Genetic Liability and Risk Prediction for Psychiatric Symptoms in Survivors of Childhood Cancer

CCSS Working Groups: Genetics; Psychology/Neuropsychology; Biostatistics/Epidemiology

Investigators

Sarah Benstock	sarah.benstock@stjude.org
Kendrick Li	kendrick.li@stjude.org
Kumar Srivastava	kumar.srivastava@stjude.org
Ellen van der Plas	evanderplas@uams.edu
Kevin Krull	kevin.krull@stjude.org
Greg Armstrong	greg.armstrong@stjude.org
Wendy Leisenring	wleisenr@fredhutch.org
Rebecca Howell	rhowell@mdanderson.org
Vikki Nolan	vikki.nolan2@stjude.org
Monica Gramatges	gramatge@bcm.edu
Yadav Sapkota	yadav.sapkota@stjude.org
Tara Brinkman	tara.brinkman@stjude.org

Background

Advances in diagnosis and treatment have resulted in a 5-year survival rate over 85% and an increasing population of survivors[1]; however, survivors remain at increased risk for experiencing long-term adverse health outcomes, including psychological distress and psychiatric disorders[2–4]. Childhood cancer survivors are 1.6 times more likely to have a psychiatric disorder compared to their peers who have not experienced cancer and just over 4 times more likely to have a psychiatric disorder compared to adult cancer survivors[4]. Furthermore, estimated population prevalences for depression and anxiety are slightly higher in survivors than the estimated lifetime population prevalences in the general population (29% for anxiety and 24% for depression[4]). Although commonly associated with cancer treatment exposures, the variability of risk in psychiatric outcomes among survivors suggests a potential role for genetic susceptibility.

There have been several studies to identify genetic associations with late and long-term effects in childhood cancer survivors. Candidate gene studies and genome-wide association studies have successfully identified several genes and genetic variants associated with cardiovascular disease, infertility, obesity, neurocognitive deficits, and subsequent malignant neoplasms in long-term childhood cancer survivors[5]. However, genetic studies for psychiatric outcomes have been more limited. To date, there has been a single genome-wide association study focusing on post-traumatic stress disorder in the Childhood Cancer Survivor Study (CCSS) that successfully identified one significantly associated genomic locus and several marginally significant loci[6]. However, for anxiety and depression, there have been only a few candidate gene studies[7,8] and one exome-wide association study in childhood cancer survivors[9]. Conversely, in the general population, a substantial number of genetic variants have been associated with major depressive disorder (MDD)[10,11] and anxiety disorders[12–14]. Furthermore, current research focuses on the assessment of depression, anxiety, or distress at a single timepoint; whereas it is well established that psychiatric outcomes can be long lasting and often require independent treatment[15,16]. Previous studies in childhood cancer survivors have identified multiple outcome trajectories for global psychological distress, depression, and anxiety[17]. More specifically, four longitudinal patterns were identified: persistently low, or no symptoms; persistently elevated symptoms; increasing symptoms; and decreasing symptoms. Predictors of the different patterns also varied, which suggests that there could be variability in the underlying genetic risk for these outcomes long-term.

Therefore, the aim of this study is to better understand the genetic risk underlying psychological symptoms in childhood cancer survivors, both cross-sectionally and longitudinally. This first set of analyses will focus on both cross-sectional measures to identify the underlying genetic architecture and potential downstream consequences at a single timepoint, like those used in other largescale GWAS. Subsequently, we will examine the genetic associations with each of the longitudinal patterns of psychological distress that have been previously identified in childhood cancer survivors[17], including increasing symptoms over time, decreasing symptoms, and persistently elevated or low symptoms. Therefore, this study will examine the genetic liability for mental health outcomes in childhood cancer survivors in multiple disease contexts. Analyses will utilize both the baseline and

two follow-up assessments of the Brief Symptom Inventory-18[18]. We will use three measures assessed by the BSI-18: depression, anxiety, and the global severity index (GSI) of psychological distress, which integrates depression, anxiety, and somatization. While a unique measure, the GSI will allow us to identify associations with a broader internalizing psychopathology phenotype, which is a recent approach in the general population that focuses on alternative assessments of psychological outcomes, like HiTOP [19–21]. This study represents the first analysis to use GWAS to identify common genetic variants associated with global psychological distress (BSI-18 global severity index), depression, and anxiety in childhood cancer survivors. Additionally, we will also develop the first risk prediction model for these outcomes in childhood cancer survivors. The results of this study will shed light on the mechanisms underlying psychological outcomes, especially those are unique to childhood cancer survivors. Additionally, results will allow us to begin to identify survivors at even higher risk for psychological distress to hopefully be able to provide earlier intervention or inform prophylactic protocols.

Specific Aims

Aim 1: To identify common genetic variants associated with baseline measures of global psychological distress and specific symptom dimensions (i.e. anxiety, depression) in childhood cancer survivors, using CCSS as the discovery cohort and SJLIFE as the replication cohort.

1.1 Conduct GWAS using a continuous phenotype using BSI-18 t-scores for the highest level of statistical power to detect associations

1.2 Conduct GWAS using a clinically relevant phenotype for mental distress

Aim 2: Use GWAS to identify common variants that are associated with changes in psychological distress, depression, and anxiety.

2.1 Run a longitudinal latent profile analysis to identify different longitudinal patterns (i.e. increased distress/symptoms, decreased distress/symptoms, persistently low or no distress/symptoms, and persistently high distress/symptoms).and class membership for survivors to be included in subsequent analyses

2.2 Conduct GWAS using multinomial logistic regression for the different longitudinal patterns identified in aim 2.1.

Aim 3: Using GWAS summary statistics from Aims 1 and 2, identify functional consequences and downstream pathways of associated variants through functional annotation and estimate the proportion of variation explained by common variants and compare results from childhood cancer survivors to findings from the general population.

3.1 Conduct functional annotation analyses and gene-based analyses using SMR and FUMA/MAGMA to ascertain potential downstream pathways of associated genetic variants from Aims 1 and 2.

3.2 Using Linkage Disequilibrium Score Regression (LDSC), estimate the variance explained by all common variants and partitioned by those in different functional regions of the genome for each set of GWAS summary statistics from Aims 1 and 2, compare these results with those from the general population.

Aim 4: Develop risk prediction models for psychiatric symptoms in in CCSS and perform external validation using the SJLIFE cohort.

4.1 Develop and validate a risk prediction model using relevant clinical variables (e.g. age at diagnosis, diagnosis, treatment) for each psychiatric outcome

4.2 Develop and validate an integrated risk prediction model for each psychiatric outcome that incorporates both the clinical factors (from 4.1) and the polygenic scores from the general population.

Methods

Study Population

This study will include adult survivors from CCSS (original and expansion cohorts) and SJLIFE with available genotype data.

For the cross-sectional analyses in Aim 1, we will include CCSS survivors who have a baseline measure for the BSI-18 in addition to genotype data and SJLIFE survivors who have at least one measure for the BSI-18 at a baseline CORE assessment in addition to genotype data.

For the longitudinal analyses in Aim 2, we will include CCSS survivors who have a baseline measure for the BSI-18 and two follow-up assessments in addition to the genotype data. More specifically, for the original CCSS cohort this will consist of baseline, follow-up 2, and follow-up 4; whereas for the expansion cohort this will consist of baseline, follow-up 5, and follow-up 7. and SJLIFE survivors who have at least three BSI-18 measurements at separate evaluations at least 180 days apart in addition to genotype data.

In all aims, as we are using SJLIFE as a replication cohort and because SJLIFE has a more racially diverse population, overlapping survivors between CCSS and SJLIFE will be included in the initial discovery GWAS and removed from the replication cohort to ensure the highest statistical power to detect associations in the discovery analyses.

Outcome Variables

The main outcomes variables for this study are from the Brief Symptom Inventory-18. The first is the global severity index (GSI), which represent global psychological distress. We will also use two of the symptom dimension subscales, specifically depression and anxiety. For the first set of GWAS analyses in Aim 1, we will use continuous outcomes that consists of all the T-Scores for GSI, depression, and anxiety. For the second set of GWAS in Aim 1, we will use ordinal outcomes. The ordinal outcome will include three groups: those who are considered to have clinically significant GSI or symptoms, those who have substantial symptoms but do not reach the clinical threshold, and those who have very few or no symptoms. Clinical significance for GSI, depression, and anxiety using the currently established T-score threshold of 63 or higher[18]. Subthreshold individuals, or those who have substantial symptoms but do not reach the clinical threshold, have T-scores between 50[22,23] and 62. The unaffected survivors have T-scores below 50.

For the longitudinal analyses, the BSI-18 T-scores from the baselines assessment and two follow-ups will be considered (FU2 and FU4 for Original cohort survivors, FU5 and FU7 for Expansion cohort survivors). Specifically, we will use the BSI-18 measures at these three timepoints to construct a nominal variable, using longitudinal latent profile analysis, that includes the different longitudinal patterns.

Sociodemographic/Clinical Variables

- Demographic variables
    - Age (at baseline and contact for specific follow-up assessments)
    - Age at diagnosis
    - Sex
    - Race/ethnicity
- Clinical Variables
    - Diagnosis
        - Leukemia
        - CNS Tumor
        - Wilms Tumor
        - Lymphoma

        - Non-Hodgkin's Lymphoma
        - Neuroblastoma
        - Soft tissue sarcoma
        - Bone tumors
    - Age at diagnosis
    - Time since diagnosis
    - Cancer treatment exposures
        - Radiation (yes/no)
            - Field-specific radiotherapy with total body irradiation (yes/no) and dose
        - Cranial Radiation Therapy
            - 0-29 Gy
            - 30-49 Gy
            - 50Gy or greater
        - Chemotherapy (yes/no)
            - Alkylating agents
            - Anthracyclines
            - Methotrexate dose
            - Platinum dose
            - Epipodophyllotoxins dose
            - Corticosteroids
        - Surgery
- First 10 Genetic Principal Components (genetically determined ancestry)

Statistical Analysis

Aim 1: To identify common genetic variants associated with baseline measures of global psychological distress and specific symptom dimensions (i.e. anxiety, depression) using CCSS as the discovery.

Descriptive statistics of all patient characteristics and cancer-related variables will be reported, including frequency and proportion for the categorical variables, and mean, standard deviation, median, and IQR for the continuous variables.

We will first conduct a multi-ancestry genome-wide association study (GWASs), accounting for different ancestries using an ADMIXTURE coefficient. Subsequently, we will conduct ancestry-specific follow-up analyses where statistical power permits. In CCSS, the outcome is the BSI-18 scores at baseline; in SJLIFE, the outcome is the first available BSI-18 scores.

Analyses will be carried out using multivariable linear regression models with a continuous BSI-18 scores as the phenotype (i.e., not applying a cut-off on BSI-18 scores) in order have the highest statistical power possible. Subsequently, we will conduct GWAS using a proportional odds regression model with an ordinal outcome, which was previously described, for global psychological distress and symptom dimensions. Using an ordinal model will allow us to use more clinically relevant phenotypes. All regression models will adjust for age at BSI-18 measurement, sex, age at cancer diagnosis, cancer treatment exposures, and include an ADMIXTURE coefficient to adjust for the multi-ancestry sample. As a sensitivity analysis, we will replace the ADMIXTURE in the regression model with the top ancestry-specific genetic principal components. The selection weights due to oversampling of ALL survivors in the expansion cohorts will be accounted for in the regression model. The threshold to determine a significant association for any SNP will be the commonly accepted genome-wide significance threshold of 5×10^{-8}. A preliminary variable selection of cancer treatment exposures will be performed using stepwise selection as well as 10-fold cross-validated elastic net with the multivariable linear regression without the genetic variants. As a sensitivity analysis, the above analysis will be repeated with survivors who had antidepressants or anxiolytics use within the last two years removed from the study sample.

Genome-wide significant associations from the trans-ancestry analysis and ancestry-specific analyses will be evaluated in the SJLIFE cohort, focusing on the genetic variants identified in the CCSS cohort. Associations that are successfully replicated will then be further explored through stratified analyses by cancer treatment exposures to evaluate potential effect modification. In the corresponding regression models, the stratifying variables will be removed from the adjustment set.

Power Analysis

Power to detect genetic associations was assessed using the power calculations from Appendix A of Visscher et al. (2017)[24]. These calculations estimate power across a range of effect sizes and minor allele frequencies with a fixed sample size and alpha level ($5x10^{-08}$). As the sample sizes were incredible similar for all phenotypes for the multi-ancestry sample, the estimates for power are essentially the same for all three phenotypes. As can be seen in the table below, we have substantial power to detect associations across effect sizes and minor allele frequencies.

Statistical Power for Cross-Sectional Multi-Ancestry Analyses in Aim 1

MAF / Beta	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
0.05	0.0002	0.002	0.006	0.01	0.01
0.1	0.06	0.39	0.68	0.81	0.84
0.15	0.64	0.98	0.99	0.99	0.99
0.2	0.99	0.99	1.00	1.00	1.00
0.25	0.99	1.00	1.00	1.00	1.00
0.3	0.99	1.00	1.00	1.00	1.00

<u>Aim 2:</u> Conduct GWAS on the longitudinal patterns of the selected mental health outcomes to determine if there are specific common variants associated with different longitudinal patterns of global psychological distress and specific symptom dimensions (i.e. increased distress/symptoms, decreased distress/symptoms, persistently low or no distress/symptoms, and persistently high distress/symptoms) with CCSS as the discovery sample and SJLIFE as the replication sample.

For Aim 2, we will start by conducting a longitudinal latent profile analysis to identify longitudinal patterns psychological distress, similar to those conducted in Brinkman et al (2013). Original and Expansion cohort survivors with complete GSI, depression, and anxiety T-scores at Baseline and two follow-up questionnaires will be included in the analysis. The indicators will include GSI, depression, and anxiety T-scores at all three time points. Multiple statistical indicators will be used to select the best fitting model, including Bayesian Information Criterion, entropy, and p-values from the Vuong–Lo–Mendell–Rubin likelihood ratio test (VLMR). We will also select models to have minimum class membership >5% to provide sufficient power for subsequent analyses. Models will be fit with two through six profiles to determine the optimal number of profiles needed to describe longitudinal patterns of distress for survivors. The latent profile analysis will be performed using the statistical software MPLUS. The distress indicators will then be described for the resulting classes using summary statistics and spaghetti plots. Subsequently, we will use GWAS with multinomial logistic modelling to identify genetic variants that are significantly associated with distinct longitudinal patterns of distress. The model will include the specific pattern as the outcome and the same covariates will be used as in Aim 1 in addition to the time since the baseline BSI-18 measure, including demographic characteristics, clinical risk factors, and an ADMIXTURE coefficient or the top principal components. Targeted replication will be conducted in SJLIFE.

Power Analysis

The same methods were used to estimate power to detect genetic associations for the GWAS incorporating longitudinal data. These calculations estimate power across a range of effect sizes and minor allele frequencies with a fixed sample size (from Table 2) and alpha level ($5x10^{-08}$).

Statistical Power for Analyses in Aim 2

MAF / Beta	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
0.05	0.00007	0.0006	0.002	0. 003	0.003
0.1	0.02	0.16	0.36	0.50	0.55

0.15	0.33	0.89	0.98	0.99	0.99
0.2	0.89	0.99	0.99	0.99	0.99
0.25	0.99	1.00	1.00	1.00	1.00
0.3	0.99	1.00	1.00	1.00	1.00

<u>Aim 3:</u> Identify functional consequences and downstream pathways of associated variants using functional annotation and estimate the proportion of variation explained by common variants using GWAS summary statistics from Aims 1 and 2 and results from childhood cancer survivors to findings from the general population.

> ***3.1*** **Conduct functional annotation analyses and gene-based analyses using SMR and FUMA/MAGMA to ascertain potential downstream pathways of associated genetic variants from Aims 1 and 2.**
>
> ***3.2*** **Using Linkage Disequilibrium Score Regression (LDSC), estimate the variance explained by all common variants and partitioned by those in different functional regions of the genome for each set of GWAS summary statistics from Aims 1 and 2, compare these results with those from the general population.**

Aim 3 will use the summary statistics from Aims 1 and 2. We will independently conduct functional annotation analyses and gene-based analyses for the results from Aim 1 and Aim 2. Aim 3.1 will use currently available programs to conduct functional annotation and gene-based analyses. We will use both SMR[25] and FUMA[26]/MAGMA[27] to characterize the potential functional consequences of associated SNPS, annotate associated SNPs with genes, identify gene-trait associations, and tissue expression of the gene. Conducting these analyses independently for the results from Aims 1 and 2 will allow us to identify both the intersections and unique implicated loci in the different outcome contexts. Subsequently, in Aim 3.2 we will use LDSC with the 1000Genomes project reference genome and our GWAS summary statistics (from aims 1 and 2) to estimate the heritability for each phenotype and partition the heritability into functional regions of the genome.

<u>Aim 4:</u> Develop (in CCSS) risk prediction models for psychiatric outcomes in childhood cancer survivors and validate it in SJLIFE.

> ***4.1*** **Develop and validate a risk prediction model using relevant clinical variables (e.g. age at diagnosis, diagnosis, treatment) for each psychiatric outcome**
>
> ***4.2*** **Develop and validate an integrated risk prediction model for each psychiatric outcome that incorporates both the clinical factors (from 4.1) and the polygenic scores from the general population.**

For Aim 4, we will begin by constructing polygenic scores for each phenotype using the discovery GWAS in CCSS. For the first risk prediction model, we will use multivariable logistic regression model for impairment in each domain (depression, anxiety) adjusting for clinically relevant variables, including age at cancer diagnosis, sex, race and ethnicity, age at BSI-18 measurement, and cancer treatment, after variable selection using 10-fold cross-validated elastic net. The selection weights due to oversampling of ALL survivors in the expansion cohorts will be accounted for in the regression model. Subsequently, we will validate the model in SJLIFE.

We will also construct an integrated risk prediction model that incorporates both the clinical factors from the original prediction model and polygenic scores generated from GWAS in the general population for depression and anxiety. The integrated risk prediction models will be limited to depression and anxiety due to available GWAS summary statistics in the general population. We will also validate the integrated model in SJLIFE and compare the performance of both risk prediction models to ascertain which model is more effective for identifying survivors at a higher risk for developing depression or anxiety using time-dependent ROC analysis.

Power Analysis

To determine the power of our ROC analyses, power analyses were conducted using the pROC R package[28]. For all psychiatric outcomes, using the trans-ancestry sample, power was equal to 1.

Example Table and Figures

Table 1. Sample Sizes for Aim 1				
Childhood Cancer Survivor Study (CCSS, Discovery)				
	Cases (N %) (T-Score ≥ 63)	Subthreshold (N %) (62 = T-Score ≥ 50)	Control (N %) (T-Score < 50)	Continuous N
Depression				
All Ancestries	843 (10.2)	1705 (20.5)	5755 (69.3)	8,303
European	769 (10.1)	1571 (20.7)	5265 (69.2)	7,605
Anxiety				
All Ancestries	558 (6.7)	1620 (19.5)	6120 (73.8)	8,298
European	517 (6.8)	1493 (19.6)	5590 (73.6)	7,600
Global Severity Index (GSI)				
All Ancestries	689 (8.3)	2142 (25.8)	5464 (65.9)	8,295
European	625 (8.2)	1969 (25.9)	5004 (65.9)	7,598
St. Jude Lifetime Cohort (SJLIFE, Replication)				
	Cases (N %) (T-Score ≥ 63)	Subthreshold (N %) (62 = T-Score ≥ 50)	Control (N %) (T-Score < 50)	Continuous N
Depression				
All Ancestries	170	289	766	1,225
Anxiety				
All Ancestries	135	301	789	1,225
Global Severity Index (GSI)				
All Ancestries	179	366	680	1,225

Table 2. Sample Sizes for Aim 2				
CCSS (Discovery)				
	Baseline	First Follow-Up	Second Follow-Up	N for analyses
Depression	6,545	6,176	6,554	6,176
Anxiety	6,547	6,177	6,553	6,177
GSI	6,544	6,173	6,551	6,173
SJLIFE (Replication)				
Depression	1,225	1,224	1,225	1,225
Anxiety	1,224	1,223	1,224	1,224
GSI	1,224	1,224	1,224	1,224

Example Table 3. Demographic and Clinical Characteristics of Discovery and Replication Cohorts

	CCSS (Discovery)	SJLIFE (Replication/Validation)
Sex		
Males		
Females		
Age (years)		
Race and Ethnicity		
White, Non-Hispanic		
Black, Non-Hispanic		
Hispanic/Latinx		
Other		
Age at Diagnosis (years)		
Diagnosis		
Leukemia		
Lymphoma		
Non-Hodgkin Lymphoma		
Soft-tissue Sarcoma		
Bone Cancer		
Neuroblastoma		
CNS Tumor		
Wilms Tumor		
Time since Diagnosis (years)		
Cancer Treatment Exposures		
Radiation (yes/no)		
Radiation Therapy field and dose		
Cranial radiation therapy		
0-29 Gy		
30-49 Gy		
50Gy or greater		
Chemotherapy (yes/no)		
Alkylating agents		
Anthracyclines		
Methotrexate dose		
Platinum dose		
Epipodophyllotoxins dose		
Corticosteroid Use (yes/no)		

Aim 1 Results Tables and Figures

Example Table 4. Significantly Associated SNPs with GSI as a continuous outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Table 5. Significantly Associated SNPs with Depression as a continuous outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

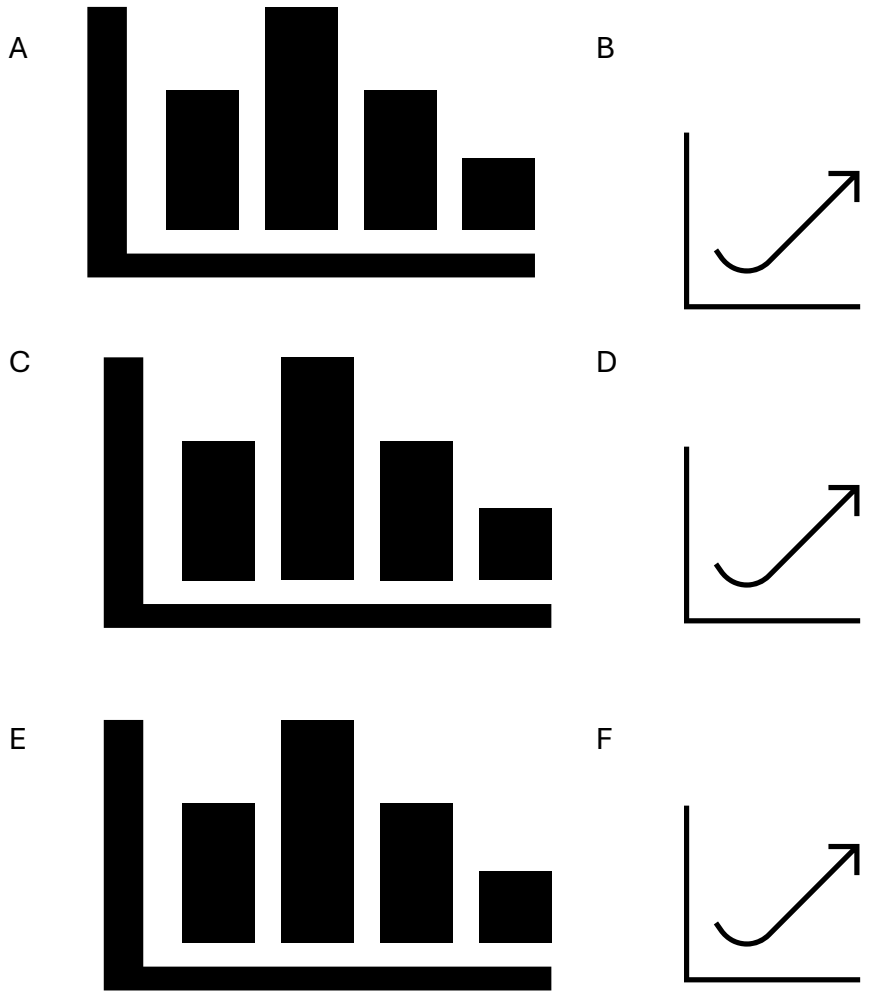
Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Table 6. Significantly Associated SNPs with Anxiety as a continuous outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Figure 1. Manhattan and Quantile-Quantile Plots for GSI, Depression, and Anxiety as Continuous Outcomes



Example Table 7. Significantly Associated SNPs with Anxiety as an ordinal outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Table 8. Significantly Associated SNPs with Anxiety as an ordinal outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Table 9. Significantly Associated SNPs with Anxiety as an ordinal outcome in Childhood Cancer Survivors

SNP	CHR	BP	A1	A2	MAF	Effect Size	P-value	Nearest Gene
rs12345	1	987654321	A	C	0.12	0.654	$7.89e^{-08}$	ACBC

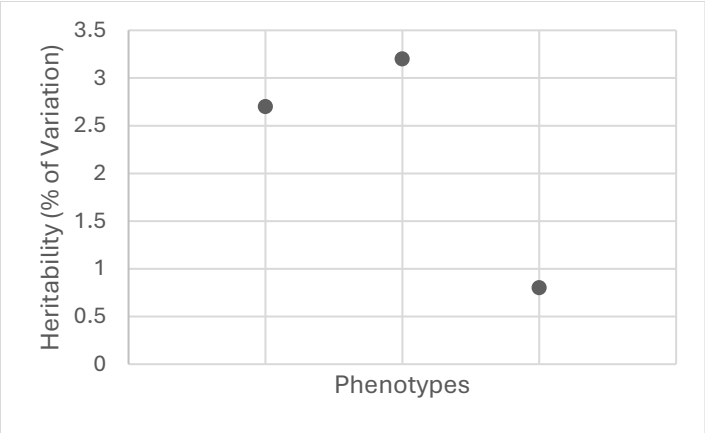
Abbreviations: SNP = single nucleotide polymorphism; CHR = chromosome; BP = Base Pair; A1 = Effect Allele; A2 = alternate allele; MAF = Minor Allele Frequency

Example Figure 2. Manhattan and Quantile-Quantile Plots for GSI, Depression, and Anxiety as Ordinal Outcomes

Aim 2 Tables and Figures will be very similar to those of Aim 1; except they will reflect the associations with the different patterns of symptoms instead of the psychological outcome

Aim 3 Example Tables and Figures

For gene-based associations with each outcome, the tables and figures will follow the same format as those for Aim 1 and Aim 2.



Example Figure 4. Heritability Estimates for GSI, Depression, and Anxiety as Continuous Outcomes

Example Table 10. Genetic Correlation Between Outcomes

	GSI	Depression	Anxiety
GSI	**1.0**		
Depression		**1.0**	
Anxiety			**1.0**

Example Table 11. Genetic Correlation Between Trajectories for Each Outcome

	Increasing	Decreasing	Persistent
Increasing	**1.0**		
Decreasing		**1.0**	
Persistent			**1.0**

References

1. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA. Cancer J. Clin.* **74**, 12–49 (2024).
2. Brinkman, T. M., Recklitis, C. J., Michel, G., Grootenhuis, M. A. & Klosky, J. L. Psychological symptoms, social outcomes, socioeconomic attainment, and health behaviors among survivors of childhood cancer: Current state of the literature. *J. Clin. Oncol.* **36**, 2190–2197 (2018).
3. Abadie, A. *et al.* Prevalence of Psychiatric Complications in Young Adults After Childhood Cancer Treatment: Results of the Long-Term Follow-Up Studies in Oncology. **9**, 247–255 (2020).
4. Osmani, V., Hörner, L., Klug, S. J. & Tanaka, L. F. Prevalence and risk of psychological distress, anxiety and depression in adolescent and young adult (AYA) cancer survivors: A systematic review and meta-analysis. *Cancer Med.* **12**, 18354–18367 (2023).
5. Gramatges, M. M. & Bhatia, S. Evidence for Genetic Risk Contributing to Long-Term Adverse Treatment Effects in Childhood Cancer Survivors. *Annu. Rev. Med.* **69**, 247–262 (2018).
6. Lu, D. *et al.* Genome-wide association study of posttraumatic stress disorder among childhood cancer survivors: results from the Childhood Cancer Survivor Study and the St. Jude Lifetime Cohort. *Transl. Psychiatry* **12**, 342 (2022).
7. Brackett, J. *et al.* Antioxidant enzyme polymorphisms and neuropsychological outcomes in medulloblastoma survivors: A report from the Childhood Cancer Survivor Study. **14**, 1018–1025 (2012).
8. Petrykey, K. *et al.* Influence of genetic factors on long-term treatment related neurocognitive complications, and on anxiety and depression in survivors of childhood acute lymphoblastic leukemia: The Petale study. *PLoS ONE* **14**, (2019).
9. Petrykey, K. *et al.* Genetic Factors and Long-term Treatment-Related Neurocognitive Deficits, Anxiety, and Depression in Childhood Leukemia Survivors: An Exome-Wide Association Study. *Cancer Epidemiol. Biomarkers Prev.* **33**, 234–243 (2024).
10. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).

11. Levey, D. F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat. Neurosci.* **24**, 954–963 (2021).
12. Strom, N. I. *et al.* Genome-wide association study of major anxiety disorders in 122,341 European-ancestry cases identifies 58 loci and highlights GABAergic signaling. Preprint at https://doi.org/10.1101/2024.07.03.24309466 (2024).
13. Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol. Psychiatry* **25**, 3292–3303 (2020).
14. Levey, D. F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am. J. Psychiatry* **177**, 223–232 (2020).
15. Sutin, A. R. *et al.* The Trajectory of Depressive Symptoms Across the Adult Lifespan. *JAMA Psychiatry Chic. Ill* **70**, 803–811 (2013).
16. Fountoulakis, K. N. *et al.* Treatment and long-term outcome of mental disorders: The grim picture from a quasi-epidemiological investigation in 54,826 subjects from 40 countries. *Psychiatry Res.* **348**, 116459 (2025).
17. Brinkman, T. M. *et al.* Longitudinal patterns of psychological distress in adult survivors of childhood cancer. *Br. J. Cancer* **109**, 1373–1381 (2013).
18. Derogatis, L. R. & Melisaratos, N. The Brief Symptom Inventory: an introductory report. *Psychol. Med.* **13**, 595–605 (1983).
19. Conway, C. C., Forbes, M. K. & South, S. C. A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers. *Clin. Psychol. Sci.* **10**, 236–258 (2021).
20. Davis, C. N. *et al.* Integrating HiTOP and RDoC frameworks Part I: Genetic architecture of externalizing and internalizing psychopathology. *Psychol. Med.* **55**, e138 (2025).
21. Kotov, R. *et al.* The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* **126**, 454–477 (2017).
22. Grassi, L., Caruso, R., Mitchell, A. J., Sabato, S. & Nanni, M. G. Screening for emotional disorders in patients with cancer using the Brief Symptom Inventory (BSI) and the BSI-18 versus a standardized psychiatric interview (the World Health Organization Composite International Diagnostic Interview). *Cancer* **124**, 2415–2426 (2018).

23. Recklitis, C. J., Blackmon, J. E. & Chang, G. Validity of the Brief Symptom Inventory-18 (BSI-18) for Identifying Depression and Anxiety in Young Adult Cancer Survivors: Comparison with a Structured Clinical Diagnostic Interview. *Psychol. Assess.* **29**, 1189–1200 (2017).
24. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
25. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
26. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
27. De Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
28. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).