Data Supplement

Somatotypes trajectories during adulthood and their association with COPD phenotypes.

Authors:

Miguel J. Divo, Marta Marin Oto, Ciro Casanova Macario, Carlos Cabrera Lopez, Juan P. de-Torres, Jose Maria Marin Trigo, Craig P. Hersh, Ana Ezponda Casajús, Cherie Maguire, Victor M. Pinto-Plata, Francesca Polverino, James C. Ross, Dawn DeMeo, Gorka Bastarrika, Edwin K Silverman and Bartolome R. Celli

1. Statistical analysis-trajectory modeling
2. Figures

S1. Regression plot demonstrating the linear relationship between current Stunkard's somatotype score and current BMI.

3. Tables

Table S1. Baseline characteristics comparing subjects from the BODE and COPDGene cohorts.

Statistical analysis for trajectory modeling

As used in the Nurses' Health Study and the Health Professionals Follow-up Study analysis[1, 2], we performed the model fitting and assessment using a censored normal model as a polynomial function of the time scale (i.e. at 18, 30 , 40, 50 and current age). The optimal number of groups and the shapes of trajectories were selected for best fit to the data using a two-stage approach based on the change in the Bayesian Information Criterion (BIC) [3] and clinical interpretability. The first stage was to determine the number of groups using a cubic form for all trajectory groups. We considered up to five groups and compared the BIC with models with 4 and 3 groups, respectively. Once observing that the model with five groups fitted the data best, we then determined in the second stage the order of the polynomial function specifying the shape of each trajectory. We compared the BIC of the five group models with different polynomial function (linear, quadratic and cubic) and found that the model with all groups up to cubic order terms demonstrated the best fit to the data. Therefore, the final trajectory model was built using a cubic function of age for each of the five trajectories. We then named the trajectory groups to describe their visual patterns based on their initial point at age 18 and the directionality of the trajectory as Lean-Flat, Lean-Increase, Medium-Increase and Medium-Parabolic and Heavy-Increase.

Once a final model was selected, we calculated the posterior predicted probability for each individual of being a member of each of the five trajectories. Participants were assigned into the trajectory group to which their posterior membership probability was highest. We performed model diagnostics to assess fit to the data as specified in [3]. A group APP value of 1 indicates that individuals have been assigned with certainty. As recommended in [3], per-group APP should be > 0.7, a criterion met by all groups found in this study. The average membership

probability for each trajectory group in the order listed above was 0.92, 0.89, 0.85, 0.87 and 0.89 respectively.

Figure S1. Regression plot demonstrating the linear relationship between current Stunkard's somatotype score (predictor) and current BMI (outcome). Red line represents the fitting line for females while the blue is for males. There is no significant difference in the relation between somatotype and BMI amongst males and females .

The prediction formula:

BMI= 17.6 (95% CI 16.7 -18.5) + 1.9 (95% CI 1.8- 2.1) x somatotype score

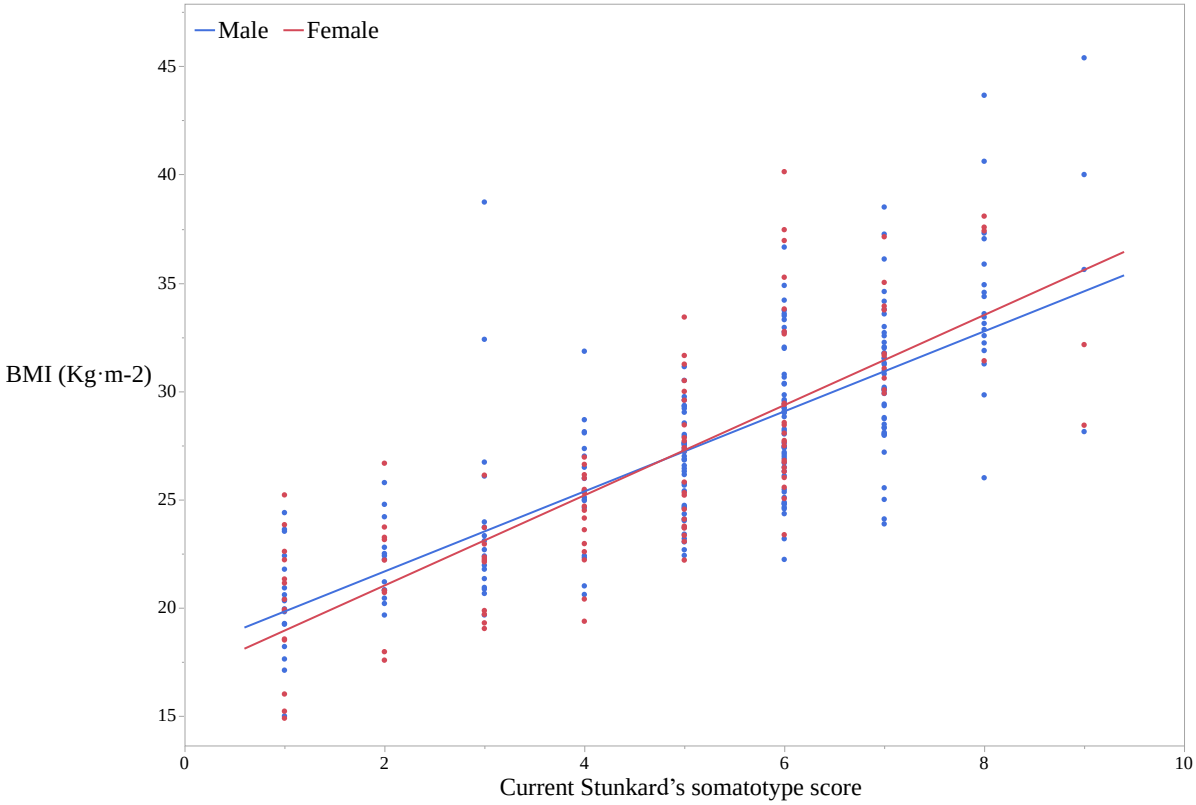

Table S1. Comparison of subjects' baseline characteristics by cohort.

	COPDGene cohort	BODE cohort	*p-value*
N	55	301	
Age, yr	72 ± 8	67 ± 9	*<0.0001*
Males %	60%	69%	0.2019
BMI ($kg \cdot m^{-2}$)	26.6 ± 4.5	27.3 ± 5.2	*0.8465*
BMI categories:			
BMI ≤23 $kg \cdot m^{-2}$	11 (20%)	62 (20%)	NS
23<BMI<30 $kg \cdot m^{-2}$	32 (58%)	153 (51%)	
BMI≥30 $kg \cdot m^{-2}$	12 (22%)	86 (29%)	
FEV_1 % predicted	72 ± 18	67 ± 22	0.0545
GOLD classification:			
GOLD 1	19 (34%)	97 (32%)	NS
GOLD 2	28 (51%)	132 (44%)	
GOLD 3	8 (15%)	59 (20%)	
GOLD 4	0	12 (4%)	

References

1. Song M, Hu FB, Wu K, Must A, Chan AT, Willett WC, Giovannucci EL. Trajectory of body shape in early and middle life and all cause and cause specific mortality: results from two prospective US cohort studies. *BMJ* 2016; : i2195–10.

2. Zheng Y, Song M, of JMAJ, 2017. Group-based trajectory of body shape from ages 5 to 55 years and cardiometabolic disease risk in 2 US cohorts. *academic.oup.com*

3. NAGIN D. Group-Based Modeling of Development. Cambridge, MA and London, England: Harvard University Press.