Title: Development and Validation of the PREMMplus Clinical Prediction Model for Multigene Hereditary Cancer Risk Assessment

Authors: Matthew B. Yurgelun, Hajime Uno, C. Sloane Furniss, Chinedu Ukaegbu, Miki Horiguchi, Anu Chittenden, Tara Fehlmann, Holly LaDuca, Judy Garber, Sapna Syngal

Abstract

Background: Current clinical prediction models provide syndrome-specific numeric estimates of an individual's likelihood of having a specific hereditary cancer syndrome (e.g. PREMM5 for Lynch syndrome; BRCAPRO for *BRCA1/2*) based on clinical data (age, sex, personal/family cancer history). With the emergence of multigene panel testing, however, the field of cancer genetics has rapidly moved away from syndrome-specific genetic testing and thus there is a need to evaluate individuals' risk of a diverse array of cancer predisposition genes/syndromes in parallel. The aim of this study was to develop and validate the PREMMplus clinical prediction model for multigene cancer risk assessment.

Patients and Methods: A cohort of 7296 individuals with prior multigene panel testing at a single institution is being used for model development using polytomous logistic regression to examine candidate predictive variables, including individuals' age, sex, ethnicity, personal and family history of various cancers, including age(s) at diagnosis. For each individual assessed by PREMMplus, a numeric score will be calculated to quantify their likelihood of carrying a pathogenic/likely pathogenic germline variant (mutation) in any of 18 cancer susceptibility genes (11 high-penetrance genes [*APC*, *BRCA1/2*, *CDH1*, *EPCAM*, *MLH1*, *MSH2*, *MSH6*, biallelic *MUTYH*, *PMS2*, and *TP53*] and 7 moderate-penetrance genes [*ATM*, *CDKN2A*, *CHEK2*, *PALB2*, *PTEN*, *RAD51C*, and *RAD51D*]), encompassing a diverse array of cancer susceptibility. An independent dataset of 20,000 individuals undergoing clinical multigene panel testing at a large commercial laboratory will be unlocked by February 26, 2020 for validation of PREMMplus.

Expected Results: We will report metrics of PREMMplus model performance in both the development and validation datasets, including sensitivity, specificity, positive-/negative-predictive values, and number needed to test with PREMMplus in order to identify a mutation carrier. The main outcomes will be the ability of PREMMplus to identify individuals with a mutation in any of 18 cancer susceptibility genes (primary) or any of 11 high-penetrance susceptibility genes (secondary) by the area under the receiver operating characteristic curve (ROC-AUC) and corresponding 0.95 confidence interval. If successful, we expect this to be the first clinically validated prediction model for multigene/multisyndromic hereditary cancer risk assessment.