

Appendix S1

SUPPLEMENTARY MATERIALS AND METHODS

Statistical analysis

Logistic regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). It is a predictive analysis and is used to describe data and to explain the relationship between one dependent binary variable and 1 or more nominal, ordinal, interval or ratio-level independent variables. The output from the logistic regression model is the estimated probability, which is used to indicate the level of confidence that a predicted value is the actual value when given an input X. Stepwise selection with a combination of the forward and backward selection techniques was applied using the likelihood-based Wald test. Both forward and backward techniques converge on the same variables, which were those included in the final model. Accordingly, the specific model was as follows:

$$\text{Score} = 100 * \left(1 - \frac{e^y}{1 + e^y} \right)$$

The model allowed us to anticipate whether the real age of an individual was higher or lower than the age predicted with the Glogau scale from a set of independent variables (the exposome factors) that were associated with a *p*-value less than 0. To calculate the score for predicting skin aging, the logistical equation described above must be solved taking into account that:

$$Y = -0.116 \times \text{age (years)} - 0.354 \times \text{smoking habit} + 0.18 \times \text{use of sunscreens} - 0.055 \times \text{use of cosmetics} + 4.35;$$

Where:

- Smoking habit: smoker = 1; ex-smoker = 2; non-smoker = 3
- Use of sunscreens: daily = 1; occasionally = 2; only when sunbathing = 3; never = 4.
- Use of cosmetics: daily cleansing = 1; moisturising lotion = 2; moisturising lotion + daily cleansing = 3; anti-aging cream = 4; anti-aging + daily cleansing = 5; anti-aging + moisturising lotion = 6; anti-aging cream + moisturising lotion + daily cleansing = 7.