

For time-varying exposures, piecewise time-varying coefficient models or joint longitudinal-survival models can be used to mitigate bias arising from temporal misalignment (Kachuri et al., 2024).

Machine learning methods provide complementary tools for capturing nonlinearities and higher-order interactions. Random forests and gradient boosting machines (GBM) are robust to data heterogeneity and missingness, while deep learning models are suitable for integrating multi-source phenotypic data (e.g., high-throughput phenotyping or wearable device data). However, from a statistical inference perspective, increased model flexibility often comes at the cost of higher estimation bias and greater uncertainty in generalization. Therefore, nested cross-validation, class imbalance reweighting, and post hoc calibration (e.g., Platt scaling or isotonic regression) are essential to control overfitting (Sima et al., 2024).

For model interpretability, SHAP values or permutation importance can be used to quantify the marginal contributions of PRS, key environmental variables, and their interactions. Stability selection across different ancestries or ecological populations can further help identify predictors with consistent effects.

### 3.3 Predictive gain and decision utility of joint models

The improvement in predictive performance achieved by joint modeling of PRS and environmental factors can be quantified from a variance decomposition perspective:

$$\Delta R^2 = R_{\text{joint}}^2 - R_G^2$$

where  $R_G^2$ ,  $R_E^2$ , and  $R_{GE}^2$  represent the contributions of genetic main effects, environmental main effects, and interaction effects, respectively, and  $\Delta R^2$  reflects the incremental gain of the joint model over PRS alone.

The magnitude of this gain depends on the variability of environmental factors and the strength of G×E interactions. When environmental exposures are modifiable and widely distributed in the target population, joint models often produce steeper risk gradients in the high-risk tail, thereby increasing net benefit in decision curve analysis (Kachuri et al., 2024; Sima et al., 2024).

In population-based studies, the evaluation of PRS is more appropriately conducted within a comparative modeling framework rather than relying on a single model specification. A baseline model incorporating key covariates, such as age, sex, and ancestry principal components, is typically used as a reference, upon which PRS, environmental exposures, and their interaction terms are incorporated. Model performance can then be assessed in a multidimensional manner, capturing both discrimination and explanatory capacity, using metrics such as AUC,  $R^2$  (or liability-scale  $R^2$ ), effect sizes per standard deviation of PRS (e.g., OR or HR), and net benefit derived from decision curve analysis. Furthermore, stratifying individuals according to lifestyle or environmental exposure levels allows for a more nuanced characterization of risk distributions and calibration patterns, thereby facilitating the evaluation of the “risk equivalence” phenomenon—namely, the extent to which adverse environmental exposures may attenuate or offset the protective effects associated with lower genetic risk (Sima et al., 2024).

In crop breeding, a two-stage decision framework can be implemented: early-generation selection using PGS to improve selection accuracy (correlation  $r$ ), followed by estimation of environment-specific interaction effects ( $\beta_{ge}$ ) in small samples from target environments. This enables a strategy of “general adaptability selection + environment-specific optimization,” with theoretical genetic gain expressed as:

$$\Delta G \approx i \cdot r \cdot \sigma_A$$

where  $i$  denotes selection intensity and  $\sigma_A$  the additive genetic standard deviation (Ruan et al., 2021). This framework demonstrates that integrating PRS with environmental information not only improves predictive accuracy but also translates directly into higher selection efficiency and practical utility.

## 4 Ethical and Population Fairness Issues

In the practical application of polygenic risk scores (PRS/PGS), ethical and population fairness concerns arise not only from social structural differences but are also deeply rooted in differences in the applicability of statistical