Adapting Breast Cancer Polygenic Risk Scores to Diverse Ethnic Groups in Israel

Hagai Levi[1], Ran Elkon[2], Ron Shamir[1]

[1] Blavatnik School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel

[2] Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

Abstract

Breast cancer (BC) is the most common cancer among women. BC early detection strategy is mammography screens starting at age 50. Recent studies established that women can be stratified according to BC risk using polygenic risk score (PRS). PRS accrues the effect of genomic variants discovered by genome-wide association studies (GWASs). However, sufficiently large GWASs exist only for European (EUR) women. PRS performance declines with the genetic distance between the discovery and target populations. Notably, Ashkenazi Jews (AJ) are genetically close to EUR (Figure 1). Here, we develop methods to adapt EUR-PRSs for Israeli populations, using transfer learning with only a few hundred genotypes from the target population. We expect that the proximity of AJ and EUR makes such adaptation feasible. We applied BC-PRS on a subgroup of AJ "proxies" and healthy AJ cohort (Figurer 2). The results suggest that accurate PRS is attainable for AJ based on extant PRS.
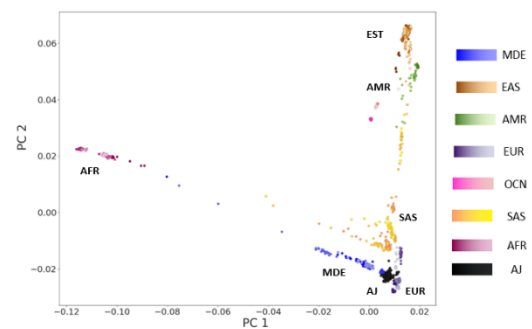
Figure 1: Principal Component Analysis demonstrating genetic distances between main global populations. MDE: middle-east; EAS: East Asia; AMR: Americans; EUR: European; OCN: Oceanians; SAS: South Asia; AFR: Africans; AJ: Ashkenazi Jews.
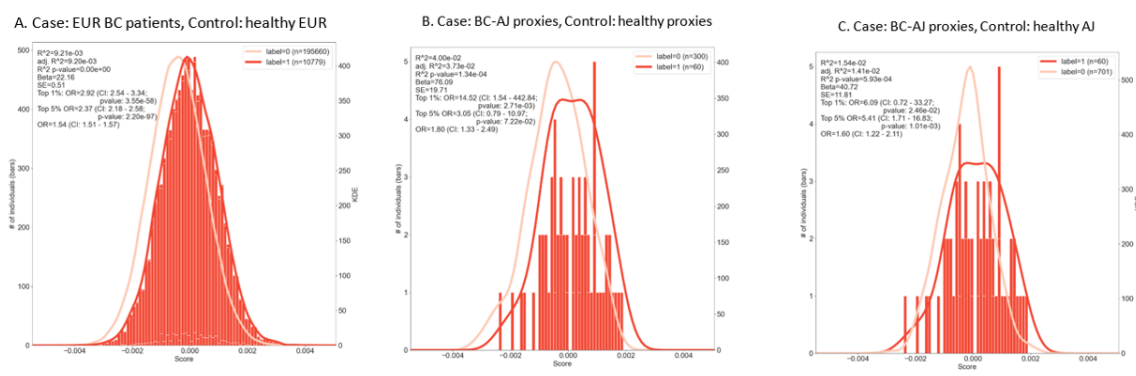

Figure 2. EUR BC-PRS distribution on (**A**) the entire EUR cohort from the UK Biobank (UKB) (**B**) BC-AJ proxies (60 cases) vs. healthy AJ proxies (300 subjects) and (**C**) BC-AJ proxies vs. healthy (real) AJ women. Light: controls; dark: cases. Bars are absolute number of women with BC. Curves are normalized kernel density estimations of each set.