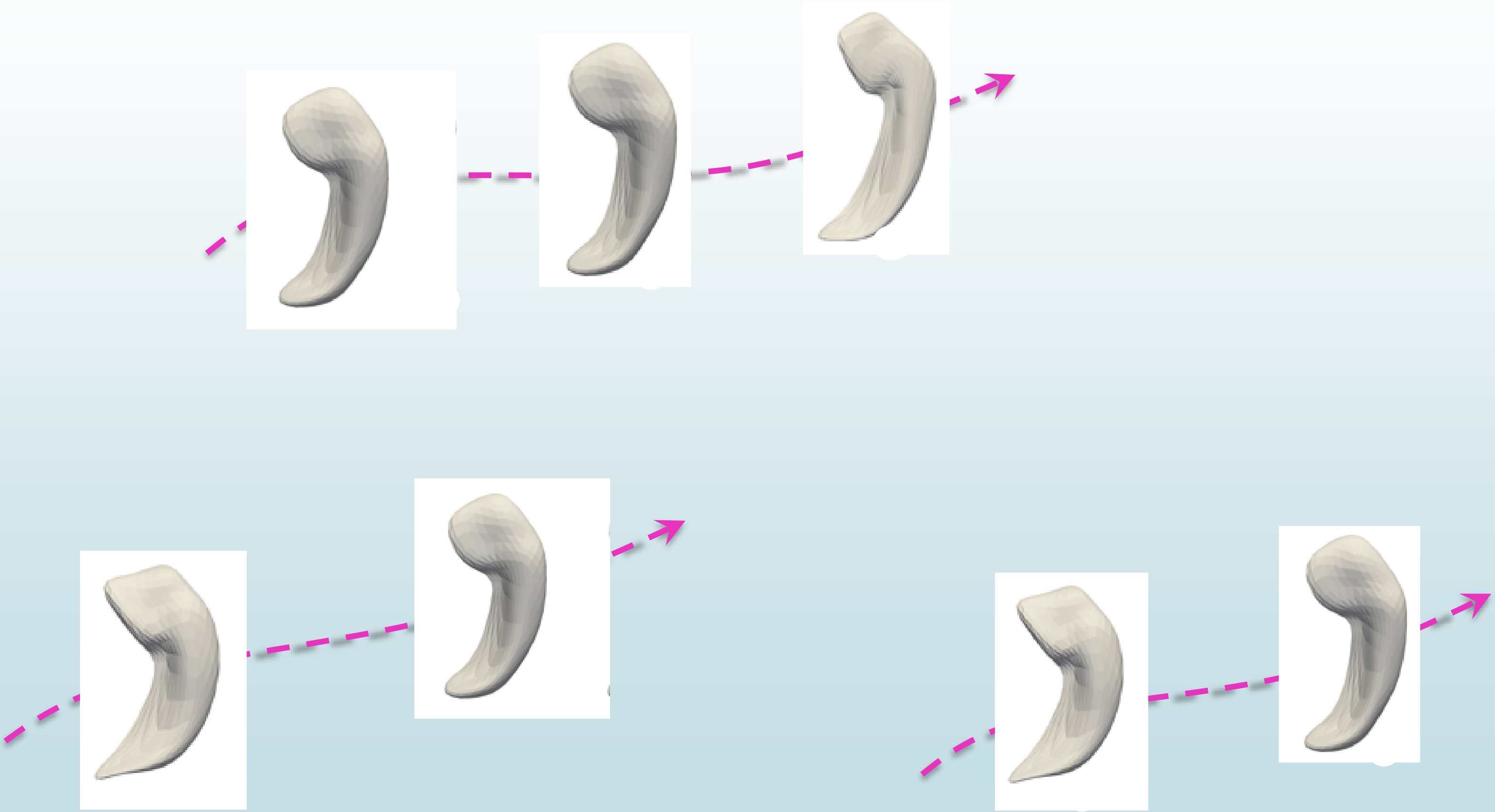
2007 WHO Reference

Now, imagine that the data look more like this:

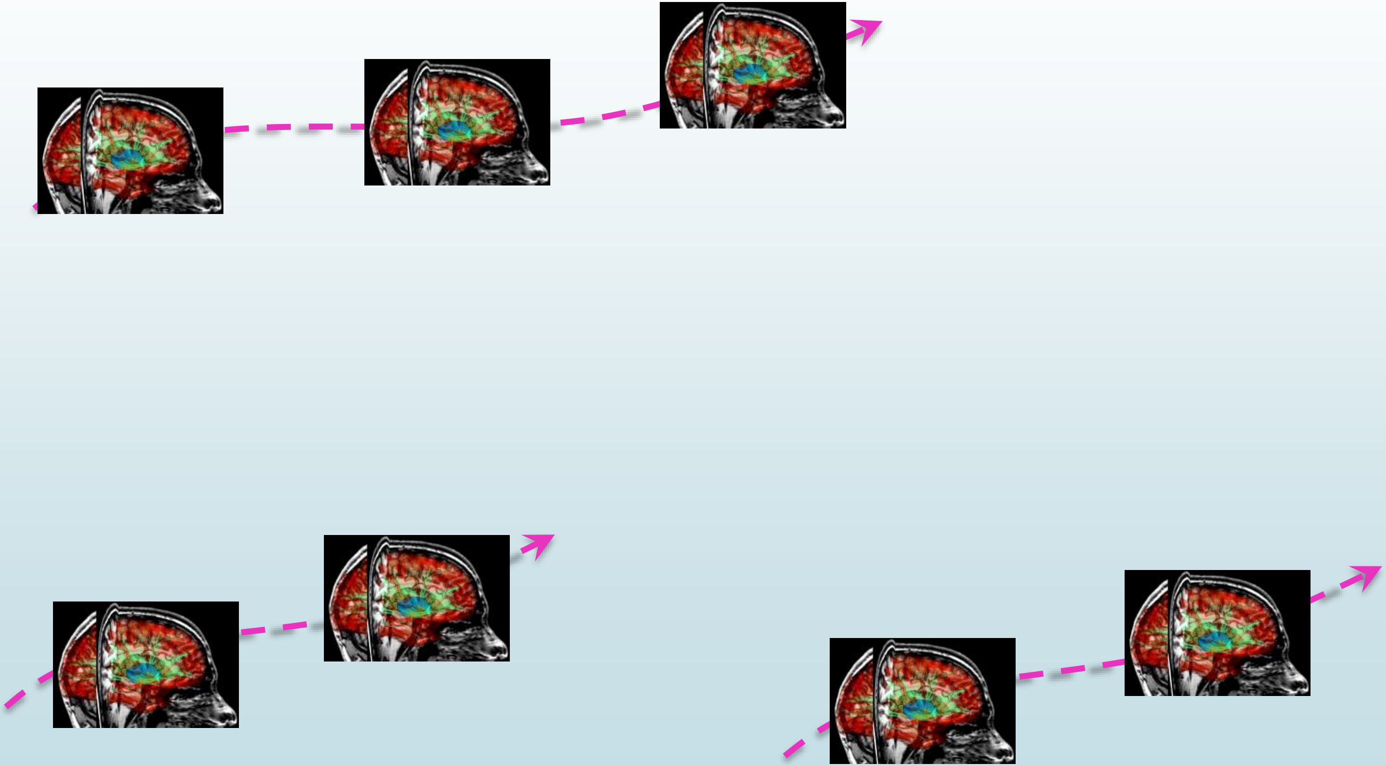


Reference timepoint

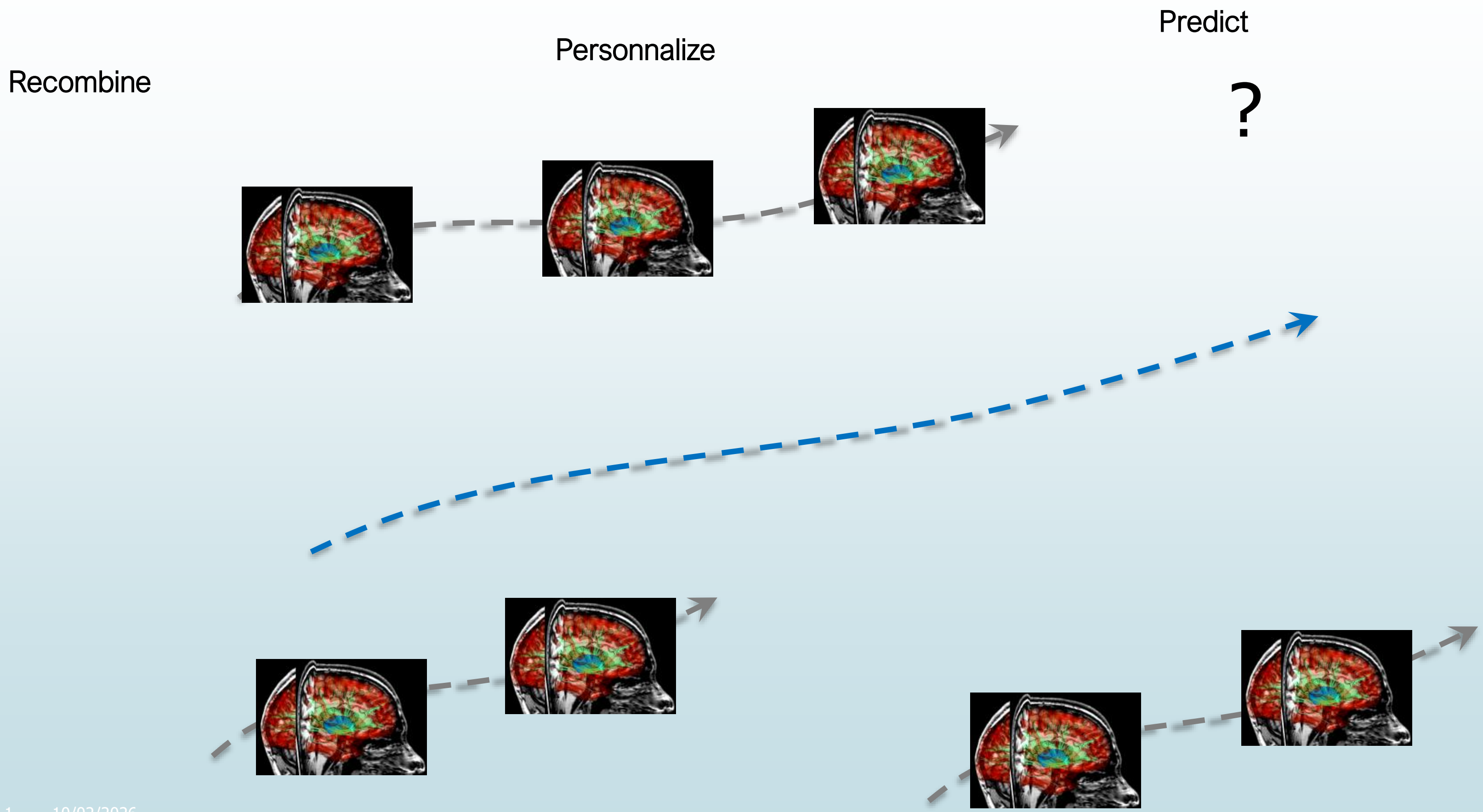
Or this:



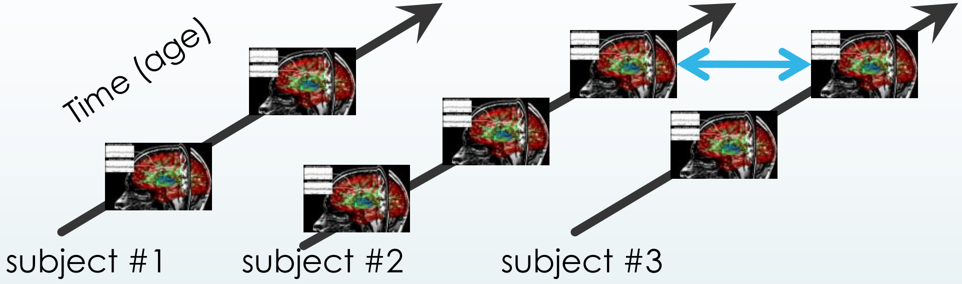
Or even this:



What is now an « average » evolution?



Longitudinal Data Analysis



How to learn representative trajectories of data changes from longitudinal data?

Temporal marker of progression
(e.g. time since drug injection, seeding, birth, etc..)

Regression
(e.g. compare measurements at same time-point)

Linear mixed-effects models
[Laird&Ware 82, Diggle et al., Fitzmaurice et al.]

vector-valued data

No temporal marker of progression
(e.g. in aging, neurodegenerative diseases, etc..)

Learning **spatiotemporal distribution** of trajectories
- *Find temporal correspondences*
- *Compare data at corresponding stages of progression*

Needs to disentangle differences in:
- *Measurements*
- *Dynamics of measurement changes*

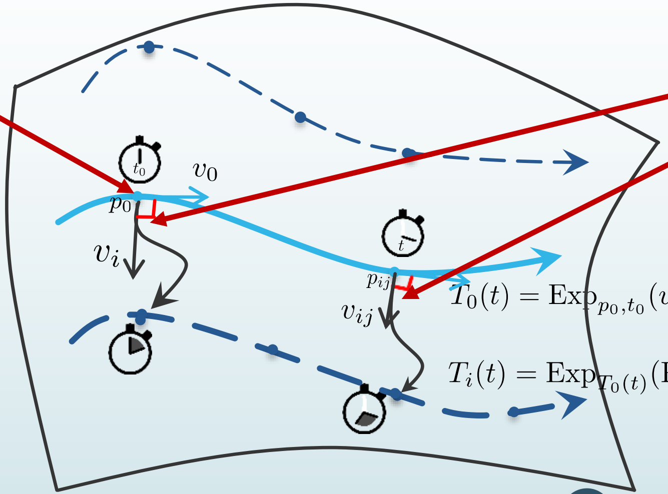
manifold-valued data
(normalized data, positive matrices, shapes, etc;;)

Spatiotemporal mixed effect model

Orthogonality ensures unique space/time decomposition

Invariance w.r.t. reference point on geodesic

Time is random variable



$$T_0(t) = \mathrm{Exp}_{p_0,t_0}(v_0)(t)$$

$$T_i(t) = \mathrm{Exp}_{T_0(t)}(\mathrm{P}^{T_0}_{t_0,t}(v_i))$$

$$y_{ij} = T_i(\psi_i(t)) + \varepsilon_{ij}$$

$$\psi_i(t) = t_0 + \alpha_i(t - t_0 - \tau_i)$$

	Acceleration factor	Time-shift	Space-shift
Random effects:	$\alpha_i \sim log\,\mathcal{N}(0, \sigma_\alpha^2)$	$\tau_i \sim \mathcal{N}(0, \sigma_\tau^2)$	$v_i = (A_1 \vert \dots \vert A_K)\, s_i$ $A_k \perp v_0$
Fixed effects:	(p_0, t_0, v_0)		

[Schiratti et al. IPMI'15, NIPS'15, PhD'17, JMLR'17]

Spatiotemporal Statistical Model

$$\begin{cases} y_{ij} = T_i(\psi_i(t)) + \varepsilon_{ij} \\ T_i(t) = \mathrm{Exp}_{T_0(t)}(\mathrm{P}^{T_0}_{t_0,t}(v_i)) \\ T_0(t) = \mathrm{Exp}_{p_0,t_0}(v_0)(t) \\ \psi_i(t) = t_0 + \alpha_i(t - t_0 - \tau_i) \\ \alpha_i \sim log\,\mathcal{N}(0, \sigma_\alpha^2) \\ \tau_i \sim \mathcal{N}(0, \sigma_\tau^2) \\ v_i = (A_1 | \dots | A_K)\, s_i \\ A_k \perp v_0 \\ (p_0, t_0, v_0) \\ (\sigma_\alpha^2, \sigma_\tau^2, A_1, \dots A_K) \end{cases}$$

- $y_{ij} = T_i(\psi_i(t)) + \varepsilon_{ij}$: Submanifold value observations
- $T_i(t) = \mathrm{Exp}_{T_0(t)}(\mathrm{P}^{T_0}_{t_0,t}(v_i))$: Parallel curve
- $T_0(t) = \mathrm{Exp}_{p_0,t_0}(v_0)(t)$: Representative trajectory
- $\psi_i(t) = t_0 + \alpha_i(t - t_0 - \tau_i)$: Linear time reparametrization

$\alpha_i \sim log\,\mathcal{N}(0, \sigma_\alpha^2)$, $\tau_i \sim \mathcal{N}(0, \sigma_\tau^2)$, $v_i = (A_1 | \dots | A_K)\, s_i$, $A_k \perp v_0$

Hidden random variables:
Acceleration factor
Time shift
Space shift

(p_0, t_0, v_0), $(\sigma_\alpha^2, \sigma_\tau^2, A_1, \dots A_K)$

Parameters:
Mean trajectory parametrization
and prior parameter

Parameter Estimation: EM algorithm

$$y = (y_1, ..., y_N),\ \ z = (z_1, ...z_N),\ \ \theta = (\sigma_z^2, \sigma_\varepsilon^2, A_1, ..., A_K, p_0, t_0, v_0)$$

- Maximum Likelihood: $$max_\theta\, p(y|\theta) = \int p(y, z|\theta) dz$$

- EM: $$\theta_{k+1} = argmax_\theta \sum_{i=1}^{N} \int \log \Big(\underbrace{p(y_i, z_i|\theta)}_{p(y_i|z_i, \theta)p(z_i|\theta)} \Big) p(z_i|y_i, \theta_k) dz_i$$

- Distribution from the curved exponential family

$$\log p(y_i, z_i|\theta) = \phi(\theta)^T S(y_i, z_i) - log(C(\theta))$$

$$\theta_{k+1} = argmax_\theta \left\{ \phi(\theta)^T \sum_{i=1}^{N} \int S(y_i, z_i) p(z_i|y_i, \theta_k) dz_i - N \log(C(\theta)) \right\}$$

Parameter Estimation: stochastic algorithm

- SA-EM: replaces integration by one simulation of the hidden variable:

  sample $z_{i,k+1}$ from $p(z_i|y_i,\theta_k)$,

  and a stochastic approximation of the sufficient statistics

$$\overline{S}_{k+1} = (1-\Delta_k)\overline{S}_k + \Delta_k \left(\frac{1}{N} \sum_{i=1}^{N} S(y_i, z_{i,k+1}) \right)$$

  Maximization step (unchanged) $\theta_{k+1} = argmax_\theta \left\{ \phi(\theta)^T \overline{S}_{k+1} - \log(C(\theta)) \right\}$

- MCMC-SAEM: replaces sampling by a single Markov Chain step

  - For each subject, sample the random effect w.r.t a transition kernel of a geometrically ergodic Markov chain targeting the conditional distribution $q(z_i|y_i,\theta_k)$

[Delyon, Lavielle, Moulines.'99] [Allassonnière et al.'10]

Parameter Estimation: stochastic algorithm

To escape from local maxima
With tempered distributions

- SA-EM: replaces integration by one simulation of the hidden variable:

sample $z_{i,k+1}$ from $p(z_i|y_i,\theta_k)$,

and a stochastic approximation of the sufficient statistics

$$\overline{S}_{k+1} = (1-\Delta_k)\overline{S}_k + \Delta_k \left(\frac{1}{N} \sum_{i=1}^{N} S(y_i, z_{i,k+1}) \right)$$

Maximization step (unchanged) $\theta_{k+1} = argmax_\theta \left\{ \phi(\theta)^T \overline{S}_{k+1} - \log(C(\theta)) \right\}$

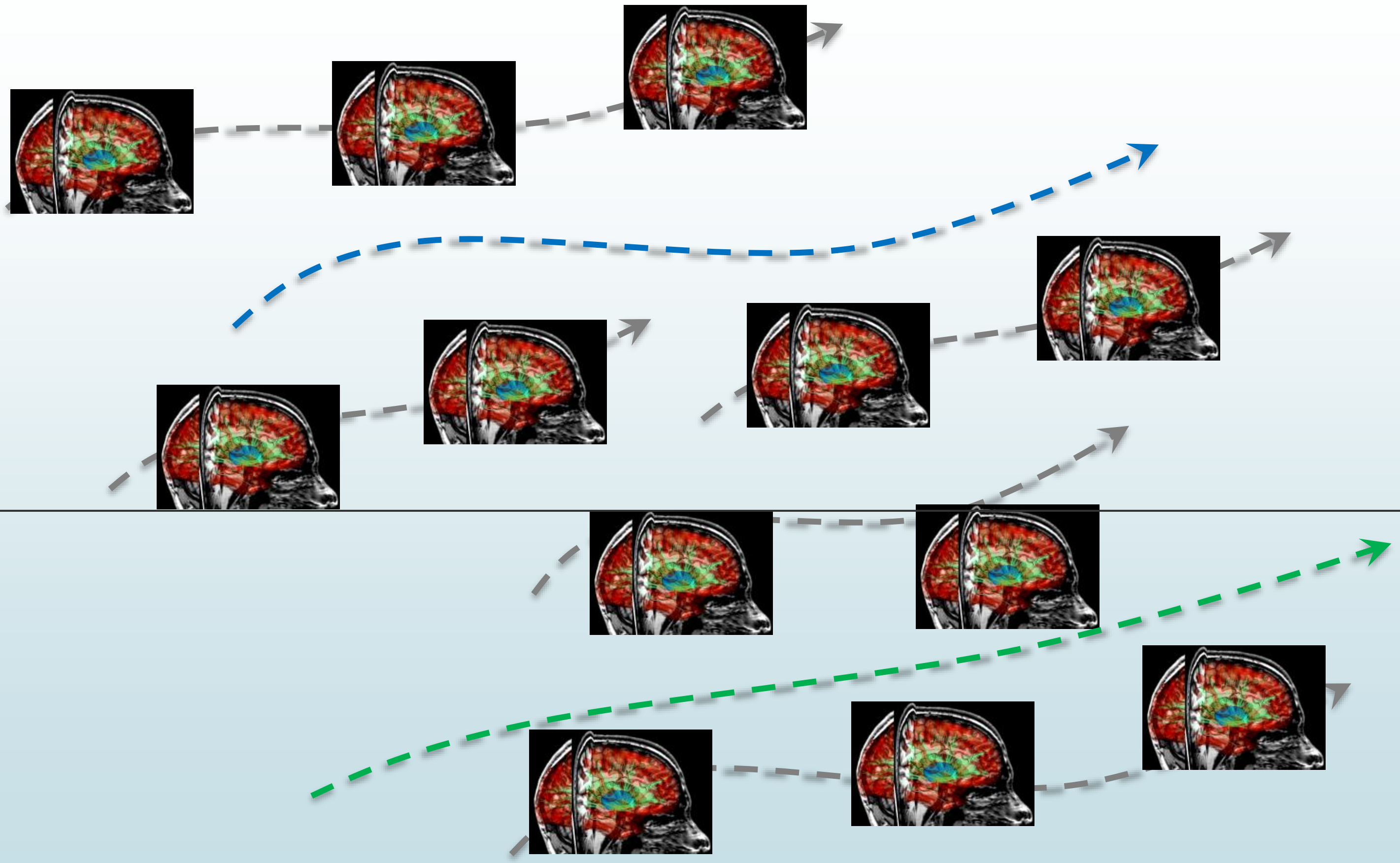
- MCMC-SAEM: replaces sampling by a single Markov Chain step

  - For each subject, sample the random effect w.r.t a transition kernel of a geometrically ergodic Markov chain targeting the conditional distribution $\tilde{q}_k(z_i, \theta_k)$

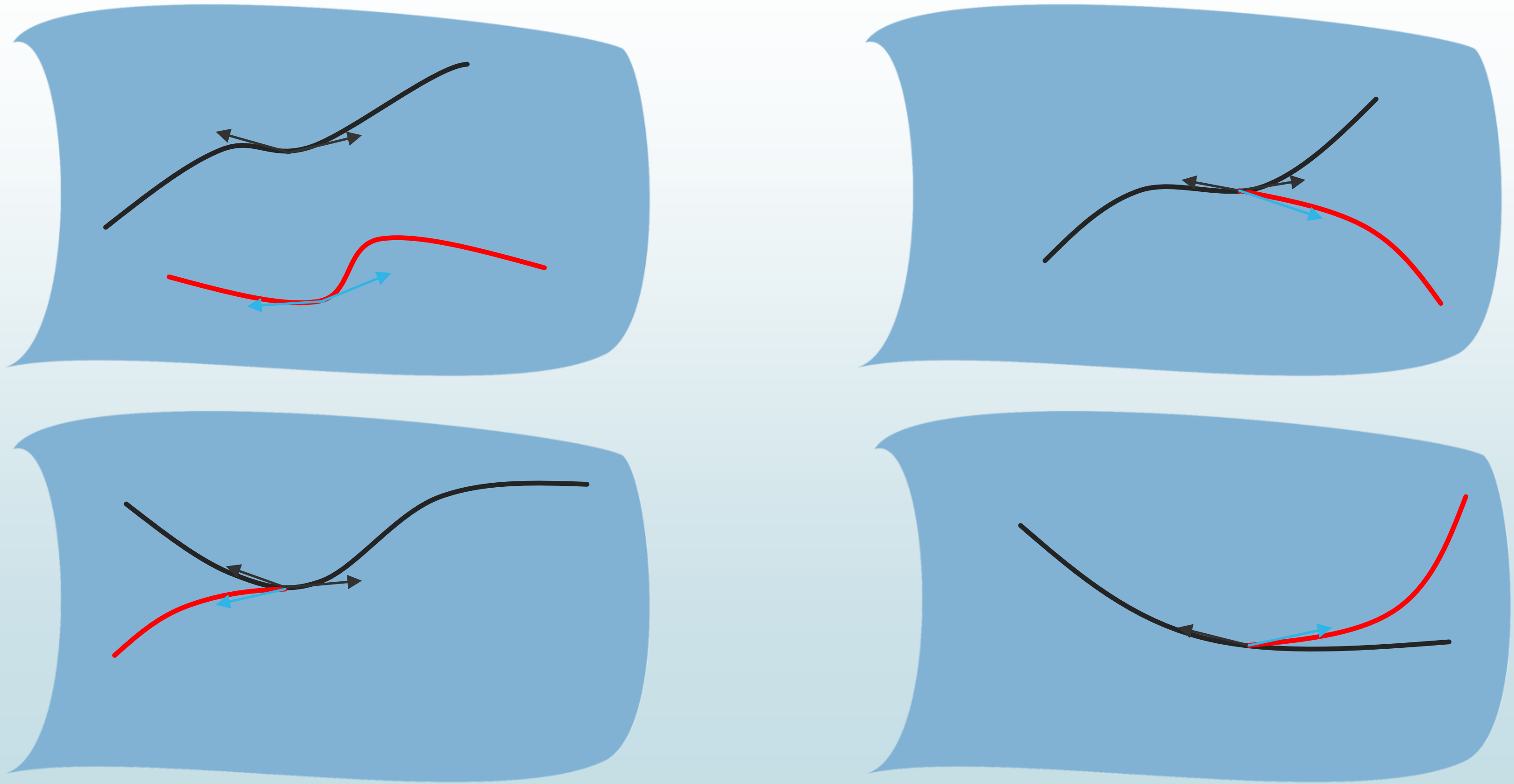
  - As long as $\tilde{q}_k$ "converges towards" $q(z_i|y_i,\theta_k)$ as $k \to \infty$

[Chevallier & Allassonnière, preprint '19, Debavelare et al. MICCAI 2019]

When facing heterogeneous populations:



Including many possible variations in the population:



- Model implemented using deformetrica.
- Soon to be released as a part of this software.

https://www.deformetrica.org/

- https://gitlab.com/icm-institute/aramislab/leaspy

Model of Antiangiogenic response

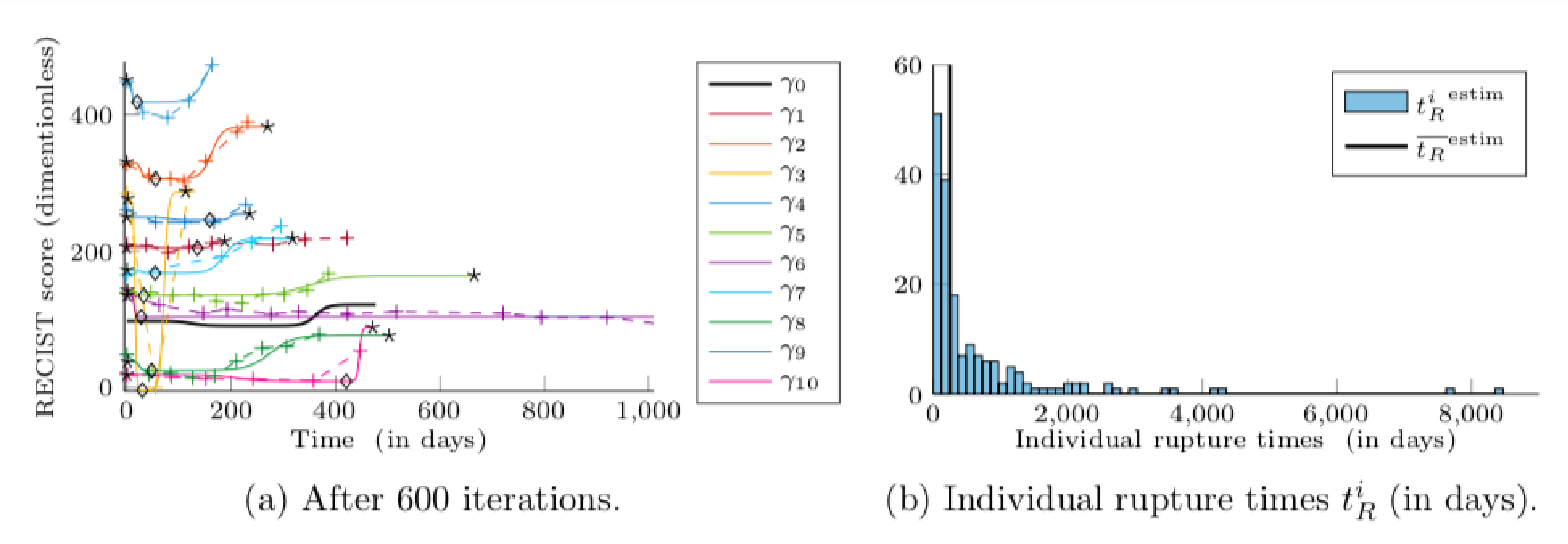
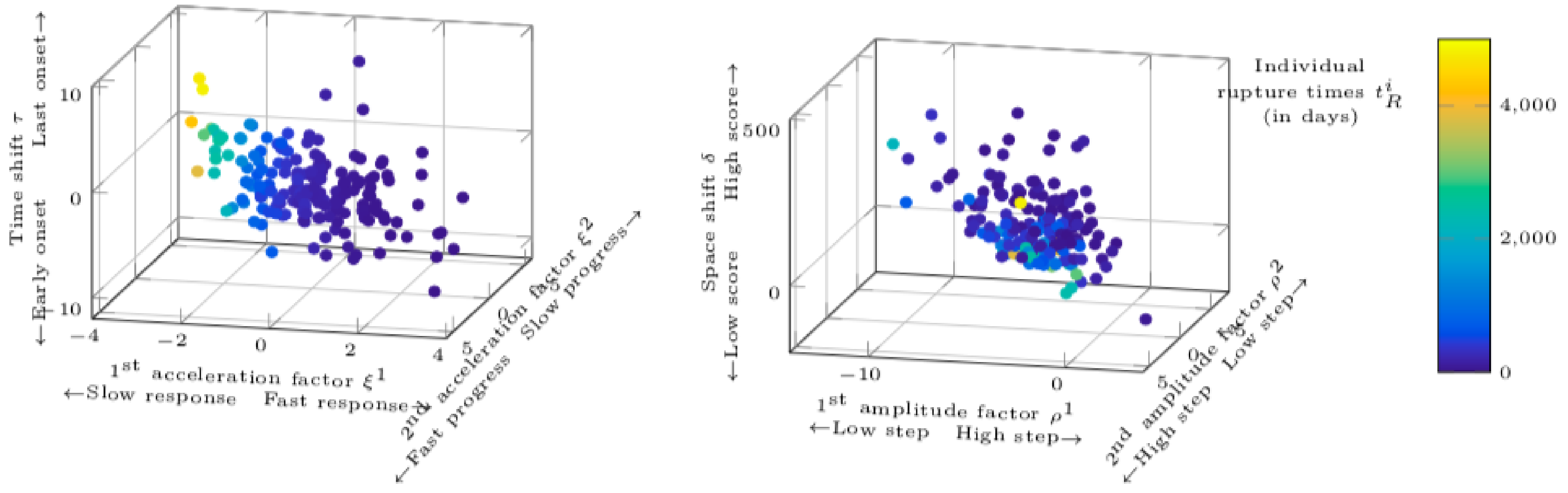


Fig. a. RECIST score for 10 patients among the 176.
Fig. b. Histogram of the estimated rupture times.

Model of Antiangiogenic response

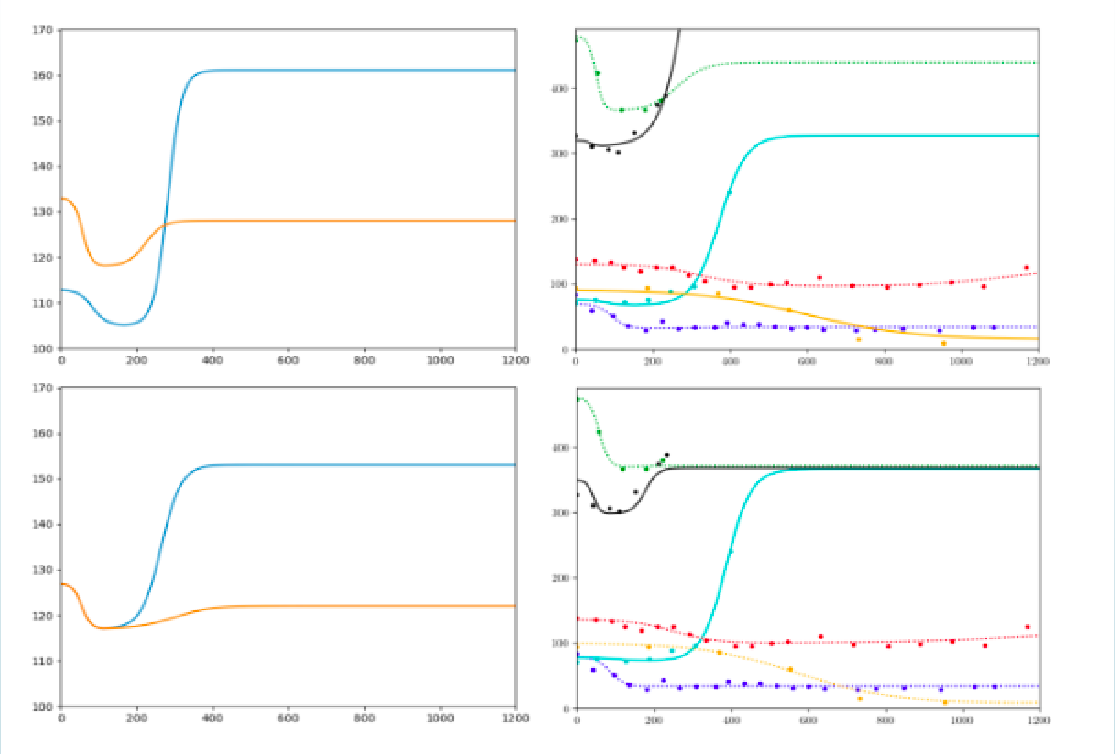


(a) The time warp.

(b) The space warp.

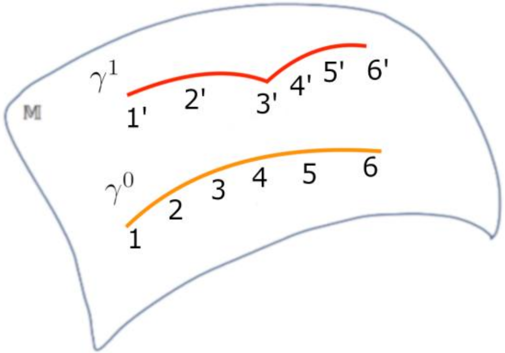
Fig. 5. *Individual random effects.* Fig. 5a: log-acceleration factors ξ_i^1 and ξ_i^2 against times shifts τ_i. Fig. 5b: log-amplitude factors ρ_i^1 and ρ_i^2 against space shifts δ_i. In both figure, the colour corresponds to the individual rupture time t_R^i. These estimates hold for the same run as Fig. 4.

Model of Antiangiogenic response considering the heterogeneity of the population

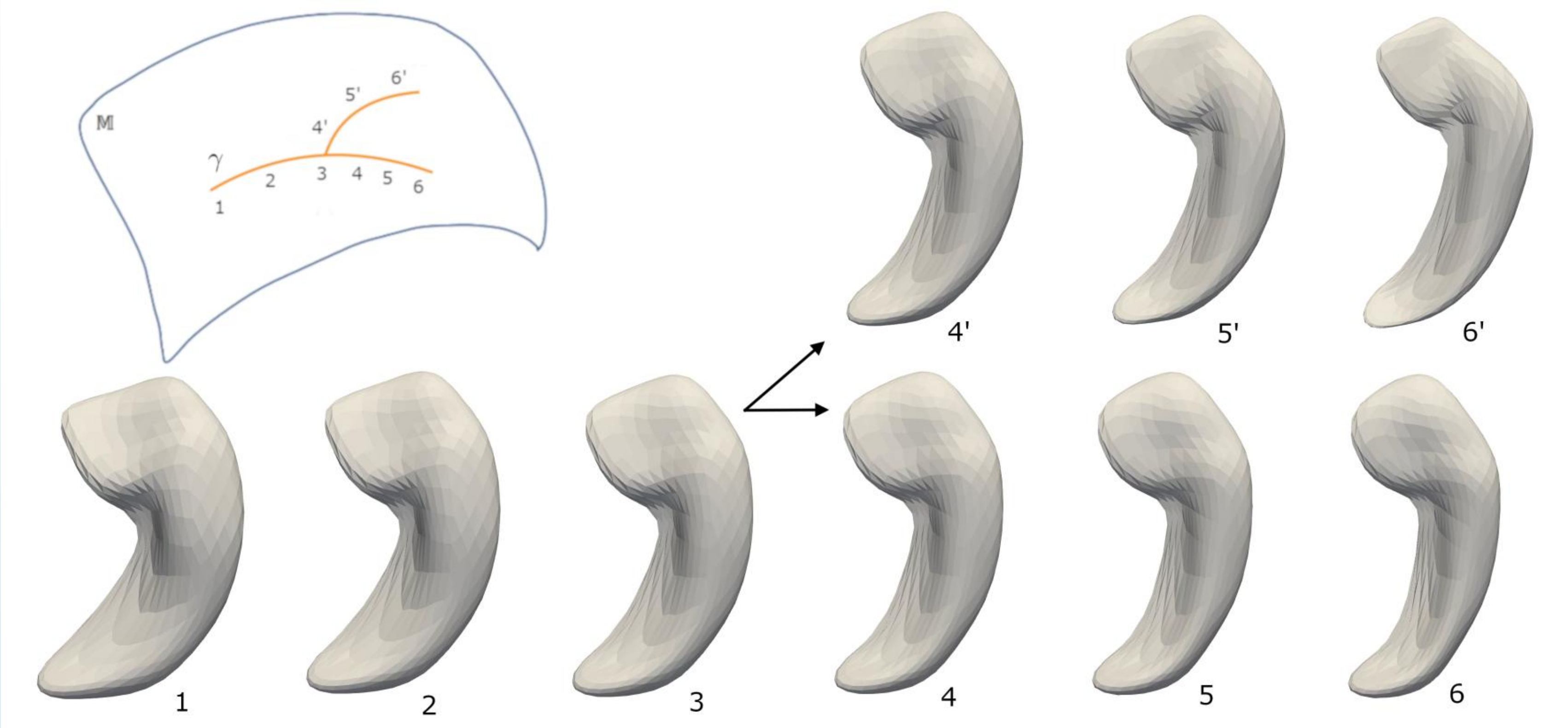


Top: two different clusters. Bottom: two clusters whose first dynamic is shared. Left: Estimated templates. Right: 6 subjects and their reconstructed trajectories. In dotted lines, subjects in the cluster of the orange template. In plain lines subjects in the cluster of the blue template

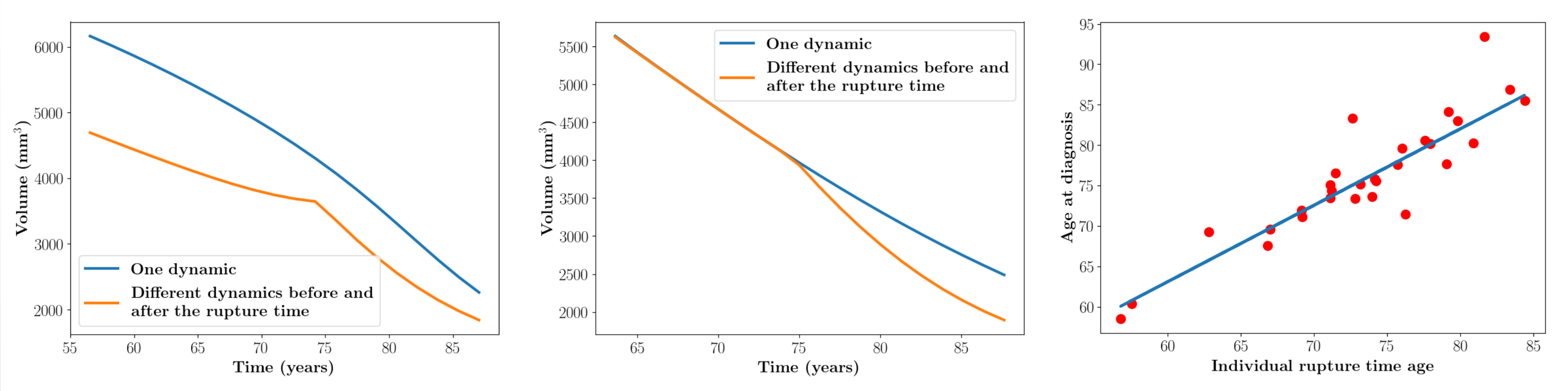
Model of Alzheimer's disease conversion in heterogeneity of the population



Model of Alzheimer’s disease conversion in heterogeneity of the population



Comparisons of these models of Alzheimer's disease conversion



Un exemple concret :

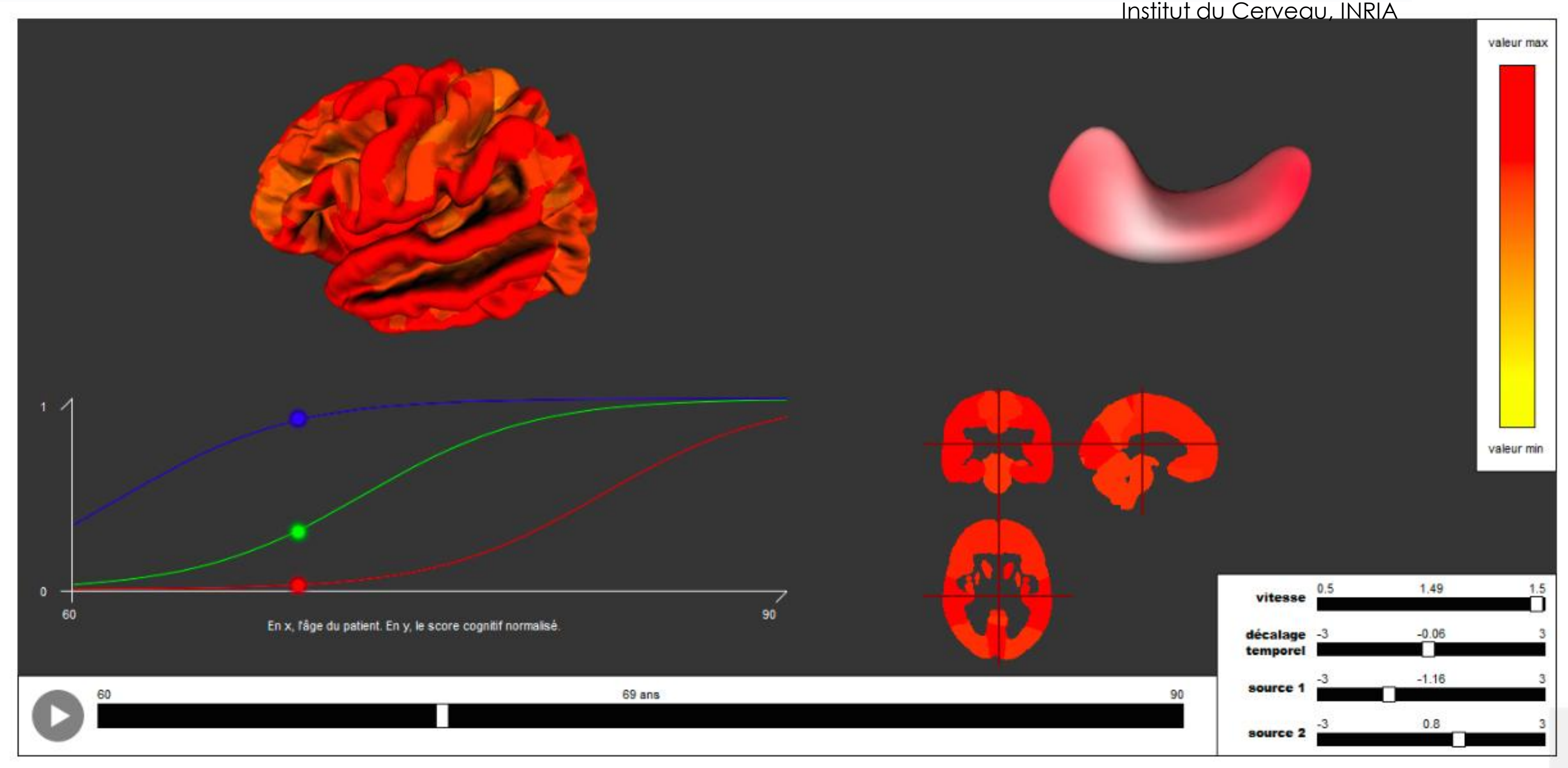
La clinique du docteur memo : The experts of brain aging

LA CLINIQUE DU

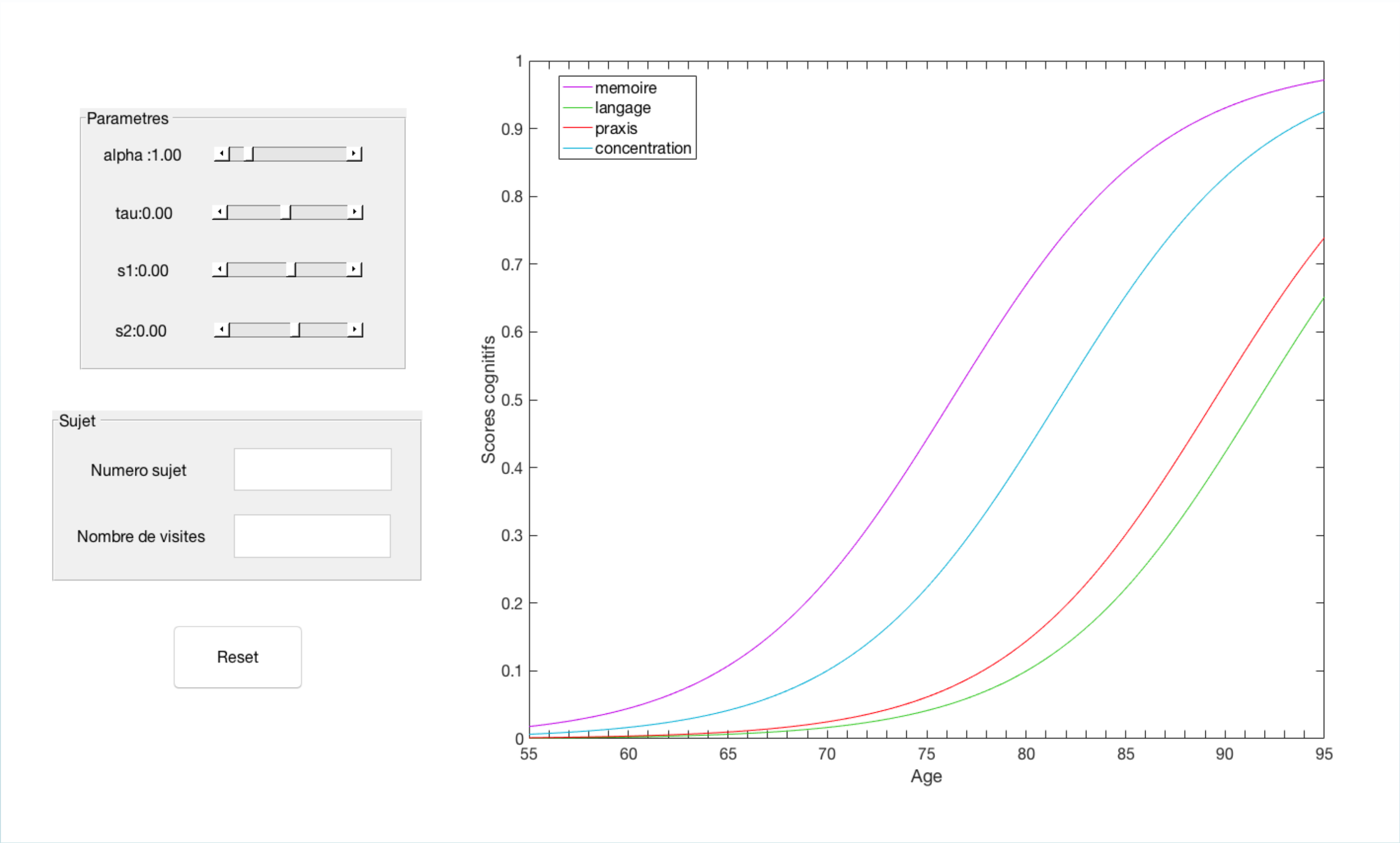
Docteur Memo

ÉVOLUTION DE LA MALADIE D'ALZHEIMER

Collaboration à l'équipe Aramis, de Stanley Durrleman,
Institut du Cerveau, INRIA

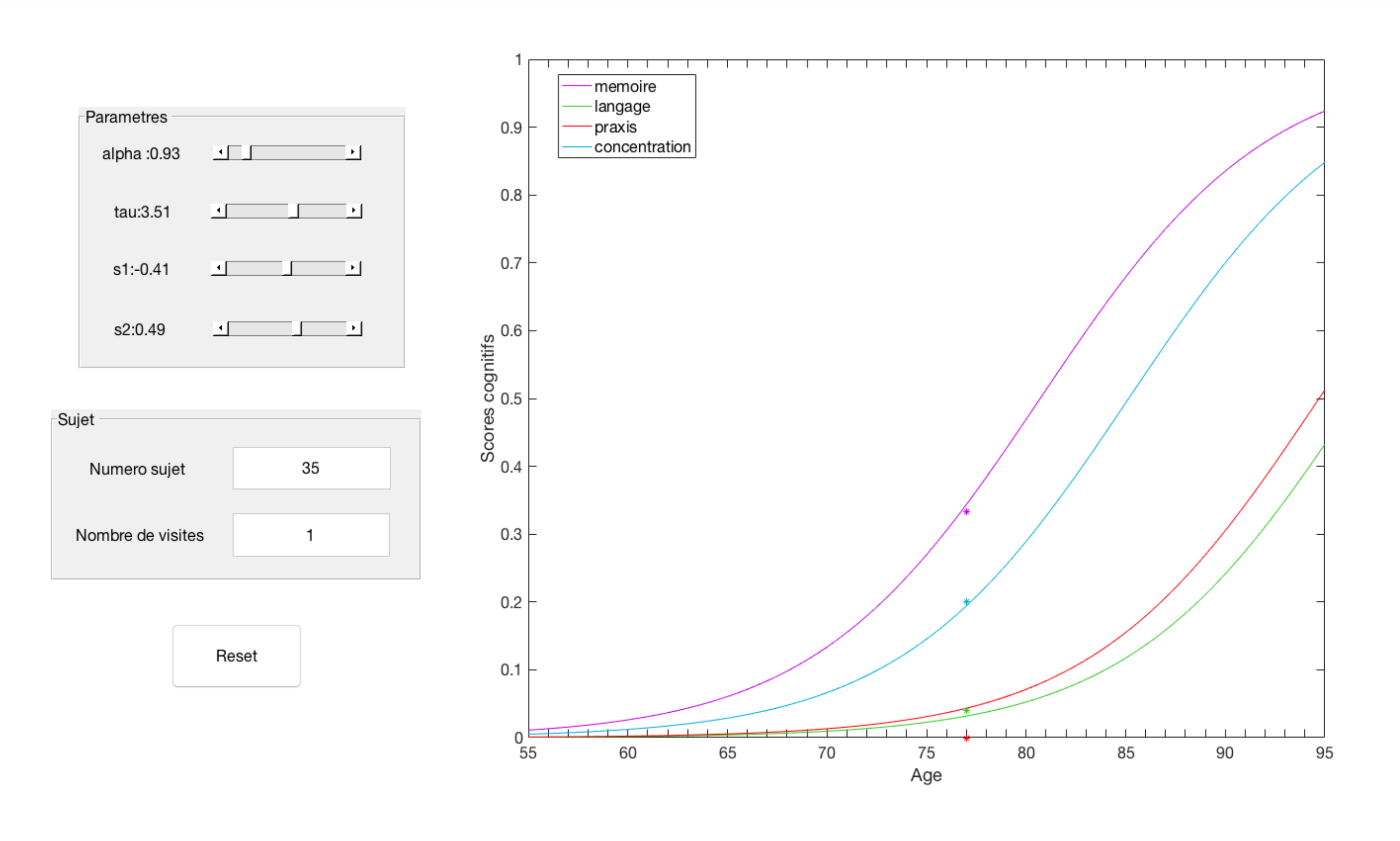


AUTOMATIC PROGNOSIS SYSTEM (PATENTED)

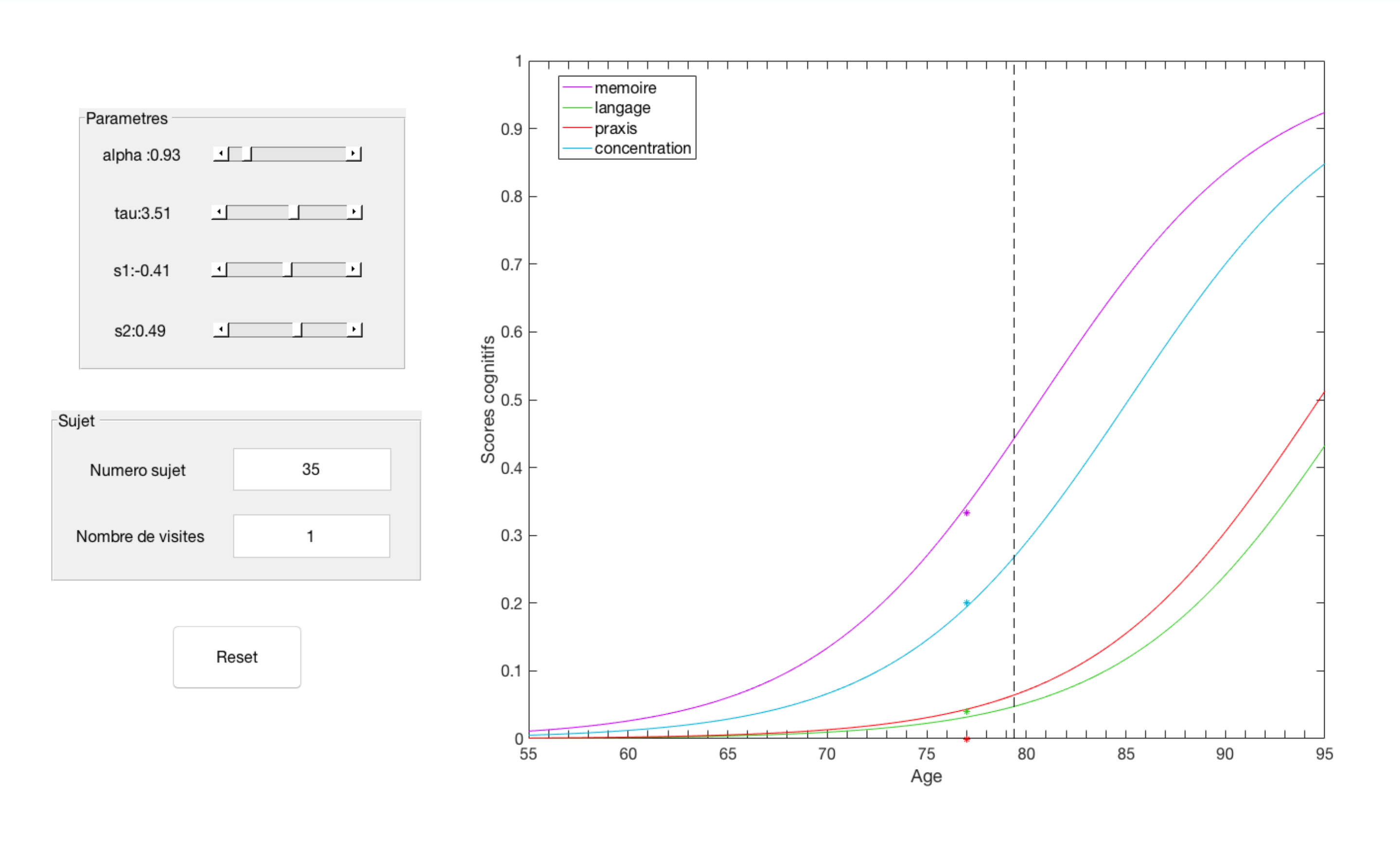


AUTOMATIC PROGNOSIS SYSTEM (PATENTED)

LA CLINIQUE DU Docteur Memo

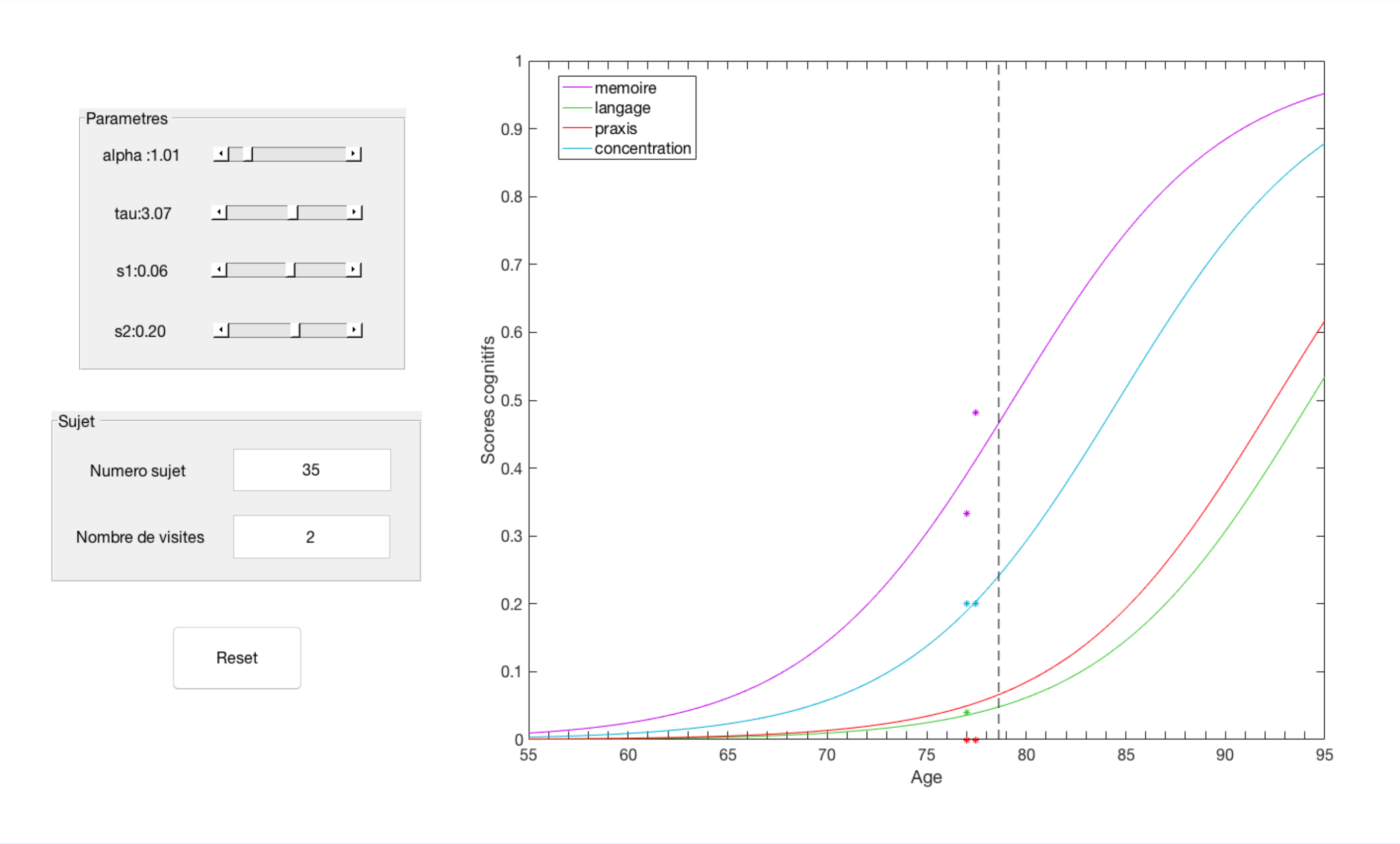


AUTOMATIC PROGNOSIS SYSTEM (PATENTED)



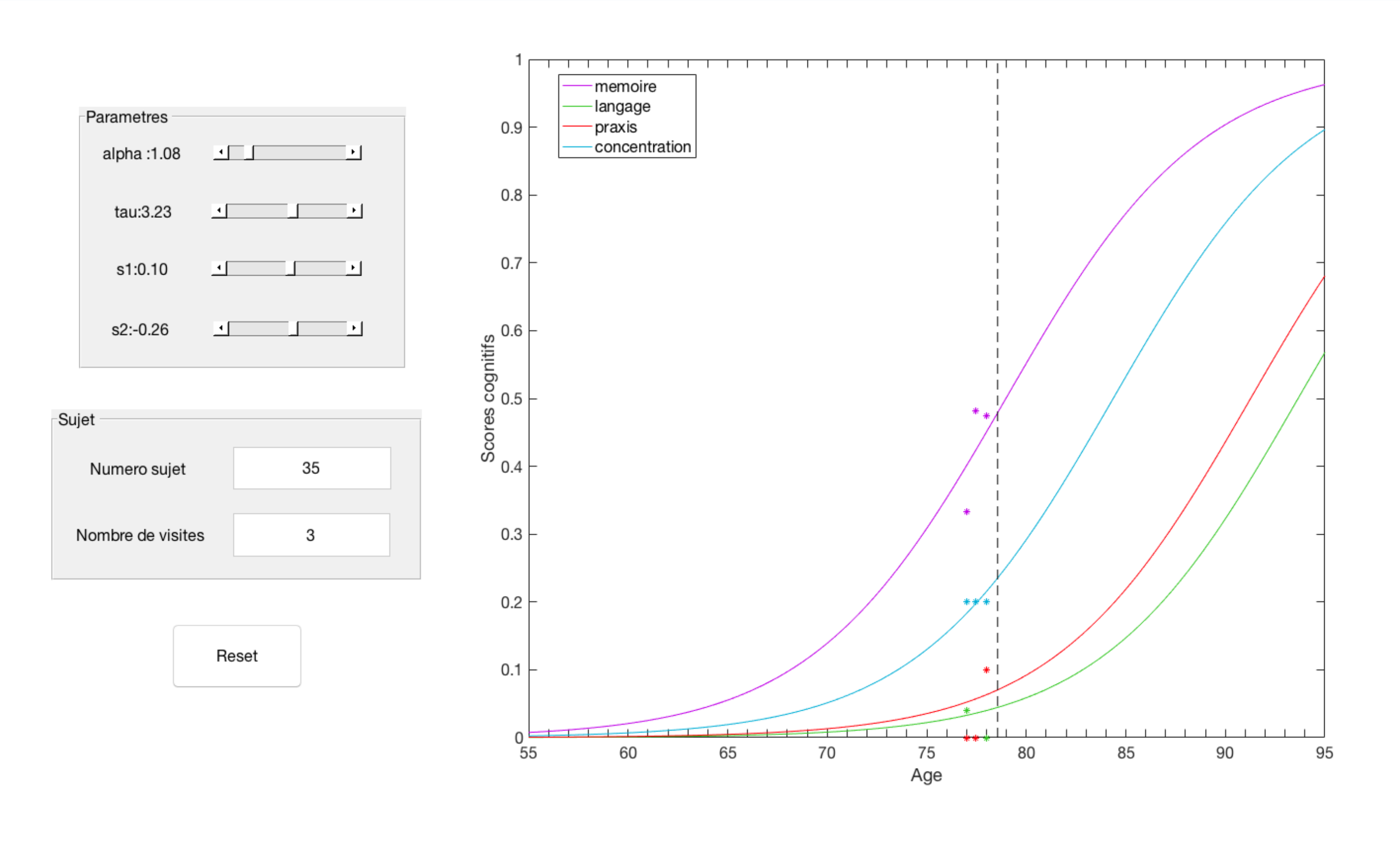
AUTOMATIC PROGNOSIS SYSTEM (PATENTED)

LA CLINIQUE DU Docteur Memo



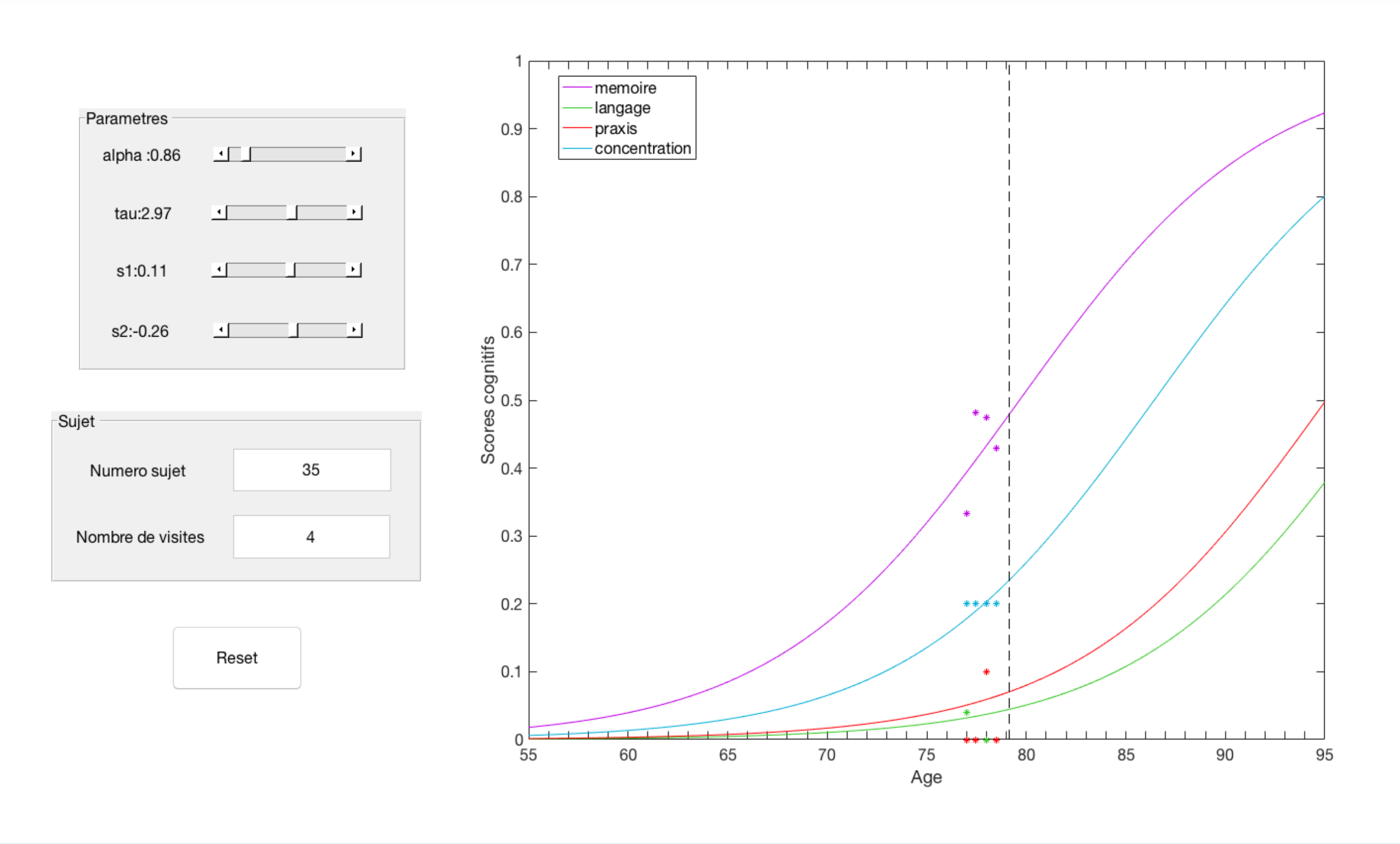
AUTOMATIC PROGNOSIS SYSTEM (PATENTED)

LA CLINIQUE DU Docteur Memo

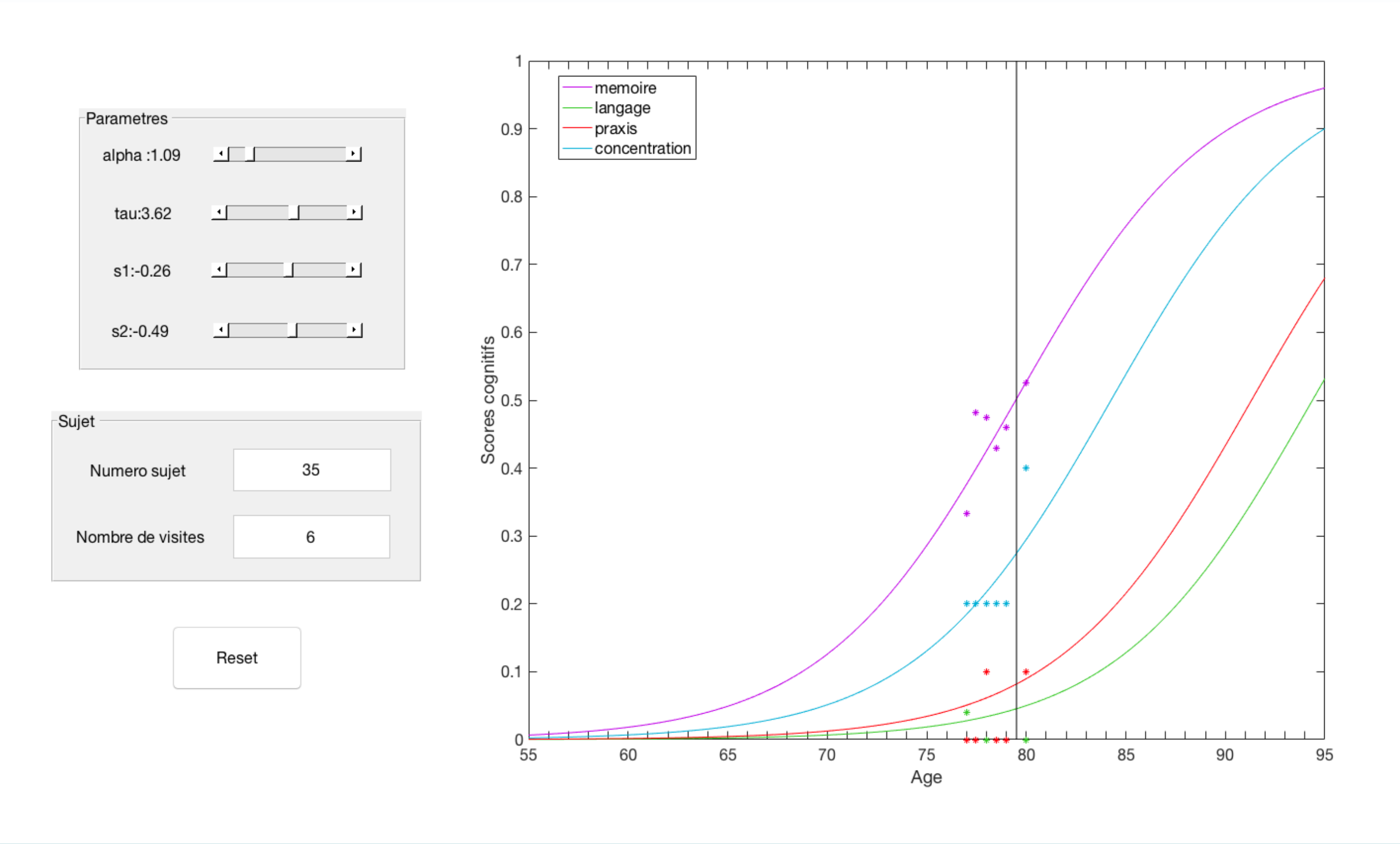


AUTOMATIC PROGNOSIS SYSTEM (PATENTED)

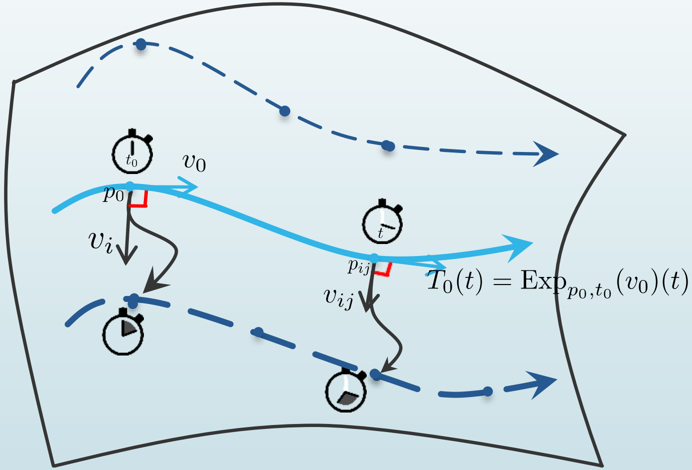
LA CLINIQUE DU Docteur Memo

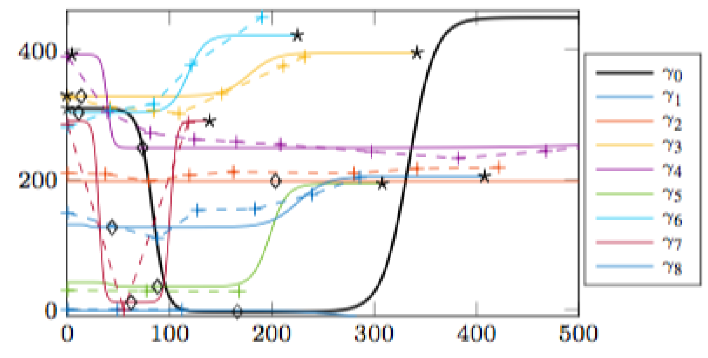


AUTOMATIC PROGNOSIS SYSTEM (PATENTED)



Thank you!





(a) After 600 iterations.