

**Additional file 4 — Accuracy of PCA-SVM using different schemes for handling missing data**

Ethnicity	# Samples	Missing Data Handling		
		rCRS	Probability	Common Region
Caucasian	2807	96.94	96.15	93.94
Asian	915	30.60	50.93	68.85
African	254	59.06	57.48	76.38
Micro-Accuracy	3976	79.25	83.27	87.05
Macro-Accuracy	3976	62.20	68.19	79.72