A whole leukocyte-DNA methylome risk score to predict pancreatic cancer

Evangelina López de Maturana[1]
Lola Alonso[1] and Núria Malats[1]
[1] Genetic and Molecular Epidemiology Group, CNIO and CIBERONC, Spain

Background & objectives

Pancreatic cancer (PC) has a low 5-year survival rate mainly because of late diagnosis. To earlier diagnose PC is necessary to discover new biomarkers to define high-risk populations. DNA methylation (DNAm) in peripheral blood has been barely explored. Here we show the potential of whole leukocyte-DNA methylome risk score (WMRS) to predict PC.

Methods

We followed a two-phase epidemiological design, including in the first-phase 338 cases and 285 controls, and 148 cases and 137 controls in the validation-phase from the PanGenEU. Our study population had epidemiological and DNAm data determined with Infinium MethylationEPIC array from leukocyte-DNA. We applied a single marker and a Bayesian kernel based regression (BKR) to train the model in the discovery phase and then, validate and evaluate the biomarkers and evaluate their predictive ability in the testing set. Both uni/multimarker approaches were adjusted for immune cell composition, age, gender, region of recruitment, diabetes and smoking status.

Results

We identified a signature of 173 CpGs associated with PC risk in the discovery phase using a single marker regression strategy. Among them, 77 (44%) were validated in the phase-2 with an AUC=0.73. An AUC=0.70 was obtained using the BKR in the discovery phase after applying a 10 times 10-fold cross validation approach, and a slightly lower value (AUC=0.66) when it was applied to predict the testing set outcomes.

The application of BKR also enabled us to estimate that the WMRS explains ~15% of the PC risk variance. We also observed that the methylome interaction with diabetes status explained a much larger proportion (~33%) than that with the smoking status (19%). In particular, methylome interaction with long-standing diabetes was greater than that with new-onset (11% vs. 6%). Moreover, the interaction with smoking status explained 19% of the total variance.

Conclusion

For the first time, we evaluated the explored and estimated the variability explained by the main effect of leukocyte-DNA methylome, as well as that of its interactions with the main risk factors for PC. Moreover, we propose a 77 CpG signature showing a good performance to predict PC.