nature medicine

Supplementary information

https://doi.org/10.1038/s41591-024-03144-x

Brain aging patterns in a large and diverse cohort of 49,482 individuals

In the format provided by the authors and unedited

Brain aging patterns in a large and diverse cohort of 49,482 individuals

Supplementary Information

	λ=0.1	λ=0.2	λ=0.4	λ=0.6
γ=12	0.734/0.665	0.749/0.673	0.734/0.641	0.710/0.627
γ=16	0.736/0.665	0.733/0.662	0.727/0.654	0.668/0.606
γ=20	**0.764/0.692**	**0.771/0.696**	0.743/0.670	0.715/0.640
γ=24	**0.764/0.694**	**0.778/0.704**	0.755/0.670	0.711/0.643

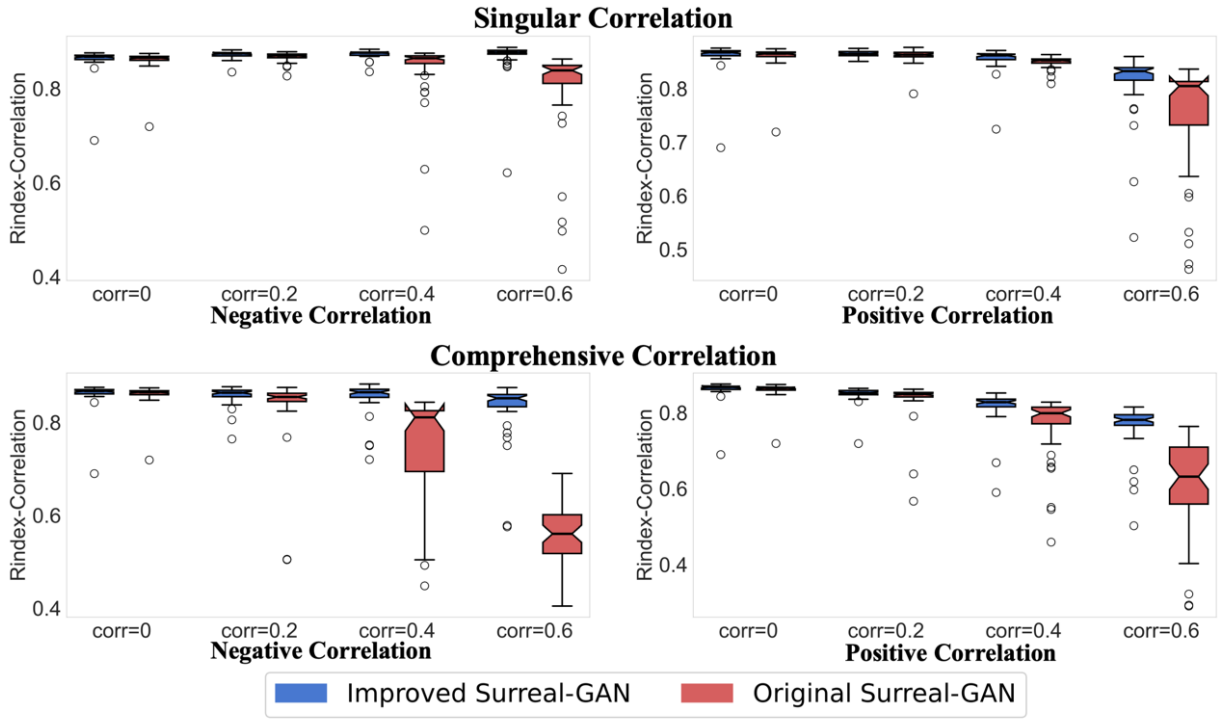
Supplementary Table 1. Agreements among repetitively trained models indicate underlying representation accuracies. We trained the models with different combinations of λ and γ on semi-synthetic data with a maximum of 15% simulated atrophy rate. Within each cell, we presented agreements/average representation accuracies. Parameters resulting in the four highest agreements are highlighted, with the highest one being colored for emphasis.

Simulated Atrophy Rate	Basic Model	Mild Model	Extra Mild Model
0-30%	0.940±0.002	0.932±0.004	0.824±0.113
0-20%	0.867±0.003	0.867±0.005	0.760±0.103
0-15%	0.775±0.004	0.778±0.008	0.704±0.081

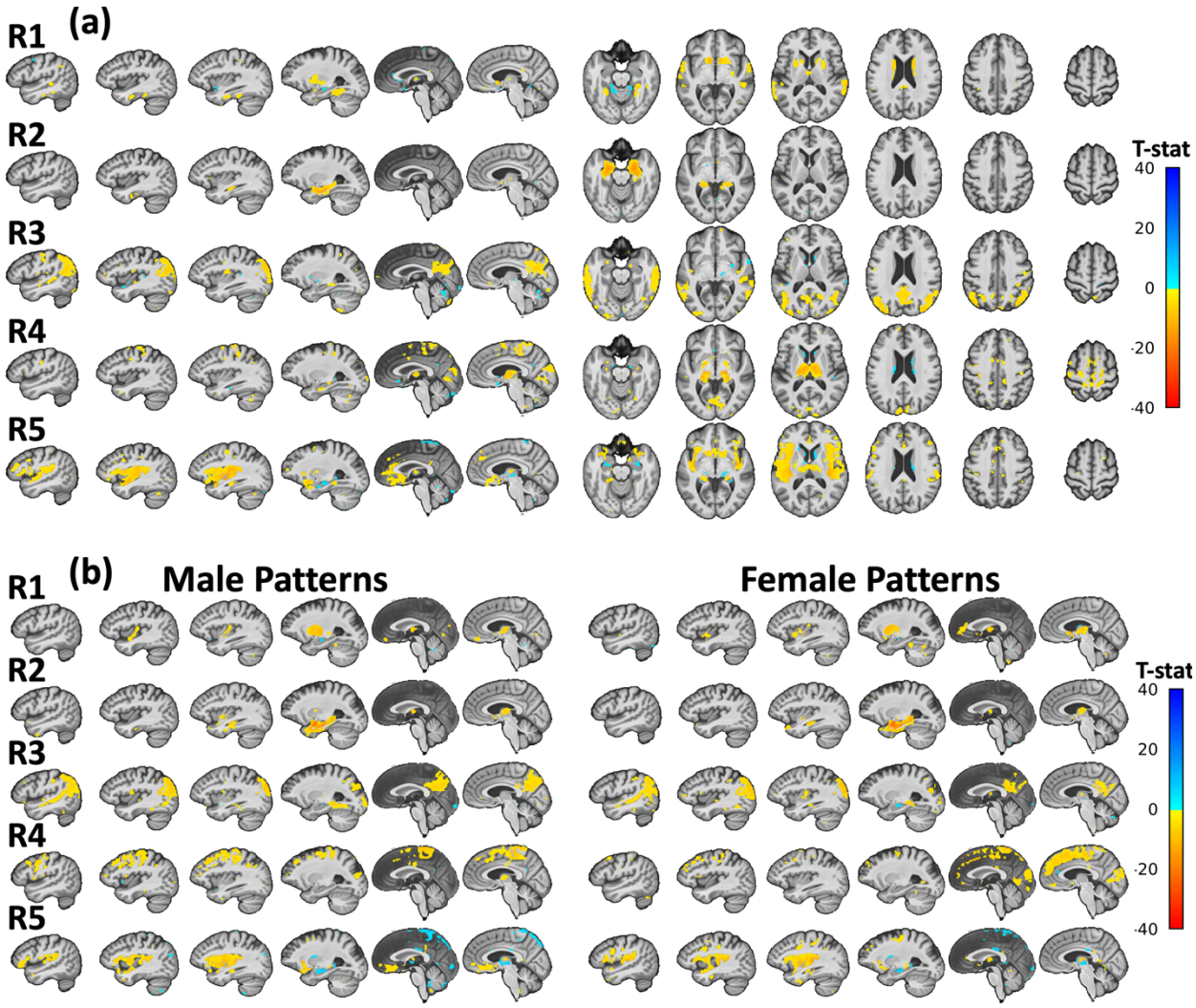
Supplementary Table 2. Performances of models trained on data with different levels of atrophy. "Basic", "Mild", and "Extra Mild" models represent models trained on semi-synthetic data with 0-30%, 0-20%, and 0-15% rates of simulated atrophy, respectively. Reducing the brain atrophy rate in semi-synthetic data resulted in a decline in model performance. However, models trained on datasets with the same patterns but a larger atrophy rate tended to perform better when tested on datasets with a maximum atrophy rate of only 15%. This implies the model's ability to better capture the ground truth when trained on datasets containing participants with a higher degree of atrophy.

Dataset	Semi-synthetic Data			Aging data
	0-15% atrophy	0-20% atrophy	0-30% atrophy	
z-scores of TAR ROIs	-0.188±1.05	-0.238±1.09	-0.353±1.22	-0.649±1.01

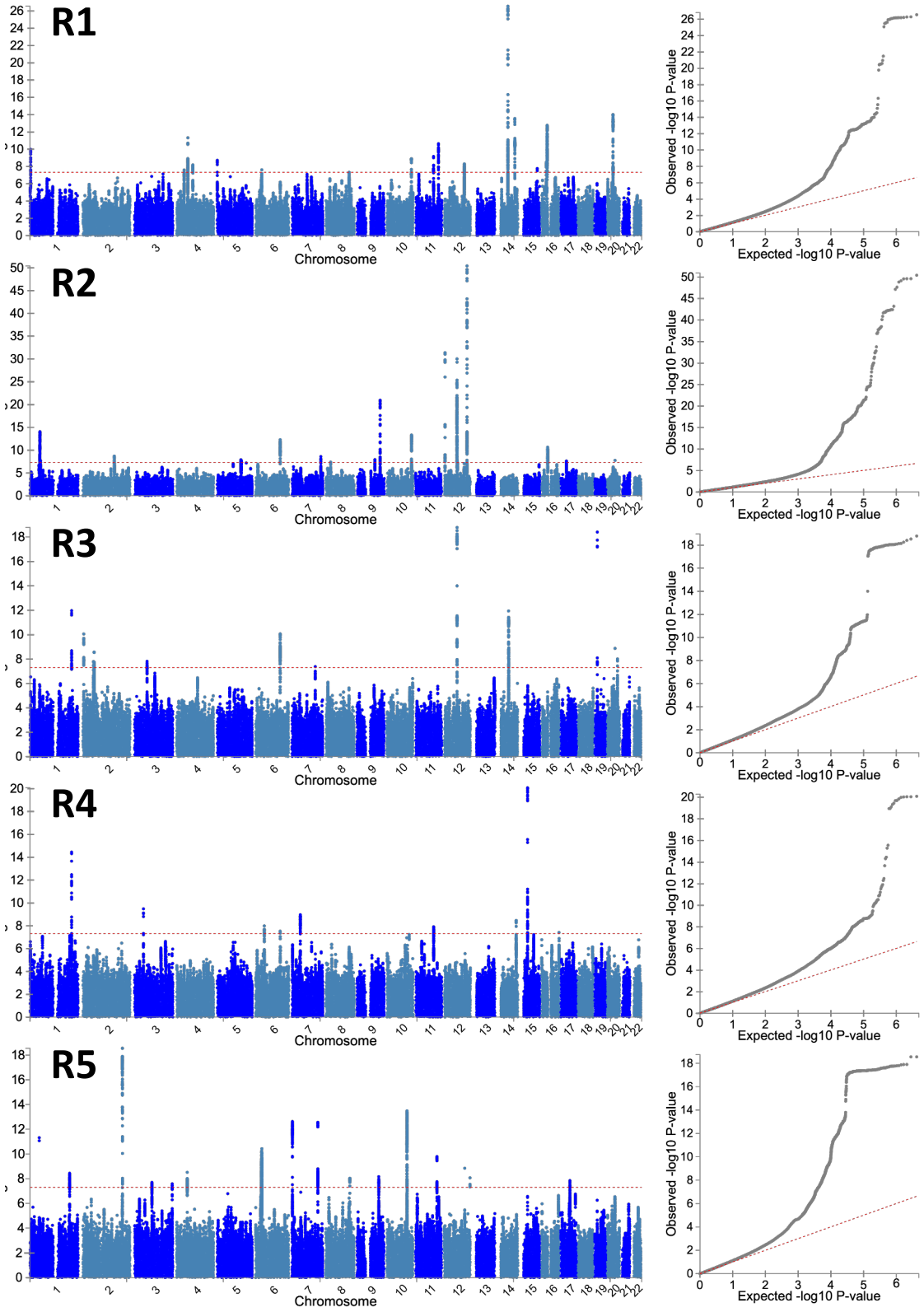
Supplementary Table 3. Quantification of atrophy rate in real and semi-synthetic datasets. Within the aging population, the z-scores of TAR group's ROIs displayed a mean value of -0.649±1.01, significantly higher than those of the semi-synthetic datasets, suggesting the presence of a strong signal that can potentially enable the model to accurately capture the underlying truth patterns.



Supplementary Figure 1. The improved Surreal-GAN significantly outperforms the original version. With increased correlations simulated among the ground truth R-indices, the original Surreal-GAN shows dramatically decreased model performances. The performance degradation is even more pronounced when correlations are introduced among all three dimensions. In contrast, the improved one demonstrates robustness, proving its capability in capturing the correlations among the ground-truth brain change patterns. Fifty Rindex-correlation values were derived for each case with 50 repeated runs. (Centerline, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range; points, outliers)



Supplementary Figure 2. Surreal-GAN identifies five dimensions with replicable imaging signatures. **(a)** Voxel-wise t-tests were performed for the five dimensions rederived using an independent replication training set, while adjusting for age, sex, intracranial volume (ICV), and the remaining four R-indices. The replicated five dimensions are associated with consistently reproducible brain changes (**Figure 1a**). This is supported by a quantitative analysis, where the replication R-indices exhibit a high Rindices-Correlation of 0.752 with the original R-indices. **(b)**Voxel-wise t-tests were performed for the male- and female-specific dimensions within their respective sex populations. For both (a) and (b), False discovery rate (FDR) correction was performed to adjust multiple comparisons with a p-value threshold of 0.001. The identified brain atrophy patterns generally replicate those revealed in **Figure 1a**.



Supplementary Figure 3. Manhattan and QQ plots of baseline GWAS results for R1-R5 in the UKBB population. Two-sided t-tests were applied for testing the significance of regression coefficients of SNPs.