

validation datasets (e.g., across ancestries, locations, or years) and perform systematic sensitivity analyses, such as replacing LD reference panels or excluding high-LD regions (Wang et al., 2023).

In addition, cross-population applications must account for data heterogeneity, including inconsistencies in phenotype definitions, missing environmental records, and platform-specific batch effects, all of which may amplify extrapolation error. Therefore, external validation design should incorporate standardized data processing, multidimensional data recording, and cross-platform harmonization to improve interpretability and reproducibility.

### 2.3 Performance Evaluation Metrics

The effectiveness of PRS/PGS models should be assessed using a multi-dimensional framework that captures predictive accuracy, calibration, and practical decision value. For binary outcomes (e.g., disease status), the area under the receiver operating characteristic curve (AUC) is a primary measure of discrimination (Lennon et al., 2024). For continuous traits, the proportion of variance explained ( $R^2$ ) and correlation coefficients ( $r$ ) reflect predictive accuracy, and can be transformed to the liability scale ( $R^2_{\text{liability}}$ ) for cross-population comparability (Sima et al., 2024).

In terms of effect size, odds ratios (OR) or hazard ratios (HR) per standard deviation of PRS are commonly reported in clinical studies (Patel et al., 2023). Cross-population performance can be quantified using the transferability ratio  $T = R^2_{\text{target}}/R^2_{\text{source}}$  and differences in AUC (Duncan et al., 2019). From a statistical interpretation perspective,  $R^2$  reflects variance explained by genetic signals, whereas AUC reflects the model's ability to rank individuals, corresponding to "variance explanation" and "predictive discrimination," respectively.

At the application level, threshold selection should be guided by decision context. In medical studies, decision curve analysis (DCA) can be used to determine clinically meaningful thresholds based on disease prevalence and intervention costs. For example,  $OR \approx 1.5$  with good calibration may support enhanced risk screening, whereas  $OR \geq 2.0$  may justify early intervention for high-risk individuals (Jung et al., 2025). In crop breeding, threshold optimization should consider heritability, environmental heterogeneity, and economic weights, and can be informed by simulations or multi-environment trial data.

To avoid over-reliance on single metrics, it is recommended to report uncertainty intervals, sensitivity analyses, and calibration curves, thereby providing a more robust and comprehensive evaluation of model performance.

## 3 Joint Prediction of PRS with Environmental and Lifestyle Factors

Within the framework of statistical genetics, PRS fundamentally captures an individual's baseline genetic risk under given genetic information. However, in real-world systems, phenotypes are jointly determined by genetic effects and environmental exposures. Therefore, jointly modeling PRS with environmental or lifestyle factors can be interpreted as a conditional extension of the predictive functional, in which environmental dependence is introduced on top of genetic main effects, thereby improving both predictive accuracy and decision utility.

### 3.1 Modeling gene-environment interaction ( $G \times E$ )

In multi-environment trials (METs) or longitudinal population studies, gene-environment interaction ( $G \times E$ ) determines the stability and transferability of predictive functions across different environmental conditions. By using PRS/PGS as a low-dimensional representation of genetic main effects and incorporating environmental variables (e.g., site, year, climate, or lifestyle exposures), a reaction norm framework can be constructed:

$$y_{ij} = \mu + \beta_g \cdot PGS_i + \beta_e \cdot E_j + \beta_{ge} \cdot (PGS_i \cdot E_j) + g_i + (gE)_{ij} + \varepsilon_{ij}$$

where  $g_i$  and  $(gE)_{ij}$  denote random genetic effects and their interaction with the environment, respectively. These can be modeled using factor-analytic (FA) or kernel-based covariance structures to capture cross-environment correlations. In high-dimensional environmental settings, dimensionality reduction techniques such as principal component analysis or ecological indices are typically applied to distinguish general adaptability from specific adaptability.