

S1 Table. GLI-based z-scores of all healthy never-smoking women (HNSW) with available lung function measurements at baseline and follow-up grouped by time of examination. Two one-sided tests for equivalence were performed to establish equivalence between the mean z-score of the GLI reference population and SALIA ($p < 0.05$).

		N	Mean	sd	p-value*	% \leq LLN
Baseline	FEV ₁	299	-0.11	0.90	<0.001	3.68
	FVC	299	0.07	0.81	<0.001	2.34
	FEV ₁ /FVC	299	-0.35	0.79	0.001	5.35
Follow-up	FEV ₁	299	0.22	1.03	<0.001	2.68
	FVC	299	0.45	0.99	0.177	1.67
	FEV ₁ /FVC	299	-0.45	0.81	0.150	8.03

FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; LLN: Lower Limit of Normal.

*: Two one-sided tests (TOST) for equivalence were performed to test on equivalence between the mean z-scores of the GLI reference population and SALIA. A good fit was reached if the null-hypothesis of a mean z-score outside of the interval $[-0.5, 0.5]$ was rejected at the 5% significance level (H_0 : mean $z \notin [-0.5, 0.5]$).