

## Supplementary Information: B

### Data set

Trajectories were defined from questionnaires with an available body mass index and age, over the 1990 to 2015 period (26 questionnaires), among participants of the GAZEL cohort who responded in 2015. Trajectories were defined separately among men and women.

### Model parameters

Trajectories model the evolution of body mass index over time. Models take into account the heterogeneity of the population by defining subgroups of individuals sharing similar BMI trends. These models were defined using latent class mixed models<sup>43</sup>.

Time was considered as age minus minimum age at inclusion (35 years for women, 40 for men). Age at the time of response was defined from the date of response to the questionnaire and the individual's date of birth.

An unstructured variance-covariance matrix was used and defined as common over latent classes. Linear growth was considered, in accordance with others studies<sup>2, 3</sup> and to reduce computational intensity. A fixed effect of period of birth (binary indicator, birth after World War II) was defined to take into account a known cohort effect on body mass index that was highlighted by Dugravot et al.<sup>4</sup> and confirmed in exploratory data analysis.

### Management of local maxima & convergence

To avoid convergence to local maxima (for models with at least two latent classes), a 2-step method was performed:

- 1) Initial values from a one-class model were randomly drawn for a large number (300) of models with a small number of iterations (50) for each. Of the 300 results, the model with the best log-likelihood was selected, regardless of convergence criterion.
- 2) Its values were used as initial values for a small number (25) of models with a large number of iterations (400) for each. Of the 25 results, converging models were identified. The one with the best log-likelihood was selected as the final model.

If models failed to converge, or reached obvious local maxima (a lower latent class number model's solution), the method was re-run with doubled numbers of models and iterations for each step. If models also failed, no results were retained.

#### Models selection

Models from one to five latent classes were performed. Selection of the appropriate number of classes was guided by both model fit and interpretability using the Bayesian Information Criterion (BIC) and the interpretability criterion derived from Nagin's criterion<sup>5</sup>. The model retained has the lowest BIC and for each latent class both sufficient prevalence ( $1\% \leq$ ) and mean posterior probabilities ( $80\% \leq$ ).

(See results in table 1 in article & B1 in annex).

For both men and women, a 3-class solution was selected. Their descriptions can be found in the main paper.

#### Software

R statistical software (version 3.4.4) was used and LCMM package (version 1.7.9) for latent class mixed modeling.

### Supplementary Information References

1. Proust-Lima C, Philipps V, Lique B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. 2015;78(2).
2. Mezuk B, , Matthew C. Lohman, Andrew K. Rock and MEP. Trajectories of Body Mass Indices and Development of Frailty: Evidence from the Health and Retirement Study. *Obesity*. 2016;24(8):1643-1647.
3. Zajacova A, Ailshire J. Body Mass Trajectories and Mortality Among Older Adults: A Joint Growth Mixture–Discrete-Time Survival Analysis. *Gerontologist*. 2014;54(2):221-231.
4. Dugravot A, Sabia S, Stringhini S, et al. Do socioeconomic factors shape weight and obesity trajectories over the transition from midlife to old age? Results from the French GAZEL cohort study. *Am J Clin Nutr*. 2010;92(1):16-23.
5. Nagin D, Odgers CL. Group-Based Trajectory Modeling in Clinical Research. *Ssrn*. 2010.

## Supplementary Information Table

Table B1: Three-class models of posterior probabilities for men and women of the GAZEL cohort.

		Men			Women		
		<b>Class 1</b> 8 278 (95%)	<b>Class 2</b> 342 (3.9%)	<b>Class 3</b> 131 (1.5%)	<b>Class 1</b> 2 736 (90%)	<b>Class 2</b> 143 (4.7%)	<b>Class 3</b> 154 (5.1%)
Posterior probability	Mean (sd)	98 (6.8)	84 (16)	85 (16)	97 (8.2)	80 (17)	88 (16)
	>0.7: % (N)	98 (8 106)	74 (254)	79 (103)	97 (2 656)	71 (101)	84 (129)
	>0.8: % (N)	97 (7 991)	64 (218)	68 (89)	94 (2 584)	57 (82)	76 (117)
	>0.9: % (N)	94 (7 752)	49 (169)	48 (63)	91 (2 479)	41 (58)	65 (100)