SUPPLEMENTAL METHODS

Phenotype and covariate data in survivor cohorts

In the Childhood Cancer Survivor Study (CCSS), information related to the original pediatric cancer diagnosis and treatment was abstracted from medical records. CCSS has abstracted dose data for selected chemotherapies: doses of IV, IT methotrexate were available in CCSS and used in the current analyses. For other chemotherapies (corticosteroids), we considered exposure (yes/no) only. Using data from review of individual radiation therapy (RT) records, the maximum tumor dose (maxTD) from RT was estimated as the total delivered dose from all overlapping RT fields across seven major treatment regions, i.e., head, neck, chest, abdomen, pelvis, arm, and leg[1]. For each of these major treatment regions, radiation dose to adjacent regions to the primary treatment site were estimated to have received 2 Gy, while radiation dose to more distal body regions from the primary treatment site was estimated at 0.2 Gy. All other phenotype data were either self-reported in CCSS questionnaires or reported by family proxies for survivors who could not complete surveys, were deceased, or <18 years old. Fracture histories were only queried in the 2007 and 2014 follow-up questionnaires. Because the 2007 follow-up questionnaire only requested the age at first fracture, we extracted fracture histories for all discovery cohort participants from the 2014 follow-up questionnaire which queried lifetime fracture history and corresponding ages of occurrence and skeletal sites of fractures, allowing study of fractures occurring after primary cancer diagnosis. All reported fractures were assigned ICD9/10 medical diagnostic codes by a trained nosologist and reviewed for relevance. Participants with incomplete fracture event histories that precluded characterization of post-diagnosis incident fracture events were excluded (Supplemental Figure 1). Premature menopause status, defined as cessation of menses before age 40 years, was ascertained using CCSS baseline and follow-up questionnaires[2], while attained height and weight were taken from the 2014 CCSS follow-up questionnaire.

St. Jude Lifetime Cohort Study (SJLIFE) participants with biobanked specimens and at least one St. Jude Children's Research Hospital (SJCRH) on-site clinical assessment visit as of June 30, 2017 were included in the SJLIFE replication analysis. Data related to primary cancer diagnosis and treatments were obtained from medical record review at SJCRH, similar to CCSS. Measured height and weight were taken from the most recent SJCRH study visit and premature menopause status was clinically assessed and graded according to the NCI Common Terminology Criteria for Adverse Event (CTCAE) v4.03 classification system.

Genotype data in CCSS and SJLIFE

For CCSS, DNA was genotyped at the Cancer Genomics Research Laboratory of the National Cancer Institute (Bethesda, MD) using the Illumina HumanOmni5Exome array. Genotypes were called with Genotyping Module v1.9 (Illumina GenomeStudio software v2011.1). Samples with excess missingness (≥8%), heterozygosity (<0.11 or >0.16), sex discordance (X chromosome heterozygosity >5% for males or <20% for females), and cryptic relatedness (identity-by-descent sharing >0.70) were removed. For the 5,739 samples meeting these quality control thresholds, genotypes were imputed using Minimac3[3] and the Haplotype Reference Consortium r1.1 reference panel. A total of 2,453 participants of European genetic ancestry (see *Ancestry* below) who also met study inclusion criteria were retained (Supplemental Figure 1). Analyses excluded rare/low-frequency SNPs (minor allele frequency <5%), as well as SNPs with excess missingness (>5%) and departures from Hardy-Weinberg equilibrium ($P<1x10^{-6}$ among participants without fracture events). Analyses were further restricted to SNPs with high imputation quality ($r^2 \geq 0.8$), leaving ~5.4 million SNPs.

In SJLIFE, sequencing for 3,006 samples was completed at the HudsonAlpha Institute for Biotechnology Genomic Services Laboratory (Huntsville, AL) using the Illumina HiSeq X10 platform to yield 150 base pair paired-end reads with an average coverage per sample of 36.8X. Variant calls were processed with GATK v3.4.0[4] and BCFtools[5]. PLINK v1.90b[6] and VCFtools v0.1.13[7] were used to perform additional quality control, applying the following sample exclusion criteria: excess missingness (≥5%), cryptic relatedness (pi-hat>0.25), and excess heterozygosity (>3 SD). Variants with Hardy Weinberg equilibrium test $P<1\times10^{-10}$ and >10% missingness across samples were removed. Analyses were restricted to 1,417 SJLIFE participants of European ancestry (see *Ancestry* below) who met replication study inclusion criteria.

Ancestry

Procedures to identify the genetic ancestry of SJLIFE and CCSS samples have been described elsewhere[8,9]. Briefly, an EIGENSTRAT-based principal component analysis[10] was performed with PLINK v1.90b for each cohort by combining cohort samples with samples from 1000 Genomes (1000G) global reference populations. Samples with principal component scores ±3 SD of the means of the first two principal components in the 1000G European population were considered to be of European ancestry.

Relevant cancer treatment covariates

Before conducting SNP association testing with fracture risk, we evaluated univariate associations between fracture risk following childhood cancer diagnosis and cancer treatments with known osteotoxic effects in the CCSS discovery cohort. Time to first post-diagnosis fracture was analyzed using Cox proportional hazards regression models with age as the time scale, with follow-up time split into 1-year intervals in a counting-process data format beginning from childhood cancer diagnosis to first fracture or censoring at completion of the 2014 follow-up questionnaire. Univariate treatment associations (as hazard ratios or HRs) adjusted for sex, attained height and weight, and premature menopause status with P<0.2 (from two-sided testing) were considered to be relevant adjustment covariates. Treatment associations meeting this criterion included any exposure to corticosteroids, intravenous (IV) and intrathecal (IT) methotrexate dose, and maxTD from RT to any of seven body regions (head, neck, chest, abdomen, pelvis, arm, leg).

Testing for sex-heterogeneous SNP effects

GWAMA computes an asymptotically X^2-distributed test statistic using summary statistics to test for allelic effect differences between sexes. A Bonferroni-corrected p-value threshold considering SNPs with suggestively significant associations (two-sided $P<1\times10^{-5}$) in sex-specific discovery analyses (P=0.05/number of evaluated SNPs) was used to assess statistical significance.

Annotation of credible sets of SNPs

The Bayesian approach used to construct 99% credible intervals or sets of SNPs assumes the causal variant was genotyped and a single causal variant is responsible for the signal at that locus. Each locus was defined as the 1 Mb window centered at the most strongly associated SNP in discovery. With association summary statistics from the discovery analysis, we calculated an approximate Bayes' Factor[11]:

$$BF_j = \sqrt{1-R_j}\,\exp\left(\frac{R_j\beta_j^2}{2\sigma_j^2}\right)$$

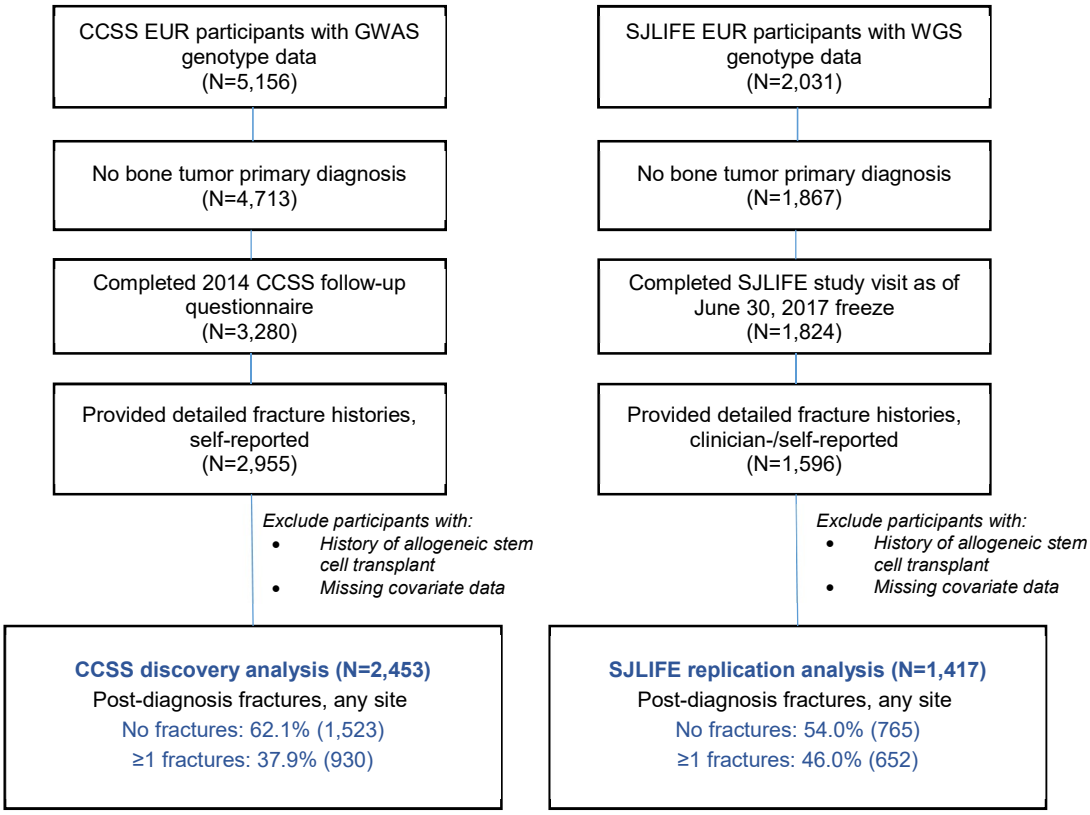
where β_j and σ_j are the allelic effect estimate (log[HR]) and standard error of the j^{th} SNP, and R_j=0.04/(σ_j^2+0.04), which incorporates a Gaussian prior $N(0,0.2^2)$ that gives higher probability to smaller effect sizes. The posterior probability that the j^{th} SNP is causal was calculated by $\pi_j = BF_j / \sum_{k=1}^{K} BF_k$. We then constructed 99% credible sets of SNPs for each locus by ranking all SNPs by their approximate Bayes' Factors and then included ranked SNPs until their cumulative posterior probability exceeded 0.99.

We examined credible set SNP associations in recent published GWAS of related phenotypes (estimated bone mineral density[12]; fracture[13]) and phenome-wide association study (PheWAS) results from the UK Biobank PheWeb (http://pheweb.sph.umich.edu:5000) for 2,419 UK Biobank phenotypes and the Michigan Genomics Initiative PheWeb (http://pheweb.sph.umich.edu/) for 1,448 ICD9 medical diagnostic codes. Coding and regulatory consequences of credible set SNPs, including Combined Annotation Dependent Depletion[14] (CADD) scores predicting variant deleteriousness (PHRED-scaled such that scores >10 represent variants with the top 10% of CADD scores, etc.), were annotated using the Ensembl Variant Effect Predictor[15] (VEP v99, genome build GRCh37). Credible set SNPs significantly associated with gene expression, i.e., expression quantitative trait loci (eQTLs; FDR≤5%), and DNA methylation levels, i.e., methylation quantitative trait loci (meQTLs; FDR<5%), were identified from the Genotype-Tissue Expression[16] (GTEx v8) project, NHLBI Genome-Wide Repository of Associations

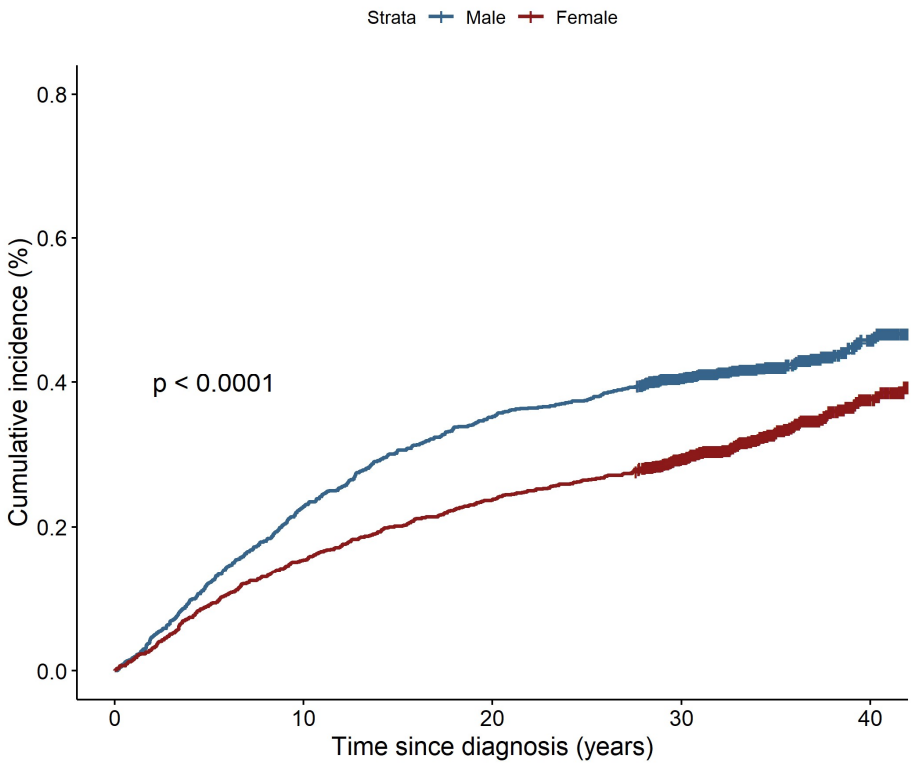
between SNPs and Phenotypes (GRASP v2.0.0.0)[17], and BIOS Consortium[18] (BIOS QTL) databases. Chromatin state annotations for regulatory states (e.g., promoters, enhancers) based on the 25-state ChromHMM model trained on 12 epigenetic marks for 127 epigenomes[19] were obtained from the Roadmap Epigenomics Consortium.

Chromatin state annotations from 25-state ChromHMM[19] were used to study enrichments for credible set SNP posterior probability in putative promoter states in a set of cell types specified *a priori* for relevance to fracture risk in survivors and were compared to a set of nine common human cell types from the Encyclopedia of DNA Elements[20] (ENCODE) Project reflecting a diversity of cell lines and tissue sources (GM12878 [B-lymphocyte], K562 [chronic myelogenous leukemia], HepG2 [hepatocellular carcinoma], HSMM [skeletal muscle myoblast], HUVEC [umbilical vein endothelial], NHEK [epidermal keratinocyte], NHLF [lung fibroblast], H1-hESC [embryonic stem cell], HMEC [mammary epithelial]). Promoter annotations based on ChromHMM were used because these annotations are based on prediction models learned on multiple directly measured experimental and imputed histone modification marks across multiple cell types rather than defining a promoter region based on one or two specific measured histone modification marks. We used pooled chromatin state annotations for active promoter (states 1-4 for active transcription start sites and upstream/downstream promoter flanks), poised promoter (states 22-23 for poised/bivalent promoters), and any promoter (active or poised) states. We then applied a permutation-based enrichment test procedure[21] to evaluate whether SNPs with higher probability of being “causal” (credible-set SNPs) are more likely to overlap promoter annotations in certain cell types (e.g., phenotype-relevant cells types vs. unrelated cell types) than expected. For each cell type, we first computed the mean posterior probability for the set of credible set SNPs overlapping promoter annotations and then generated a null distribution by calculating the mean poster probability of credible set SNPs overlapping randomly permuted promoter annotations (i.e., shifting promoter annotations’ genomic locations by a random distance selected from a uniform distribution of 1 to 100,000 bases in either direction of the observed promoter site) for 100,000 permutations. Relative fold enrichments were estimated by the ratio of the observed to expected posterior probability and test p-values were calculated by the proportion of permutations with an expected posterior probability that was equal or greater than the observed. A Bonferroni-corrected p-value threshold ($P<0.05$/number of cell types) was used to identify significant enrichments for credible set SNPs overlaps with cell-specific promoter annotations.

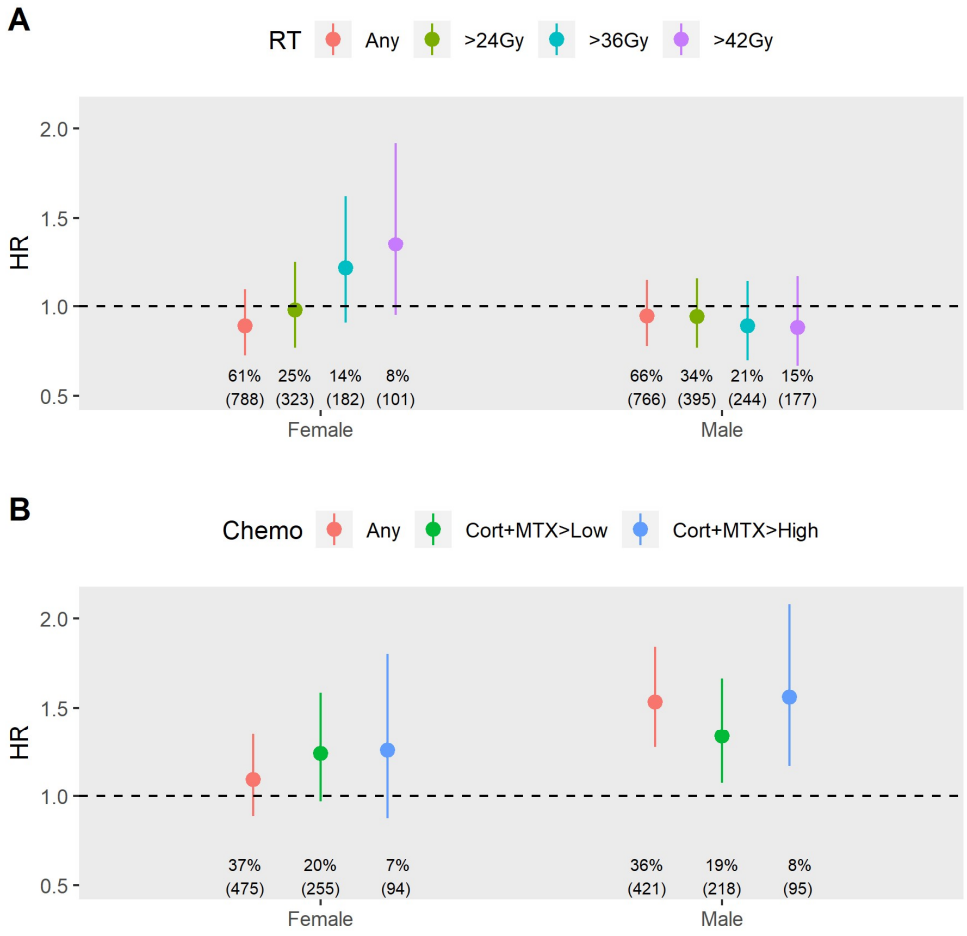
SUPPLEMENTAL FIGURES



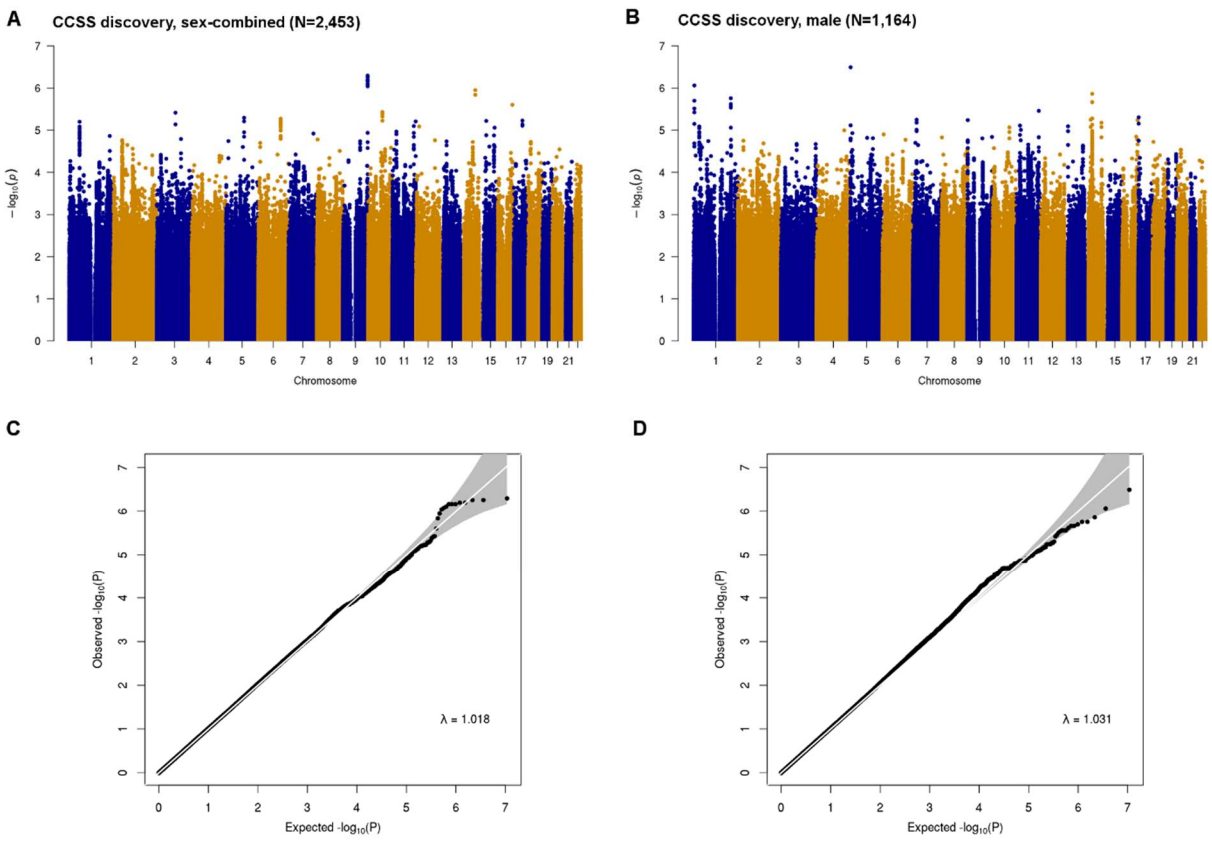
Supplemental Figure 1: Overview of major inclusion/exclusion criteria for final analytic samples.



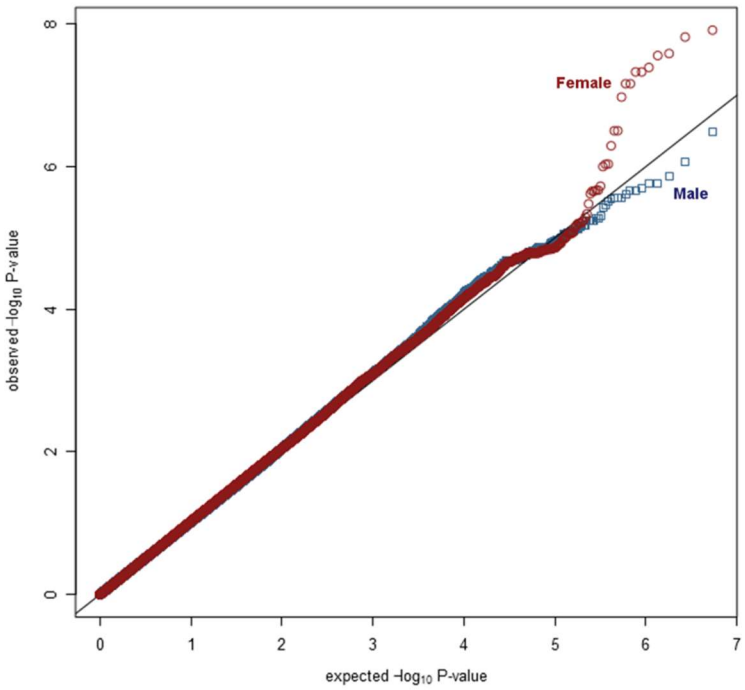
Supplemental Figure 2: Cumulative incidence curves for post-diagnosis fracture by sex in the CCSS discovery cohort (1,289 female survivors; 1,164 male survivors). The p-value from the log-rank test comparing the fracture risk probability distributions between sexes is provided in the lower left corner.



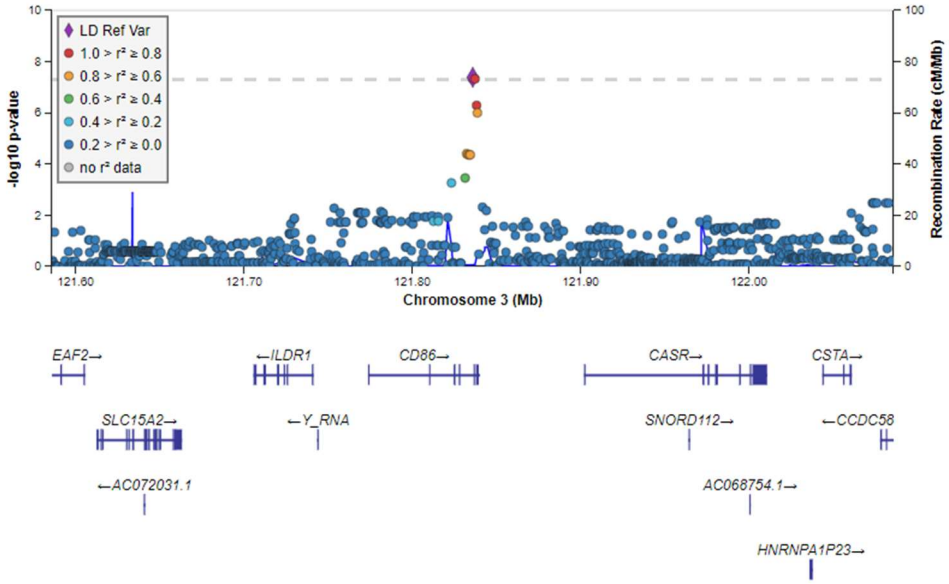
Supplemental Figure 3: Treatment threshold effects on risk of fracture following childhood cancer diagnosis by sex in the CCSS discovery cohort. Each panel shows HR estimates (dots) and respective 95% CIs (whiskers) grouped by sex, with colors corresponding to treatment threshold definitions. The top panel (A) compares adjusted fracture risk associations with threshold indicator variables for maximum cumulative radiation dosimetry dose across 7 body regions, with thresholds increasing from left to right (none versus any dose [orange], ≤24 Gy versus >24Gy [blue], ≤36 Gy versus >36Gy [teal], ≤42 Gy versus >42Gy [purple]). The bottom panel (B) compares adjusted fracture risk associations with threshold indicator variables for composite chemotherapy defined by corticosteroid exposure and IV/IT methotrexate dose, with IV/IT methotrexate dose thresholds increasing from left to right (none versus any dose [orange]; no corticosteroid exposure and ≤median IV/IT methotrexate versus corticosteroid exposure and >median IV/IT methotrexate dose [green], no corticosteroid exposure and ≤3rd quartile IV/IT methotrexate versus corticosteroid exposure and >3rd quartile IV/IT methotrexate dose [green]). Proportions and frequencies of participants meeting threshold definitions are provided at the bottom of each panel.



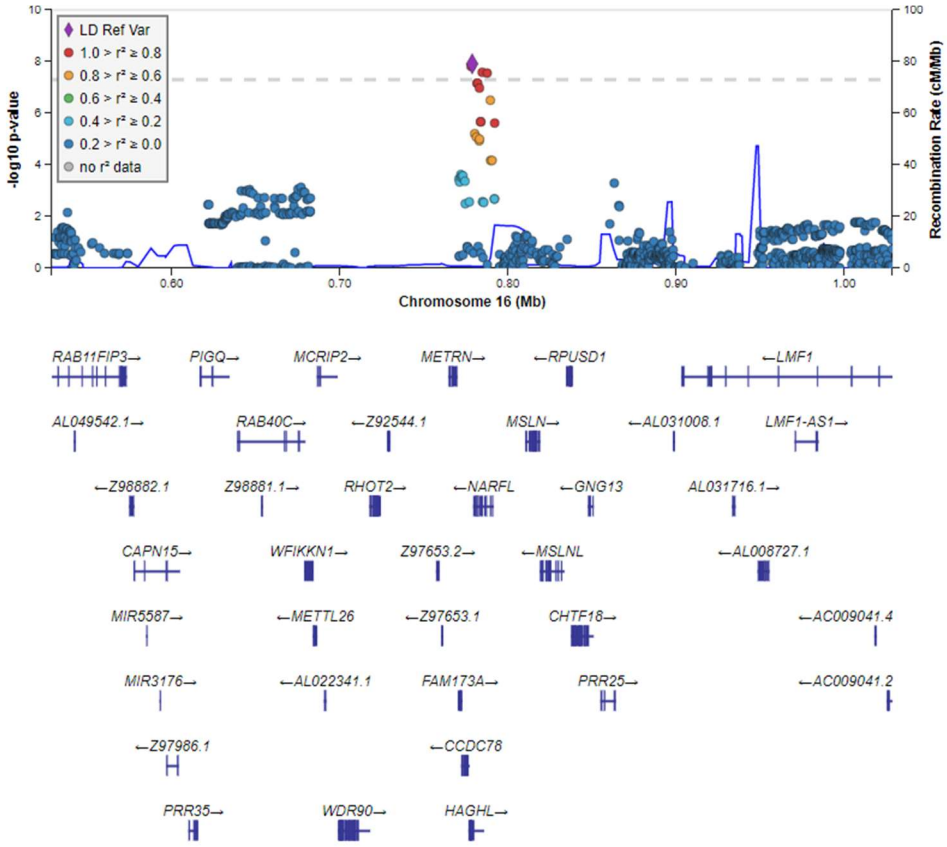
Supplemental Figure 4: Manhattan and quantile-quantile (QQ) plots for SNP association test p-values from post-diagnosis fracture risk GWAS in sex-combined and male survivors in CCSS. Manhattan plots illustrate $-\log_{10}$ p-values for SNP associations with post-diagnosis fracture risk (y-axis) against SNP genomic positions (x-axis), while QQ plots show observed $-\log_{10}$ p-values (y-axis) against those expected under the null distribution of no association (x-axis). Panels A and C show Manhattan and QQ plots, respectively, for the sex-combined discovery analysis, while panels B and D are Manhattan and QQ plots, respectively, for the male discovery analysis.



Supplemental Figure 5: QQ plots for SNP association test p-values from post-diagnosis fracture risk GWAS in sex-specific CCSS samples. The QQ plots of results from genome-wide association analyses performed in female survivors (red) and male survivors (blue) show observed -$\log_{10}$ p-values (y-axis) against those expected under the null distribution of no association (x-axis).



Supplemental Figure 6: LocusZoom plot of female-specific association p-values ($-\log_{10}P$) for the *CD86* locus. Results within a 500-kb window of the SNP with the strongest association with post-diagnosis fracture risk in this window (rs4315642, represented by the purple diamond) are shown. SNP color coding corresponds to the magnitude of LD with the top SNP (in r^2, 1000G EUR).



Supplemental Figure 7: LocusZoom plot of female-specific association p-values ($-\log_{10}P$) for the *HAGHL* locus. Results within a 500-kb window of the SNP with the strongest association with post-diagnosis fracture risk in this window (rs12448432, represented by the purple diamond) are shown. SNP color coding corresponds to the magnitude of LD with the top SNP (in r^2, 1000G EUR).

SUPPLEMENTAL TABLES

Supplemental Table 1: Demographic characteristics of childhood cancer survivors in CCSS discovery and SJLIFE replication cohorts, split by sex

	CCSS Discovery			SJLIFE Replication		
	Sex-combined (N=2,453)	Female (N=1,289)	Male (N=1,164)	Sex-combined (N=1,417)	Female (N=646)	Male (N=771)
Characteristic	% (N) or median (IQR)	% (N) or median (IQR)	% (N) or median (IQR)	% (N) or median (IQR)	% (N) or median (IQR)	% (N) or median (IQR)
Sex						
Female	52.5% (1289)			45.6% (646)		
Male	47.5% (1164)			54.4% (771)		
Age at follow-up[a] (years)	42 (36-48)	42 (36-48)	43 (37-48)	31 (26-39)	31 (26-39)	32 (26-38)
Height at follow-up[a] (cm)	168 (163-178)	163 (157-168)	178 (170-183)	169 (162-177)	162 (157-167)	176 (170-181)
Weight at follow-up[a] (kg)	77 (64-91)	68 (59-82)	84 (75-96)	79 (65-95)	70 (60-86)	86 (73-100)
Age at primary cancer diagnosis (years)	5 (2-12)	5 (2-12)	6 (3-12)	6 (3-12)	6 (3-13)	7 (3-12)
Primary cancer diagnosis						
Leukemia	35.6% (874)	38.9% (501)	32.0% (373)	35.1% (497)	35.6% (230)	34.6% (267)
Hodgkin disease	15.0% (367)	15.9% (205)	13.9% (162)	12.5% (177)	13.3% (86)	11.8% (91)
Kidney tumors	12.6% (309)	14.7% (190)	10.2% (119)	7.3% (104)	9.6% (62)	5.4% (42)
Soft tissue sarcoma	9.7% (237)	9.0% (116)	10.4% (121)	7.5% (106)	7.1% (46)	7.8% (60)
Central nervous system tumors	9.2% (226)	5.7% (74)	13.1% (152)	14.3% (203)	12.4% (80)	16.0% (123)
Neuroblastoma	9.1% (224)	10.9% (141)	7.1% (83)	4.7% (66)	4.8% (31)	4.5% (35)
Non-Hodgkin lymphoma	8.8% (216)	4.8% (62)	13.2% (154)	7.5% (106)	5.4% (35)	9.2% (71)
Other	--	--	--	11.2% (158)	11.8% (76)	10.6% (82)

a. In CCSS, follow-up time is defined by completion of the 2014 follow-up questionnaire, while in SJLIFE follow-up time is defined as the most recently completed follow-up survey.

Supplemental Table 2: Follow-up of first fractures after primary cancer diagnosis in CCSS and SJLIFE

	CCSS Discovery (N=2,453)		SJLIFE Replication (N=1,417)	
Characteristics	Female (N=1,289)	Male (N=1,164)	Female (N=646)	Male (N=771)
Total number of first fractures	429	501	246	406
Total follow-up in person-years	36,005	29,234	12,288	12,799
Median age at first fracture (IQR), in years	18 (11-31)	16 (11-25)	16 (10-25)	16 (11-22)

Supplemental Table 3: Distribution of first fracture sites assessed in CCSS and SJLIFE

	CCSS Discovery (N=2,453)			SJLIFE Replication (N=1,417)		
Fracture sites	Sex-Combined % (n)	Female % (n)	Male % (n)	Sex-Combined % (n)	Female % (n)	Male % (n)
Wrist, forearm	23.4% (218)	21.4% (92)	25.1% (126)	20.9% (136)	22.8% (56)	19.7% (80)
Rib	3.2% (30)	4.0% (17)	2.6% (13)	4.4% (29)	5.3% (13)	3.9% (16)
Shin	2.7% (25)	2.1% (9)	3.2% (16)	3.5% (23)	4.1% (10)	3.2% (13)
Hip/pelvis	1.4% (13)	2.3% (10)	0.6% (3)	1.4% (9)	1.2% (3)	1.5% (6)
Kneecap	1.1% (10)	1.4% (6)	0.8% (4)	0.9% (6)	0.8% (2)	1.0% (4)
Spine	0.9% (8)	0.9% (4)	0.8% (4)	6.7% (44)	6.9% (17)	6.7% (27)
Shoulder	0.6% (6)	0.2% (1)	1.0% (5)	0.8% (5)	0.4% (1)	1.0% (4)
Skull, face	3.5% (33)	2.6% (11)	4.4% (22)	4.3% (28)	2.4% (6)	5.4% (22)
Fingers, toes	20.3% (189)	22.1% (95)	18.8% (94)	12.4% (81)	13.0% (32)	12.1% (49)
Lower limb[a]	25.7% (239)	29.1% (125)	22.8% (114)	24.8% (162)	28.5% (70)	22.7% (92)
Upper limb[b]	37.3% (347)	33.8% (145)	40.3% (202)	35.7% (233)	35.0% (86)	36.2% (147)
Other	7.0% (65)	4.9% (21)	8.8% (44)	5.8% (38)	5.3% (13)	6.2% (25)
Total	930	429	501	652	246	406

a. Includes fractures of the upper leg, kneecap, lower leg, shin, foot, ankle.
b. Includes fractures of the upper arm, forearm, wrist, hand.

Supplemental Table 4: Distribution of cancer treatment exposures in CCSS and SJLIFE

	CCSS, Discovery			SJLIFE, Replication		
Treatments	**Sex-combined (N=2,453)** % (N) or median (IQR)	**Women (N=1,289)** % (N) or median (IQR)	**Men (N=1,164)** % (N) or median (IQR)	**Sex-combined (N=1,417)** % (N) or median (IQR)	**Women (N=646)** % (N) or median (IQR)	**Men (N=771)** % (N) or median (IQR)
Chemotherapy receipt (any)						
IV methotrexate	18.5% (454)	18.1% (233)	19.0% (221)	29.2% (414)	28.2% (182)	30.1% (232)
IT methotrexate	38.4% (941)	37.9% (488)	38.9% (453)	38.3% (543)	37.6% (243)	38.9% (300)
Glucocorticoids	47.2% (1,158)	47.0% (606)	47.4% (552)	48.3% (685)	46.9% (303)	49.5% (382)
Methotrexate dose, >0 mg/m^2 dose						
IV methotrexate	3,051 (805 - 6,058)	3,120 (596 - 6,550)	2,951 (923 - 5,510)	1,567 (211 - 2,952)	1,681 (370 – 2,922)	1,515 (185 - 3,153)
IT methotrexate	126 (71 - 222)	132 (72 - 223)	120 (68-222)	158 (93 - 233)	171 (84 - 235)	150 (96 - 233)
Radiation therapy receipt (any, >high scatter)						
Any site	63.0% (1,545)	61.0% (786)	65.2% (759)	48.2% (683)	48.9% (316)	47.6% (367)
Radiation to head regions (brain, neck, other head)	45.9% (1,125)	43.2% (557)	48.8% (568)	38.5% (545)	37.3% (241)	39.4% (304)
Radiation to trunk regions (chest, abdomen, pelvis)	37.0%% (908)	37.4% (482)	36.6% (426)	25.7% (364)	26.6% (172)	24.9% (192)
Radiation to limb regions (arm, leg)	1.4% (34)	1.7% (22)	1.0% (12)	3.7% (52)	4.3% (28)	3.1% (24)
Radiation therapy dosimetry maximum dose, >high scatter, cGy						
Any site	2,400 (2,000 - 3,900)	2,400 (1,800 - 3,600)	2,500 (2,000 - 4,100)	2,600 (2,100 - 4,500)	2,600 (2,100 - 4,000)	2,600 (2,100 - 5,070)
Head regions	2,400 (1,800 - 3,800)	2,400 (1,800 - 3,500)	2,400 (2,000 - 4,200)	2,600 (2,100 - 4,500)	2,600 (2,100 - 3,700)	2,600 (2,100 - 5,300)
Trunk regions	3,000 (2,000 - 3,900)	2,900 (2,000 - 4,000)	3,000 (2,000 - 3,800)	2,600 (2,100 - 3,500)	2,600 (2,175 - 3,600)	2,600 (2,100 - 3,500)
Limb regions	4,750 (3,625 - 5,900)	4,700 (3,000 - 5,850)	4,750 (4,275 - 5,725)	2,650 (2,000 - 4,600)	2,700 (2,000 - 4,522)	2,650 (2,000 - 3,500)

Supplemental Table 5: Multivariable models of adjusted cancer treatment associations in CCSS discovery samples

Covariate	HR (95% CI)	Z	P
Sex-combined model[a] (N=2,453)			
Corticosteroids (any vs. none)	1.13 (0.96-1.32)	1.49	0.14
IV methotrexate dose (100 g/m^2)	1.20 (1.00-1.45)	1.97	0.05
IT methotrexate dose (100 mg/m^2)	1.07 (0.99-1.15)	1.73	0.08
Radiation dosimetry dose (10 Gy)	0.99 (0.95-1.03)	-0.55	0.58
Female-specific model[b] (N=1,289)			
Corticosteroids (any vs. none)	1.08 (0.86-1.38)	0.67	0.50
IV methotrexate dose (100 g/m^2)	1.02 (0.76-1.37)	0.13	0.90
IT methotrexate dose (100 mg/m^2)	0.99 (0.88-1.12)	-0.14	0.89
Radiation dosimetry dose (10 Gy)	0.98 (0.92-1.05)	-0.56	0.58
Male-specific model[c] (N=1,164)			
Corticosteroids (any vs. none)	1.15 (0.93-1.42)	1.32	0.19
IV methotrexate dose (100 g/m^2)	1.46 (1.15-1.85)	3.12	1.8×10^{-3}
IT methotrexate dose (100 mg/m^2)	1.11 (1.02-1.22)	2.31	0.02
Radiation dosimetry dose (10 Gy)	0.99 (0.94-1.04)	-0.41	0.68

a. Adjusted for sex, height, weight, and premature menopause status.
b. Adjusted for height, weight, and premature menopause status.
c. Adjusted for height, weight.

Supplemental Table 6: Functional and regulatory annotations of 99% credible set SNPs at genome-wide significant loci

SNP	CHR	BP	EA	Genes (5 kb)	Functional consequences[a]	Predicted amino acid changes[a]	CADD[b]	Regulatory consequences[a]	Strongest eQTL[c]	GTExv8 top 3 eGenes[c]	Thyroid eGenes[c] (top 3)	Strongest meQTL[d]	Active promoter[e]	Poised promoter[e]	Active enhancer[e]	Weak enhancer[e]	State[e]: Osteoblast | Chondrocytes	State[e]: Fetal brain	State[e]: Ovary	Other relevant QTL[f]	Published GWS associations[g]
rs1406815	16	778158	G	CCDC78; HAGHL; NARFL	HAGHL (non coding transcript exon variant; synonymous variant; missense variant)	p.Arg50Gly	2.575	promoter (HAGHL, CCDC78)	NARFL (Tissue=Thyroid, EAdir=+, P=2.0e-43)	NARFL; HAGHL; WFIKKN1	NARFL; HAGHL; WFIKKN1	NA	69	48	0	10	1_TssA | 22_PromP	1_TssA	1_TssA	NA	NA
rs12448432	16	778820	A	CCDC78; HAGHL; NARFL	HAGHL (non coding transcript exon variant; synonymous variant; missense variant; NMD transcript variant)	p.[Ala202Thr; Ala94Thr; Ala84Thr; Ala21Thr]	2.036	promoter (HAGHL, CCDC78)	NARFL (Tissue=Thyroid, EAdir=+, P=1.1e-57)	NARFL; HAGHL; WFIKKN1	NARFL; HAGHL; WDR90	NA	25	45	11	39	22_PromP | 22_PromP	23_PromBiv	23_PromBiv	HAGHL eQTL [probe ILMN_15715], treated osteoblasts (dexamethasone, PGE2)	NA
rs3829492	16	781633	A	HAGHL; NARFL	NA (HAGHL, non-coding/intronic variant)	NA	16.11	NA	NARFL (Tissue=Thyroid, EAdir=+, P=1.3e-62)	NARFL; HAGHL; C16orf13	NARFL; HAGHL; WDR90	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	0	0	0	8_TxWk | 8_TxWk	7_Tx3'	7_Tx3'	NA	NA
rs12443759	16	782132	T	HAGHL; NARFL	NA (HAGHL, non-coding/intronic variant)	NA	4.136	NA	NARFL (Tissue=Thyroid, EAdir=+, P=3.6e-63)	NARFL; HAGHL; C16orf13	NARFL; HAGHL; WDR90	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	0	0	0	8_TxWk | 8_TxWk	7_Tx3'	7_Tx3'	NA	NA
rs61112891	16	783156	C	HAGHL; NARFL	HAGHL (non coding transcript exon variant); NARFL (non coding transcript exon variant)	NA	1.736	TF binding site (NARFL)	NARFL (Tissue=Thyroid, EAdir=+, P=2.2e-62)	NARFL; HAGHL; C16orf13	NARFL; HAGHL; WDR90	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	0	0	0	7_Tx3' | 7_Tx3'	7_Tx3'	7_Tx3'	NA	NA
rs12051048	16	783864	A	HAGHL; NARFL	HAGHL (non coding transcript exon variant)	NA	3.14	NA	NARFL (Tissue=Thyroid, EAdir=+, P=7.2e-79)	NARFL; WFIKKN1; HAGHL	NARFL; HAGHL; WFIKKN1	NA	0	0	0	0	7_Tx3' | 7_Tx3'	7_Tx3'	7_Tx3'	NA	NA
rs12051245	16	783865	C	HAGHL; NARFL	HAGHL (non coding transcript exon variant)	NA	2.563	NA	NARFL (Tissue=Thyroid, EAdir=+, P=7.2e-79)	NARFL; WFIKKN1; HAGHL	NARFL; HAGHL; WFIKKN1	NA	0	0	0	0	7_Tx3' | 7_Tx3'	7_Tx3'	7_Tx3'	NA	NA
rs9928077	16	784765	T	HAGHL; NARFL	HAGHL (non coding transcript exon variant)	NA	11.63	NA	NARFL (Tissue=Thyroid, EAdir=+, P=9.0e-58)	NARFL; HAGHL; WFIKKN1	NARFL; HAGHL; WDR90	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	0	0	0	7_Tx3' | 7_Tx3'	7_Tx3'	7_Tx3'	NA	NA
rs12597563	16	787738	C	NARFL	NARFL (5 prime UTR variant; missense variant; NMD transcript variant)	p.Pro46Ala	4.118	NA	NARFL (Tissue=Thyroid, EAdir=+, P=1.4e-47)	NARFL; HAGHL; WFIKKN1	NARFL; HAGHL; WFIKKN1	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	0	2	0	7_Tx3' | 7_Tx3'	7_Tx3'	7_Tx3'	NA	NA
rs10794640	16	789618	A	NARFL	NA	NA	0.045	promoter (NARFL)	NARFL (Tissue=Thyroid, EAdir=+, P=7.9e-48)	NARFL; HAGHL; C16orf13	NARFL; HAGHL; C16orf13	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	27	9	0	14	12_TxEnhW | 12_TxEnhW	10_TxEnh5'	17_EnhW2	NA	NA
rs11648796	16	792190	G	NARFL	NA	NA	2.211	promoter (NARFL)	NARFL (Tissue=Thyroid, EAdir=+, P=7.7e-55)	NARFL; WFIKKN1; HAGHL	NARFL; HAGHL; WFIKKN1	MSLN; NARFL; HAGHL (EAdir=+, Probe=cg27144592, P=3.3e-310)	0	2	0	0	25_Quies | 25_Quies	25_Quies	25_Quies	NA	Height (EAdir=+, P=1.0E-13)

Major abbreviations: EA, effect allele; NEA, non-effect allele; EAdir, EA association direction; QTL, quantitative trait loci; eQTL, expression QTL; meQTL methylation QTL; GWS, genome-wide significant; #, number.

a. Functional and overall regulatory consequences were annotated with Ensembl Variant Effect Predictor (VEP v99).
b. Combined Annotation Dependent Depletion (CADD) scores reflect variant deleteriousness, PHRED-scaled such that scores >10 represent variants with the top 10% of CADD scores, >20 with top 1% of CADD scores, etc.
c. eQTL variant annotations with FDR≤5% were based on GTEx v8, where eGenes are genes with at least one significant (FDR≤5%) cis-SNP association.
d. BIOS QTL was used to annotate significant (FDR<5%) meQTL variants.
e. Chromatin state annotations were taken from the 25-state (ChromHMM) model based on 12 epigenetic marks for 127 epigenomes (Roadmap Epigenomics Consortium). ChromHMM annotations include: 1_TssA (active TSS), 22_PromP (poised promoter), 23_PromBiv (bivalent promoter), 7_Tx3' (transcribed 3' preferential), 8_TxWk (weak transcription); 12_TxEnhW (transcribed and weak enhancer), 10_TxEnh5' (transcribed 5' preferential and enhancer); 12_TxEnhW (transcribed and weak enhancer), 17_EnhW1 (weak enhancer), 25_Quies (quiescent).
f. Other QTL annotations were taken from the NHLBI Genome-Wide Repository of Associations between SNPs and Phenotypes (GRASP v2.0.0.0).
g. NHGRI-EBI GWAS Catalog was used to annotate variants with published GWS phenotype associations.

Supplemental Table 7: Cancer treatment-stratified associations between replicated *HAGHL* SNP and post-diagnosis fracture risk in female survivors from CCSS and SJLIFE

					CCSS				SJLIFE			
Head/neck RT	SNP	Chr	BP	Strata	N	N_{cases}	HR (95% CI)	P	N	N_{cases}	HR (95% CI)	P
	rs1406815	16	778158	None	501	175	1.22 (0.95-1.57)	0.11	331	115	1.38 (1.03-1.85)	0.03
	rs1406815	16	778158	Any	788	254	1.88 (1.54-2.28)	$2.4x10^{-10}$	315	131	1.14 (0.83-1.57)	0.43
	rs1406815	16	778158	>24Gy	195	57	3.05 (1.95-4.76)	$9.1x10^{-7}$	145	54	1.48 (0.85-2.57)	0.17
	rs1406815	16	778158	>36Gy	117	39	3.79 (1.95-7.34)	$8.2x10^{-5}$	61	22	3.08 (1.09-8.74)	0.03
	rs12448432	16	778820	None	501	175	1.25 (0.97-1.61)	0.09	331	115	1.38 (1.03-1.85)	0.03
	rs12448432	16	778820	Any	788	254	1.86 (1.53-2.26)	$4.2x10^{-10}$	315	131	1.12 (0.81-1.54)	0.50
	rs12448432	16	778820	>24Gy	195	57	2.90 (1.87-4.52)	$2.3x10^{-6}$	145	54	1.48 (0.85-2.57)	0.17
	rs12448432	16	778820	>36Gy	117	39	3.51 (1.83-6.75)	$1.6x10^{-4}$	61	22	3.08 (1.09-8.74)	0.03
	rs9928077	16	784765	None	501	175	1.22 (0.95-1.57)	0.13	331	115	1.38 (1.03-1.85)	0.03
	rs9928077	16	784765	Any	788	254	1.86 (1.53-2.26)	$4.2x10^{-10}$	315	131	1.13 (0.82-1.56)	0.46
	rs9928077	16	784765	>24Gy	195	57	2.90 (1.87-4.52)	$2.3x10^{-6}$	145	54	1.48 (0.85-2.57)	0.17
	rs9928077	16	784765	>36Gy	117	39	3.51 (1.83-6.75)	$1.6x10^{-4}$	61	22	3.08 (1.09-8.74)	0.03
Trunk RT	rs1406815	16	778158	None	501	175	1.22 (0.95-1.57)	0.11	334	115	1.36 (1.01-1.84)	0.04
	rs1406815	16	778158	Any	788	254	1.88 (1.55-2.28)	$2.2x10^{-10}$	312	131	1.20 (0.88-1.64)	0.25
	rs1406815	16	778158	>24Gy	278	89	2.07 (1.49-2.86)	$1.2x10^{-5}$	117	53	0.97 (0.56-1.70)	0.92
	rs1406815	16	778158	>36Gy	144	51	2.47 (1.57-3.90)	$9.3x10^{-5}$	40	22	0.31 (0.08-1.14)	0.08
	rs12448432	16	778820	None	501	175	1.25 (0.97-1.61)	0.09	334	115	1.36 (1.01-1.84)	0.04
	rs12448432	16	778820	Any	788	254	1.86 (1.53-2.27)	$3.9x10^{-10}$	312	131	1.18 (0.86-1.62)	0.30
	rs12448432	16	778820	>24Gy	278	89	2.02 (1.46-2.79)	$2.3x10^{-5}$	117	53	0.96 (0.55-1.68)	0.88
	rs12448432	16	778820	>36Gy	144	51	2.40 (1.53-3.78)	$1.5x10^{-4}$	40	22	0.34 (0.09-1.24)	0.10
	rs9928077	16	784765	None	501	175	1.22 (0.95-1.57)	0.13	334	115	1.36 (1.01-1.83)	0.04
	rs9928077	16	784765	Any	788	254	1.86 (1.53-2.27)	$3.9x10^{-10}$	312	131	1.20 (0.87-1.64)	0.27
	rs9928077	16	784765	>24Gy	278	89	2.02 (1.46-2.79)	$2.3x10^{-5}$	117	53	0.96 (0.55-1.68)	0.88
	rs9928077	16	784765	>36Gy	144	51	2.40 (1.53-3.78)	$1.5x10^{-4}$	40	22	0.34 (0.09-1.24)	0.10
Chemotherapy	rs1406815	16	778158	None	644	212	1.63 (1.31-2.03)	$1.1x10^{-5}$	315	117	1.31 (0.96-1.77)	0.09
	rs1406815	16	778158	Any	475	164	1.33 (1.03-1.70)	0.03	255	101	1.25 (0.90-1.74)	0.19
	rs1406815	16	778158	>Med	323	106	1.10 (0.79-1.53)	0.56	170	69	1.10 (0.74-1.65)	0.63
	rs1406815	16	778158	>High	192	68	1.08 (0.71-1.64)	0.71	88	35	0.94 (0.48-1.83)	0.86
	rs12448432	16	778820	None	644	212	1.63 (1.31-2.03)	$1.2x10^{-5}$	315	117	1.28 (0.94-1.74)	0.11
	rs12448432	16	778820	Any	475	164	1.34 (1.05-1.73)	0.02	255	101	1.25 (0.89-1.74)	0.19
	rs12448432	16	778820	>Med	323	106	1.10 (0.79-1.53)	0.56	170	69	1.09 (0.73-1.64)	0.68
	rs12448432	16	778820	>High	192	68	1.08 (0.71-1.64)	0.71	88	35	0.94 (0.48-1.83)	0.86
	rs9928077	16	784765	None	644	212	1.63 (1.31-2.03)	$1.2x10^{-5}$	315	117	1.28 (0.94-1.74)	0.11
	rs9928077	16	784765	Any	475	164	1.31 (1.02-1.68)	0.04	255	101	1.24 (0.89-1.74)	0.21
	rs9928077	16	784765	>Med	323	106	1.08 (0.78-1.50)	0.66	170	69	1.12 (0.75-1.69)	0.58
	rs9928077	16	784765	>High	192	68	1.03 (0.68-1.56)	0.89	88	35	0.89 (0.45-1.74)	0.73

Abbreviations: RT, radiation therapy; Chr, chromosome; BP, genomic base position, GRCh37/hg19 reference; HR, hazard ratio; CI, confidence interval; Gy, Gray; Med, medium.

Strata thresholds for each treatment were defined as no exposure ("None"), any exposure ("Any"), >median dose exposure, >3rd quartile dose exposure. Head/neck RT includes RT to the head or neck; trunk RT includes RT to chest, abdomen, or pelvis; chemotherapy combines any exposure to corticosteroids and IT/IV methotrexate dose. All reported HRs (95% CI) are adjusted for the same covariates as the main analysis, with the addition or exclusion of specific treatment covariates as appropriate to the stratification (e.g., for head/neck RT stratification, models were not adjusted for any site RT dose, but were adjusted height, weight, premature menopause status, genetic ancestry, corticosteroids exposure, IT and IV methotrexate dose, and trunk RT dose).

Supplemental Table 8: Phenome-wide association study (PheWAS) results for credible-set SNPs

SNP	Chr	BP	CCSS HR	CCSS P	99% credible set posterior probability	GWAS phenotypes with P<threshold ($P<2.1x10^{-5}$, 2,419 phenotypes), listed in order by p-value, UK Biobank PheWeb[a] (N~337K)	ICD-9 category with top SNP association, MGI PheWeb[b] (1,448 codes, N up to ~24K)	ICD-9 codes with P<threshold, MGI PheWeb[b] ($P<3.5x10^{-5}$)	MGI PheWeb[b] musculoskeletal ICD-9 codes with $P<5x10^{-3}$
rs1406815	16	778158	1.55	$1.5x10^{-8}$	0.240	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy, unspecified back disorders
rs12448432	16	778820	1.55	$1.2x10^{-8}$	0.288	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy, unspecified back disorders
rs3829492	16	781633	1.54	$6.9x10^{-8}$	0.060	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy, senile osteoporosis
rs12443759	16	782132	1.54	$6.9x10^{-8}$	0.060	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy, senile osteoporosis
rs61112891	16	783156	1.54	$1.1x10^{-7}$	0.042	Height[c], mass[d], weight, hip circumference, forced vital capacity	Circulatory system	None	Arthropathy, senile osteoporosis, ganglion cyst
rs12051048	16	783864	1.44	$2.1x10^{-6}$	0.004	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Unspecified back disorders, arthropathy
rs12051245	16	783865	1.44	$2.1x10^{-6}$	0.004	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Unspecified back disorders, arthropathy
rs9928077	16	784765	1.54	$2.6x10^{-8}$	0.151	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy, unspecified back disorders
rs12597563	16	787738	1.57	$2.8x10^{-8}$	0.126	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Arthropathy
rs10794640	16	789618	1.54	$3.1x10^{-7}$	0.015	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Ganglion cyst
rs11648796	16	792190	1.44	$2.4x10^{-6}$	0.003	Height[c], mass[d], weight, hip circumference, forced vital capacity	Musculoskeletal	None	Unspecified back disorders, arthropathy

Abbreviations: Chr, chromosome; BP, base position (GRCh37); HR, hazard ratio; MGI, Michigan Genomics Initiative.

a. UK Biobank PheWeb refers to the PheWAS browser for UK Biobank GWAS conducted by the Neale lab (http://pheweb.sph.umich.edu:5000/).
b. MGI PheWeb refers to the PheWAS browser for ICD-9 billing codes derived from electronic health records conducted by the Michigan Genomics Initiative (http://pheweb.sph.umich.edu/).
c. Height corresponds to multiple height phenotypes, including standing height, sitting height, and comparative height at age 10 years.
d. Mass corresponds to multiple measured and predicted body mass phenotypes, including whole body, arm, and leg mass measures.

Supplemental Table 9: Credible-set SNP associations with bone mineral density and fracture risk in general population GWAS (UK Biobank)

				CCSS post-diagnosis fracture GWAS (N=1,289 female survivors)			SNP associations with estimated bone mineral density (eBMD), N=142,487 (PMID 30598549)				SNP associations with fracture risk, N=426,795 (PMID 28869591)			
SNP	CHR	BP	EA	HR	P	Posterior probability	EAF	Beta	SE	P	EAF	ln(OR)	SE (ln(OR))	P
rs1406815	16	778158	G	1.55	$1.5x10^{-8}$	0.240	0.21	-0.0070	0.0041	0.07	0.21	0.0010	0.0080	0.90
rs12448432	16	778820	A	1.55	$1.2x10^{-8}$	0.288	0.21	-0.0073	0.0041	0.06	0.21	0.0020	0.0081	0.81
rs3829492	16	781633	A	1.54	$6.9x10^{-8}$	0.060	0.19	-0.0104	0.0043	0.01	0.18	0.0074	0.0084	0.38
rs12443759	16	782132	T	1.54	$6.9x10^{-8}$	0.060	0.19	-0.0104	0.0043	0.01	0.18	0.0075	0.0084	0.37
rs61112891	16	783156	C	1.54	$1.1x10^{-7}$	0.042	0.19	-0.0104	0.0043	0.02	0.19	0.0077	0.0084	0.36
rs12051048	16	783864	A	1.44	$2.1x10^{-6}$	0.004	0.23	-0.0071	0.0040	0.07	0.23	0.0031	0.0078	0.69
rs12051245	16	783865	C	1.44	$2.1x10^{-6}$	0.004	0.23	-0.0070	0.0040	0.07	0.23	0.0031	0.0078	0.69
rs9928077	16	784765	T	1.54	$2.6x10^{-8}$	0.151	0.21	-0.0066	0.0041	0.09	0.21	0.0033	0.0080	0.68
rs12597563	16	787738	C	1.57	$2.8x10^{-8}$	0.126	0.19	-0.0029	0.0043	0.47	0.19	0.0046	0.0084	0.58
rs10794640	16	789618	A	1.54	$3.1x10^{-7}$	0.015	0.16	-0.0063	0.0046	0.16	0.16	0.0097	0.0089	0.27
rs11648796	16	792190	G	1.44	$2.4x10^{-6}$	0.003	0.23	-0.0053	0.0042	0.17	0.23	0.0042	0.0081	0.60
rs11648796	16	792190	G	0.68	$3.9x10^{-5}$	0.003	0.23	-0.0053	0.0042	0.17	0.23	0.0042	0.0081	0.60

Abbreviations: Chr, chromosome; BP, base position (GRCh37); HR, hazard ratio; OR, odds ratio.

References

1. Howell, R.M., Smith, S.A., Weathers, R.E., Kry, S.F. & Stovall, M. Adaptations to a Generalized Radiation Dose Reconstruction Methodology for Use in Epidemiologic Studies: An Update from the MD Anderson Late Effect Group. *Radiat. Res.* **192**, 169-188 (2019).
2. Mostoufi-Moab, S. *et al.* Endocrine abnormalities in aging survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *J. Clin. Oncol.* **34**, 3240 (2016).
3. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284-1287 (2016).
4. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303 (2010).
5. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
6. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575 (2007).
7. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
8. Sapkota, Y. *et al.* Genome-Wide Association Study in Irradiated Childhood Cancer Survivors Identifies HTR2A for Subsequent Basal Cell Carcinoma. *J. Invest. Dermatol.* https://doi.org/10.1016/j.jid.2019.02.029 (2019).
9. Sapkota, Y. *et al.* Whole-genome sequencing of childhood cancer survivors treated with cranial radiation therapy identifies 5p15. 33 locus for stroke: A report from the St. Jude Lifetime Cohort study. *Clin. Cancer Res.*, clincanres. 1231.2019 (2019).
10. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904 (2006).
11. Wakefield, J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am. J. Hum. Genet.* **81**, 208-227 (2007).
12. Kemp, J.P. *et al.* Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat. Genet.* **49**, 1468 (2017).
13. Morris, J.A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* **51**, 258-266 (2019).
14. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310 (2014).
15. McLaren, W. *et al.* The ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
16. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
17. Leslie, R., O'Donnell, C.J. & Johnson, A.D. GRASP: analysis of genotype–phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics* **30**, i185-i194 (2014).
18. Bonder, M.J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nat. Genet.* **49**, 131 (2017).
19. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
20. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
21. Gaulton, K.J. *et al.* Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat. Genet.* **47**, 1415-1425 (2015).