Study title: Treatment-specific genetic risk scores for late effects prediction in childhood, adolescent, and young adult cancer survivors (NCI award 1R21CA261833-01)

Primary working groups: Genetics; Biostatistics/Epidemiology
Secondary working groups: Second Malignancy; Chronic Disease

Investigators:

Cindy Im, PhD, MPH	cim1@ualberta.ca
Yan Yuan, PhD	yyuan@ualberta.ca
Gregory T. Armstrong, MD, MSCE	greg.armstrong@stjude.org
Daniel A. Mulrooney, MD, MS	daniel.mulrooney@stjude.org
Zhaoming Wang, PhD	zhaoming.wang@stjude.org
Kirsten K. Ness, PT, PhD, FAPTA	kiri.ness@stjude.org
Melissa M. Hudson, MD	melissa.hudson@stjude.org
Leslie L. Robison, PhD	les.robison@stjude.org
Joseph P. Neglia, MD, MPH	jneglia@umn.edu
Michael A. Arnold, MD, MPH	michael.arnold@childrenscolorado.org
Lindsay M. Morton, PhD	mortonli@mail.nih.gov
Eric J. Chow, MD, MPH	ericchow@uw.edu
Smita Bhatia, MD, MPH	sbhatia@peds.uab.edu
Yutaka Yasui, PhD	yutaka.yasui@stjude.org

Background and rationale

Advances in treatments have dramatically improved long-term survival after childhood, adolescent, and young adult (CAYA) cancer to >80%.[1,2] However, curative treatments for CAYA cancer have significant consequences: ongoing evaluations in established cohort studies have shown that survivors face greater risks for a broad range of chronic health conditions (CHCs) compared to general population or sibling controls.[3-6] These risks vary in magnitude with specific therapeutic exposures. For example, female CAYA cancer survivors have ~4-fold to >20-fold greater risk for subsequent breast cancer, chiefly depending on chest radiation therapy (RT) dose.[7,8] Similarly, survivors have up to 6-fold greater risk of cardiomyopathy and myocardial infarction, depending on anthracycline and cardiac-directed RT doses,[9,10] and ~2-fold greater risk of diabetes mellitus, which varies with abdominal RT exposure.[11,12] The inter-individual variation in risk for developing late effects is substantial among survivors with similar therapeutic exposures, suggesting genetic risk factors may contribute.

Considerable progress has been made in developing genetic risk profiles based on polygenic risk scores (PRS), which combine the effects of many germline genetic variants (SNPs) identified in genome-wide association studies (GWAS) into a single score. PRS can improve the clinical stratification of individuals by their risk for many common diseases.[13-17] However, PRS are based on genetic effect size estimates that reflect the biases and statistical imprecision of their respective GWAS.[16,18,19] Preliminary evaluations of PRS derived from general population GWAS in survivors have begun to emerge, but the reported magnitudes of association between PRS and risk for various CHCs are frequently attenuated in cancer survivors.[20-22] These results suggest general population PRS may not be the primary driver of genetic risk for late effects among survivors. Assessments of interaction associations between PRS and cancer treatments[20,22] suggest these types of interactions also do not fully capture the modifying effects of cancer therapies in survivors.

Novel methods to develop PRS that are not be limited to the genetic risk signals detected in the general population may substantially refine late effects risk prediction among survivors. In

this proposal, we aim to "generalize" PRS to CAYA cancer survivors, by adapting recent analytic methods[23-25] that modify PRS based on large GWAS (conducted with general population European ancestry samples) to improve their generalizability to diverse (racial/ethnic minority) populations. These methods involve selecting different SNPs and/or updating SNP weights, leveraging associations from GWAS in much larger (i.e., N>100,000) European samples with results from smaller (N<20,000) GWAS in target ethnic populations. With the availability of whole-genome/whole-exome sequencing (WGS/WES) and genotype data in ~13,000 survivors of the St. Jude Lifetime Cohort Study (SJLIFE) and Childhood Cancer Survivor Study (CCSS), we can develop a comparable methodological approach to create "survivor-enriched" PRS that appropriately account for treatment-specific genetic effects.

The main objective of this proposal is to develop new genetic risk prediction tools based on results from genome-wide gene-treatment (GxT) interaction association analyses conducted in survivor cohorts for a diverse preliminary set of CHCs that are among the leading causes of morbidity and mortality in survivors of CAYA cancer. CHCs of interest include subsequent malignant neoplasms (breast cancer; basal cell carcinoma, multiple subsequent malignant neoplasms) and cardiovascular/endocrine diseases and related complex traits (congestive heart failure/cardiomyopathy; hypertension; coronary artery disease; diabetes mellitus; body mass index; systolic/diastolic blood pressure). For each phenotype, we intend to:

(a) Identify treatment-specific genetic effects in the survivor data (i.e., from genome-wide GxT interaction association analyses);
(b) Develop survivor-enriched PRS that reconciles treatment-specific genetic effects with genetic risk associations included in PRS from published GWAS/meta-analyses with large (i.e., N>50,000) general population samples; and
(c) Compare the predictive ability of both general population PRS and survivor-enriched PRS models against other clinical/genetic risk prediction models.

The proposed research is significant in its potential to help identify survivors at high risk for adverse health conditions with personalized genetic risk profiles, enhancing the current treatment risk-based guidelines for survivorship care.

Specific hypotheses and aims

We hypothesize that survivor-enriched PRS can better predict CHC risks (or related trait variation) in CAYA cancer survivors by considering treatment-specific genetic effects.

Specific Aim 1 (Primary): Create a comprehensive catalog of treatment-specific genetic effects for selected CHCs among CAYA cancer survivors.

- We plan to perform genome-wide association analyses to evaluate treatment-specific genetic effects in SJLIFE/CCSS survivors for each CHC, including subsequent neoplasms and specific cardiovascular and endocrine diseases. The identified treatment-specific genetic effects will inform the development of survivor-enriched PRS (Aim 2).

Specific Aim 2 (Primary): Construct novel survivor-enriched PRS for the selected CHCs that appropriately account for treatment-specific genetic effects, and provide unbiased evaluations of their prediction performance in an independent cohort of SJLIFE/CCSS survivors.

- Our approach will create PRS models that include established genetic risk signals seen in the general population and results from GWAS performed in survivors (Aim 1). The improvement in prediction performance by using survivor-enriched PRS will be

evaluated over models with survivor-specific clinical risk factors and published (general population) PRS.

Specific Aim 3 (Secondary): Investigate whether survivor-enriched PRS improves CHC risk prediction in racial/ethnic minority survivor subgroups.

- PRS based on European ancestry GWAS generalize poorly to racial/ethnic minority populations. Since treatment-specific genetic effects may be trans-ethnic, survivor-enriched PRS may be more informative in predicting CHCs in these subgroups.

Analysis Plan

Research Strategy

Our strategy is to develop an approach to develop prediction models featuring PRS for 10 CHCs/complex traits (Table 1), incorporating both published results from general population GWAS and survivor-enriched PRS, which appropriately account for treatment-specific genetic effects observed among survivors. We also aim to provide unbiased evaluations of the prediction performance of survivor-enriched PRS.

Table 1: Chronic health conditions in SJLIFE/CCSS

Chronic Health Conditions (CHCs)	SJLIFE (N=4,402)	CCSS (N=8,739)	Reference GWAS
Subsequent Neoplasms	**% cases (N)**	**% cases (N)**	
Breast cancer	1.9% (40) N=2,090 females	4.4% (197) N=4,525 females	PMID: 29059683 (N=228,951)
Basal cell carcinoma	3.3% (146)	7.3% (641)	PMID: 31427789 (N=60,692)
Multiple subsequent, malignant	1.2% (52)	5.1% (450)	PMID: 31427789 (N=386,581)
Cardiovascular	**% cases (N)**	**% cases (N)**	
Hypertension	20.3% (893)	19.4% (1,692)	PMID: 31427789 (N=289,307)
Coronary artery disease	3.3% (146)	3.0% (259)	PMID: 26343387 (N=184,305)
Congestive heart failure or cardiomyopathy	9.0% (397)	5.2% (454)	PMID: 30586722 (N=390,142)
Endocrine	**% cases (N)**	**% cases (N)**	
Diabetes mellitus	19.8% (872)	6.3% (549)	PMID: 28566273 (N=159,208)
Complex traits	**median (IQR)**	**% cases (N)**	
Body mass index (kg/m^2)	23.1 (18.5-28.8)	-	PMID: 25673413 (N=234,069)
Systolic blood pressure (mmHg)	118 (110-126)	-	PMID: 31427789 (N=up to 361,411)
Diastolic blood pressure (mmHg)	68 (63-74)	-	

Our analysis plan entails three major tasks (Figure 1):

(1) *Resource Building*, to select suitable reference GWAS resources for baseline comparison PRS and to create the treatment-specific genetic effects catalog with results from genome-wide GxT interaction association analyses in survivors (Aim 1);
(2) *PRS Development and Tuning*, to create survivor-enriched PRS incorporating treatment-specific genetic effects (Aim 2); and
(3) *Model Testing*, to provide robust performance evaluations of survivor-enriched PRS in independent data (Aim 2).

This analysis plan requires up to three pre-specified CAYA cancer survivor datasets. The first survivor dataset will be dedicated to discovering treatment-specific genetic effects. For some phenotypes, we may reserve a second survivor dataset for PRS "tuning", which involves evaluating prediction metrics for a range of hyperparameters to select the best-performing PRS (e.g., assuming various selections for p-value and genetic variant linkage disequilibrium cut-offs). For phenotypes that are only measured in SJLIFE (e.g., blood pressure), both GxT interaction GWAS and PRS tuning will be performed in the same dataset using a cross-validation algorithm.

A final survivor dataset will be set aside strictly for testing the relative predictive performance of tuned survivor-enriched PRS, general population PRS, and appropriate clinical models in both European and racial/ethnic minority survivor subgroups (Aim 3).

Study Population

For the phenotype-specific genome-wide GxT interaction association analyses, we will use survivor cohort data with greater discovery statistical power. We plan to primarily use the CCSS original cohort for discovery analyses for second cancer-related phenotypes, since CCSS has more participants with long-term follow-up and second cancers are confirmed via pathology findings or medical record review. For all other phenotypes, we intend to use SJLIFE for discovery analyses since most phenotypes are clinically ascertained in SJLIFE. For some analyses, we will consider combining the array-based CCSS original cohort genotype data with whole genome sequencing data for the SJLIFE and CCSS expansion cohorts to generate up to three separate randomly-sampled datasets for PRS development and evaluation (i.e., for discovery GWAS, PRS selection/tuning, validation).

Outcomes of interest:

Outcomes of interest are the ten CHCs/complex traits described in Table 1. For disease phenotypes, we will use the chronic disease CTCAE definitions in CCSS and SJLIFE. For complex traits, we plan to use SJLIFE campus visit clinical assessment data. Minimum CTCAE grades for disease phenotypes that will be considered in each phenotype-specific analysis will be carefully considered by the proposal investigators, and will generally start at grades assigned for signs of symptomatic disease and/or use of medications for CHC treatment.

Subject population:

The study population will include up 13,141 long-term survivors of childhood cancer with genotype data enrolled in either CCSS or SJLIFE. Genetic ancestry has been previously determined via principal components analysis. We plan to perform analyses in both European and non-European ancestral subgroups. For all phenotypes, we will only include study participants with both genotype and phenotype data, and who also meet inclusion criteria for phenotype-specific analyses (e.g., limiting breast cancer analyses to female survivors only) as well as inclusion criteria typically applied in GWAS (i.e., meets missingness, heterozygosity, sex discordance, and relatedness thresholds).

Explanatory variables:

The primary explanatory variables of interest are polygenic risk scores comprised of common (>1% minor allele frequency) imputed or sequenced genetic variants identified in genome-wide GxT interaction association analyses performed in survivor cohort datasets.

Additional covariates:

- Study cohort participation status (SJLIFE/CCSS)
- Sex
- Age at last follow-up/contact
- Age at death
- Cancer diagnosis group
- Year of cancer diagnosis

- Age at diagnosis
- Subsequent breast cancer (Yes/No); include diagnosis ages for all occurrences
- Subsequent basal cell carcinoma (Yes/No); include diagnosis ages for all occurrences
- Multiple subsequent malignant neoplasms (Yes/No): excluding BCC; include diagnosis ages for all occurrences
- Congestive heart failure/cardiomyopathy: all CTCAE grades ≥3, including grade onset age
- Coronary artery disease: all CTCAE grades ≥3, including grade onset age
- Hypertension: all CTCAE grades ≥2, including grade onset age
- Diabetes mellitus (abnormal glucose metabolism): all CTCAE grades ≥2, including grade onset age
- Dyslipidemia: all CTCAE grades ≥2, including grade onset age
- Genetic ancestry (calculated via principal components analysis)
- Height
- Weight
- Body mass index
- Smoking history
- Alcohol use history
- Physical activity (CDC-based definition)
- Radiation therapy, total prescribed dose to any of seven major regions: head, neck, chest, abdomen, pelvis, legs, arms (Yes/No; dose)
- Total body irradiation (Yes/No; dose)
- Allogeneic bone marrow transplant history (Yes/No)
- Hypothalamic-pituitary axis RT (Yes/No; dose)
- Heart RT (Yes/No; dose)
- Alkylating agents (Yes/No; dose [cyclophosphamide ED])
- Anthracyclines (Yes/No; dose [doxorubicin ED])
- Heavy metals (Yes/No; dose [cisplatin ED])
- Epipodophyllotoxin (Yes/No; dose)
- Vinca alkaloids (Yes/No)
- Ifosfamide (Yes/No)
- Glucocorticoids (Yes/No)
- Splenectomy history (Yes/No)
- Nephrectomy history (Yes/No)

Analytic approach:

Aim 1: Identify and catalog treatment-specific genetic effects in survivors of CAYA cancer.

Methods and data analysis*:* For each of the 10 phenotypes, we will: (a) characterize etiological pathways and evaluate relevant clinical risk factors and their associations with the CHC/trait; and (b) identify treatment-specific genetic effects in the survivor data that are set aside for resource building.

- Etiological pathways: In order to define relevant clinical risk factors for each specific CHC/trait and determine how they should be modeled statistically in downstream genetic analyses, a necessary first step is to characterize the many potential etiological pathways driving therapy-associated CHC/trait risks. Sex, ancestry (race/ethnicity), attained age, CAYA cancer diagnosis, age at CAYA cancer diagnosis, and treatment exposures and doses will be considered as key biological variables, as appropriate. Published risk prediction models in the survivorship literature will also be considered, as appropriate.

- Discovering treatment-specific genetic effects: We will use Cox regression for time-to-event phenotypes or generalized linear models[26] (GLMs) for continuous and prevalence-binary phenotypes. We will test for GxT interaction associations on a genome-wide scale, including using recently developed "2-stage" approach with a screening stage followed by an interaction testing stage, to increase power while controlling type I error.[27,28] We will also consider analyses stratified by treatment exposures (and sex or other key biological variables, if appropriate) coupled with tests for the heterogeneity of stratum-specific genetic effects.[29] These genome-wide association analyses will be conducted with all measured or imputed common (minor allele frequency >1%) biallelic genetic variants, assuming an additive genetic inheritance model.

Aim 2: Develop and evaluate survivor-enriched PRS that appropriately account for treatment-specific genetic effects seen in survivors of CAYA cancer.

Methods and data analysis: For each of the selected phenotypes, we will: (a) identify relevant published "baseline" (general population) PRS; (b) innovate an approach to create survivor-enriched PRS; and (c) compare the predictive ability of survivor-enriched PRS against other clinical/genetic risk prediction models.

- Baseline genome-wide PRS: Genome-wide PRS derived from current methods with suitable reference GWAS/meta-analyses (Table 1) will serve as baseline PRS.
- Survivor-enriched PRS: We propose to: (a) directly apply and/or (b) adapt existing methods[23-25] that have been shown to improve the trans-ethnic generalizability of PRS. Direct applications of these methods involve treating subgroups of survivors with specific treatments like a distinct ethnic population. The potential drawback is that these methods would effectively re-weight every selected genetic variant with association estimates from GWAS conducted in survivors, dominating the final survivor-enriched PRS. Thus, we propose a key adaptation: re-weight SNPs only if they exhibit both statistical evidence for treatment effect modification in survivor cohort analyses and *in silico* evidence of regulatory or functional activity.
- PRS tuning and testing: We intend to "develop", "tune", and "test" (three key steps of machine learning and prediction model-building) the survivor-enriched PRS in independent survivor datasets using up to 3 major splits of SJLIFE/CCSS data. Using established performance metrics (see below), prediction models including survivor-enriched PRS will be compared to models with: (a) survivor-specific clinical risk factors only (e.g., therapy-based risk scores), and (b) baseline genome-wide PRS.
- Performance metrics: Association statistics (e.g., odds or hazard ratio, linear regression coefficient) will be evaluated to compare disease risk or mean SD change between ranked non-overlapping tiers of individuals with PRS meeting pre-specified percentile cut-off values (e.g., top 1%, top 10%, top 20%). Incremental R^2 is another prediction reliability statistic,[30] defined by the increment in R^2 (or Nagelkerke's R^2 for binary phenotypes) upon adding PRS to a (nested) model of other covariates. Established prediction (discrimination and calibration) metrics[31] such as the mean squared error (MSE) for continuous phenotypes, and the area under [receiver operating characteristic or ROC] curve (AUC) and Brier score for binary phenotypes will also be evaluated.

Aim 3 (Secondary): Evaluate the prediction performance of survivor-enriched PRS for selected phenotypes in survivors from racial/ethnic minority populations.

Methods and data analysis: We will evaluate survivor-enriched PRS developed in survivors of European ancestry in non-Hispanic Black and Hispanic survivor subgroups, using the methods described above.

Proposed Figures and Tables

Table 1: Demographic and clinical characteristics of childhood cancer survivors in training/validation datasets

	Training		Validation	
Variable	**No.**	**%**	**No.**	**%**
Gender Male Female				
Age at evaluation (years) 18-25 26-35 36-45 46-55 >55				
Body mass index (kg/m^2) <25 ≥25 and <30 ≥30 and <35 ≥35 and <40 ≥40				
Smoking history Ever smoker (yes)				
Physical activity Met CDC physical activity recommendations				
Age at diagnosis (years) 0-4 5-9 10-14 >14				
Diagnosis Leukemia Hodgkin lymphoma Non-Hodgkin lymphoma Central nervous system malignancy Kidney Neuroblastoma Soft tissue sarcoma Other malignancy				
Relevant cancer therapy exposures (Depends on phenotype)				
Phenotype Distribution or case status				

Table 2: Genetic variants in survivor-enriched PRS, adjusted univariate gene-by-treatment interaction associations with phenotype risk

Survivor PRS type	PRS LD/P thresholds	rsID	Chr	BP	NEA	EA	EAF	log OR PRS weights (95% CI)	P	Nearest gene, within 5 kb

Abbreviations: PRS, polygenic risk score; LD, linkage disequilibrium; rsID, genetic variant identifier (dbSNP build 151); Chr, chromosome; BP, base position, GRCh38 (hg38) build; NEA, non-effect (reference) allele; EA, effect (risk) allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, p-value; kb, kilobases.

Table 3: Prediction performance of phenotype risk models in training and validation samples

Dataset	Risk prediction model	OR (95% CI)	OR P	Nagelkerke's R^2	AUC (95% CI)	AUC P	AP (95% CI)
Training	Clinical model						
	General population PRS						
	Survivor-enriched PRS						
	Composite PRS General population PRS Survivor-enriched PRS						
Validation	Clinical model						
	General population PRS						
	Survivor-enriched PRS						
	Composite PRS General population PRS Survivor-enriched PRS						

Abbreviations: PRS, polygenic risk score; OR, odds ratio; CI, confidence interval; P, p-value; AUC, area under the receiver operating characteristic curve; AP, area under the precision-recall curve.

References:

1. Armstrong GT, Chen Y, Yasui Y *et al.* Reduction in late mortality among 5-year survivors of childhood cancer. ***N Engl J Med***. 2016;374(9):833-42.
2. Howlader N, Noone A, Krapcho M *et al.* SEER Cancer Statistics Review, 1975-2016. Bethesda, MD: National Cancer Institute. 2019.
3. Bhakta N, Liu Q, Ness KK *et al.* The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). ***Lancet***. 2017;390(10112):2569-82.
4. Robison LL & Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. ***Nat Rev Cancer***. 2014;14(1):61.
5. Oeffinger KC, Mertens AC, Sklar CA *et al.* Chronic health conditions in adult survivors of childhood cancer. ***N Engl J Med***. 2006;355(15):1572-82.
6. Hudson MM, Ness KK, Gurney JG *et al.* Clinical ascertainment of health outcomes among adults treated for childhood cancer. ***JAMA***. 2013;309(22):2371-81.
7. Henderson TO, Moskowitz CS, Chou JF *et al.* Breast cancer risk in childhood cancer survivors without a history of chest radiotherapy: a report from the Childhood Cancer Survivor Study. ***J Clin Oncol***. 2016;34(9):910.
8. Moskowitz CS, Chou JF, Wolden SL *et al.* Breast cancer after chest radiation therapy for childhood cancer. ***J Clin Oncol***. 2014;32(21):2217.
9. Mulrooney DA, Yeazel MW, Kawashima T *et al.* Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. ***BMJ***. 2009;339:b4606.
10. Bates JE, Howell RM, Liu Q *et al.* Therapy-related cardiac risk in childhood cancer survivors: an analysis of the childhood cancer survivor study. ***J Clin Oncol***. 2019;37(13):1090.
11. Meacham LR, Sklar CA, Li S *et al.* Diabetes mellitus in long-term survivors of childhood cancer: increased risk associated with radiation therapy: a report for the childhood cancer survivor study. ***Arch Intern Med***. 2009;169(15):1381-8.
12. Friedman DN, Moskowitz CS, Hilden P *et al.* Radiation Dose and Volume to the Pancreas and Subsequent Risk of Diabetes Mellitus: A Report from the Childhood Cancer Survivor Study. ***J Natl Cancer Inst***. 2020;112(5):525-32.
13. Natarajan P, Young R, Stitziel NO *et al.* Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. ***Circulation***. 2017;135(22):2091-101.
14. Mega JL, Stitziel NO, Smith JG *et al.* Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. ***Lancet***. 2015;385(9984):2264-71.
15. Khera AV, Chaffin M, Aragam KG *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. ***Nat Genet***. 2018;50(9):1219.
16. Torkamani A, Wineinger NE & Topol EJ. The personal and clinical utility of polygenic risk scores. ***Nat Rev Genet***. 2018;19(9):581.
17. Lambert SA, Abraham G & Inouye M. Towards clinical utility of polygenic risk scores. ***Hum Mol Genet***. 2019;28(R2):R133-R42.
18. Chatterjee N, Shi J & García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. ***Nat Rev Genet***. 2016;17(7):392.
19. Young AI, Benonisdottir S, Przeworski M & Kong A. Deconstructing the sources of genotype-phenotype associations in humans. ***Science***. 2019;365(6460):1396-400.
20. Wang Z, Liu Q, Wilson CL *et al.* Polygenic determinants for subsequent breast cancer risk in survivors of childhood cancer: The St Jude Lifetime Cohort Study (SJLIFE). ***Clin Cancer Res***. 2018;24(24):6230-5.

21. Im C, Qin N, Wang Z *et al.* Generalizability of "GWAS hits" in clinical populations: Lessons from childhood cancer survivors. ***Am J Hum Genet***. 2020;107(4):636-53.
22. Sapkota Y, Li N, Pierzynski J *et al.* Contribution of polygenic risk to hypertension among long-term survivors of childhood cancer. ***JACC: CardioOnc***. 2021;3(1):76-84.
23. Grinde KE, Qi Q, Thornton TA *et al.* Generalizing polygenic risk scores from Europeans to Hispanics/Latinos. ***Genet Epidemiol***. 2019;43(1):50-62.
24. Coram MA, Fang H, Candille SI, Assimes TL & Tang H. Leveraging multi-ethnic evidence for risk assessment of quantitative traits in minority populations. ***The American Journal of Human Genetics***. 2017;101(2):218-26.
25. Márquez-Luna C, Loh PR, Consortium SATD, Consortium STD & Price AL. Multiethnic polygenic risk scores improve risk prediction in diverse populations. ***Genet Epidemiol***. 2017;41(8):811-23.
26. McCullagh P & Nelder JA. Generalized Linear Models 2nd Edition Chapman and Hall. ***London, UK***. 1989.
27. Kooperberg C & LeBlanc M. Increasing the power of identifying gene× gene interactions in genome-wide association studies. ***Genet Epidemiol***. 2008;32(3):255-63.
28. Murcray CE, Lewinger JP & Gauderman WJ. Gene-environment interaction in genome-wide association studies. ***Am J Epidemiol***. 2009;169(2):219-26.
29. Mägi R & Morris AP. GWAMA: software for genome-wide association meta-analysis. ***BMC Bioinformatics***. 2010;11(1):288.
30. Martin AR, Kanai M, Kamatani Y *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. ***Nat Genet***. 2019;51(4):584.
31. Gail MH & Pfeiffer RM. On criteria for evaluating models of absolute risk. ***Biostatistics***. 2005;6(2):227-39.