Study title: Genetic risk prediction profiles for fracture among childhood cancer survivors

Working groups: Genetics; Epidemiology/Biostatistics; Chronic Disease

Investigators:

Cindy Im, PhD, MPH	cim1@ualberta.ca
Gregory T. Armstrong, MD, MSCE	greg.armstrong@stjude.org
Smita Bhatia, MD, MPH	sbhatia@peds.uab.edu
Charles A. Sklar, MD	sklarc@mskcc.org
Kiri K. Ness, PT, PhD, FAPTA	kiri.ness@stjude.org
Carmen L. Wilson, PhD	carmen.wilson@stjude.org
Kevin Oeffinger, MD	oeffingk@mskcc.org
Wendy Leisenring, ScD	wleisenr@fredhutch.org
Lindsay M. Morton, PhD	mortonli@mail.nih.gov
Yadav Sapkota, PhD	yadav.sapkota@stjude.org
Melissa M. Hudson, MD	melissa.hudson@stjude.org
Leslie L. Robison, PhD	les.robison@stjude.org
Yutaka Yasui, PhD	yutaka.yasui@stjude.org

Background and rationale

Although advances in treatment regimens for childhood cancer have dramatically improved five-year survival rates[1], specific curative therapies can adversely affect bone metabolism[2-4]. Cranial radiation influences bone metabolism through injury to the hypothalamic-pituitary axis, affecting relevant sex and growth hormone secretions[4,5], while methotrexate and glucocorticoids have been observed to inhibit osteoblast (bone-forming) and increase osteoclast (bone-resorbing) cell activity[4,6-10]. Other determinants of bone quality such as nutrition (e.g., Vitamin D, calcium intake) and physical activity (e.g., load-bearing exercise) may also be negatively influenced by the pediatric cancer treatment experience[4,11]. In a recent investigation of chronic health conditions in ~1,700 adult survivors of pediatric cancers in the St. Jude Lifetime Cohort Study (SJLIFE), the prevalence of osteoporosis among at-risk survivors was estimated to be 9.6%[12]. This estimate in the relatively young SJLIFE cohort (median age=32 years) is more comparable to osteoporosis prevalence estimates in population-based cohorts comprised of older adults (up to 10.7%, ages 50-60 years)[13].

While there is an abundance of evidence suggesting childhood cancer survivors may be at increased risk for developing therapy-related bone morbidities, studies of bone mineral density (BMD, a clinical predictor for osteoporosis) and fractures (an endpoint measure for osteoporosis) have shown that survivors exhibit substantial variation in BMD levels and fracture risk despite common treatment exposures[2-4,7,10,11,14]. Genetic susceptibility, combined with clinical exposures, may contribute to this uncharacterized variation in BMD levels and fracture risk in childhood cancer survivors. Recent large-scale meta-analyses (>10,000 participants) have identified dozens of genome-wide significant associations between single nucleotide polymorphisms (SNPs) and BMD and/or fracture risk in healthy adults and children[15-19]. Genome-wide association analyses of BMD or fracture risk in the childhood cancer survivor population, however, have been limited to date. Our recent genome-wide analysis of SNPs involved in interactions between genomic regions with putative regulatory effects on gene transcription (i.e., enhancer-promoter SNP interactions) is among the first genetic analyses of BMD in childhood cancer survivors[20]. In this study, we demonstrated that regulatory networks of SNPs may modify the adverse effects of specific cytotoxic treatments on BMD in survivors of acute lymphoblastic leukemia in SJLIFE. For example, the functional annotation of one of our replicated regulatory SNP interactions that

appeared to modify the effects of methotrexate on BMD suggested the implicated SNPs may increase expression of *SP7*, a gene that encodes an osteogenic transcription factor, Osterix (Osx). Exposure to methotrexate has been linked to decreased Osx expression, reductions in osteocytes, and lower bone volume in rats[21]. As such, the absence of this regulatory SNP interaction may circumvent BMD recovery in survivors exposed to methotrexate.

As an extension of our study of genetic factors associated with BMD in SJLIFE survivors, we propose to investigate genetic variants associated with fracture risk and develop genetic risk profiles to identify childhood cancer survivors at risk for fractures, using high-density genotype and/or imputed genotype data from the Childhood Cancer Survivor Study (CCSS). Along with the potential benefit of developing clinical tools to improve surveillance of childhood cancer survivors with the greatest risk for fracture, the results of this study may elucidate novel mechanisms that describe how inherited genetic variation influences the effects of cancer treatments on bone metabolism and morbidities in survivors.

Specific hypotheses and aims

Hypotheses: We hypothesize that:

- Individual common or low-frequency genetic variants contribute to fracture risk among long-term survivors of childhood cancer;
- Individual common or low frequency variants and complex genetic variants modify the effects of cancer treatments that diminish bone mineral density and increase fracture risk; and
- Risk profiles that consider genetic factors along with clinical predictors may substantively improve prediction models for fracture risk.

Aim 1: Identify common/low-frequency genetic variants (minor allele frequency [MAF] ≥1%) associated with fracture risk among long-term survivors of childhood cancer of European ancestry in CCSS ("specified study cohort").

Aim 2: Identify individual common/low-frequency genetic variants (MAF ≥1%) and complex genetic variants (e.g., SNP interactions) that modify treatment effects on fracture risk in the specified study cohort.

Aim 3: Assess the extent to which specific sets of genetic predictors improve the discriminatory performance of non-genetic prediction models for fracture risk in childhood cancer survivors.

- Goal 3a: Develop and validate individual prediction models that consider clinical indicators (e.g., sociodemographic and lifestyle factors, cancer treatment history) to predict fracture risk among survivors.

- Goal 3b: Investigate whether the inclusion of a genetic risk profile based on common/low-frequency variants associated with fractures (Aim 1 results) substantively improves clinical prediction modeling for fracture risk.

- Goal 3c: Investigate whether the inclusion of a genetic risk profile based on common/low-frequency variants (Aim 1 results) and single/complex variants that modify treatment effects on fracture risk (Aim 2 results) substantively improves clinical prediction modeling for fracture risk.

Analysis framework

All proposed analyses will be conducted using existing CCSS data. The proposed outcomes of interest, study covariates, and methodological approach are described below.

Outcomes of interest:

Data for outcome of interest will be extracted from self-reported occurrence(s) of fracture, specifically from surveys collected at Follow-Up 4 (2007, question F11) and Follow-up 5 (2014, question G11). Participants are asked whether they had "ever broken a bone", and to provide details about all occurrences of fractures (age of occurrence; fracture site). While genetic analyses of fracture risk would ideally only consider low-trauma fractures confirmed by radiographic report, such data is not available in CCSS. Consequently, we plan to investigate fracture sites typically associated with low-trauma fracture (i.e., fractures of the spine or hip) separately, along with all fractures irrespective of site.

If feasible, our primary outcome of interest is the fracture rate, where we will consider the time to the first fracture event and fractures as recurrent events. If the data for a fracture rate analysis is inadequate (see bulleted criteria below), we will consider low versus high frequency fracture history (i.e., dichotomous outcome with low frequency defined by 0-1 fractures and high frequency defined by ≥2 fractures). The final outcome definition will determined based on the following factors:

- Analytic power and missingness;
- Strength of association evidence for clinical factors known to influence bone health;
- Overall fit of models for fracture outcomes with clinical factors known to influence bone health; and
- Strength of association evidence among SNPs reported to have robust associations with BMD and/or fracture risk in the general population.

Subject population:

The study population will include up to ~5,300 long-term survivors of childhood cancer with genotype data and of European ancestry enrolled in the Original CCSS Cohort. European ancestry will be determined via principal components analysis, using the 1000 Genomes Phase III EUR cohort as a reference. Among the entire study sample with genotype and phenotype data (N=5,264), 46.3% (N=2,438) reported occurrence of at least one fracture. For this analysis, we will only include study participants with both genotype and phenotype data, and who also meet inclusion criteria typically applied in GWAS (i.e., meets missingness, heterozygosity, sex discordance, and relatedness thresholds). Participants with a bone tumor primary diagnosis will be excluded from this analysis, since these participants could have pathological fractures.

Explanatory variables:

Primary:

The primary explanatory variables of interest are the ~3.1 million genetic variants genotyped using the high-density Illumina HumanOmni5Exome array. For some analyses, we will also consider genotypes imputed with Minimac3 software, using the Haplotype Reference Consortium (release 1.1) as a reference.

Additional covariates:

- Sex
- Ancestry (calculated via principal components analysis)
- Height
- Weight
- Smoking history
- Self-reported physical activity
- Self-reported diagnosis of a problem affecting balance or equilibrium
- Cancer diagnosis
- Age at diagnosis
- Age at last follow-up
- Mortality data
- Use of medications/agents known to promote bone health (e.g., vitamin D, hormone replacement therapy, bisphosphonates, calcium supplements)
- Methotrexate (any exposure)
- Cumulative chemotherapy dose (alkylating agent score)
- Glucocorticoids (any exposure to dexamethasone, prednisone)
- Radiation to the central nervous system (dose)
- Pelvic radiation (dose)
- Radiation to fracture site (dose)

Analytic approach:

*Analyses of common and low-frequency variants***:**

To conduct single-marker and treatment interaction association tests for common and low-frequency variants (MAF ≥1%), Aims 1 and 2 will employ Cox proportional hazards regression to study associations with first fracture risk, and related methods to study recurrent fracture risk. Association analyses will assume an additive genetic inheritance model, adjusting for typical covariates applied in BMD or fracture risk GWAS (e.g., age, sex, ancestry, height, weight), as well as other relevant clinical covariates (e.g., cancer treatments). As described previously, the best outcome definition will be selected based on the strength of association evidence between clinical factors known to impact bone health and the results of a "look-up" analysis for single SNPs reported to have robust associations with BMD and/or fracture risk in prior GWAS. A genome-wide significance threshold of $p<5x10^{-8}$ will be used to control for Type I error inflation while considering ~1 million independent common variants. We will perform replication studies and meta-analysis with the SJLIFE cohort with whole-genome sequencing (WGS) data, specifically among participants of European ancestry that are not simultaneously enrolled in CCSS (N≈2,100). Approaches to investigate complex genetic variants associated with fracture risk will be developed, similar to our regulatory SNP interaction analysis of BMD in SJLIFE[20].

Biological annotation and functional validation:

We will contextualize candidate genetic susceptibility factors by annotating implicated variants with data from multiple public bioinformatics databases (ENCODE, GTEx, KEGG, etc.) and investigate the biological plausibility of reported associations with ancillary bioinformatics analyses. If feasible, functional validation of candidate genetic susceptibility factors from Aims 1 and 2 will be undertaken with SJCRH lab scientists through the prediction modeling R01 of Drs. Yasui/Zhang.

Clinical and genetic risk profile prediction modeling:

We plan to develop a "clinical prediction model" consisting of pre-selected clinical, sociodemographic, and lifestyle indicators using the specified CCSS cohort as the training dataset (N≈5,300). The existing SJLIFE cohort with WGS data will be treated as the independent test dataset (N≈2,100). Selection of the most influential set of clinical predictors for inclusion in the final clinical prediction model will employ a variable selection procedure (e.g., penalized regression such as LASSO, Elastic Net) and area under the curve (AUC) and concordance (C) statistics derived from 10-fold cross validation in the training set to reduce the possibility of over-fitting. Similar to previous prediction modeling analyses in CCSS[22], individual clinical risk scores will be computed from the linear combination of estimated regression coefficients and observed predictors, which will inform the creation of distinct risk groupings (e.g., low, moderate, high risk). AUC and C-statistics will be used to evaluate the discriminatory and predictive power of the final clinical prediction model in the test SJLIFE cohort. If possible, secondary validation of the clinical prediction model will be conducted in the SJLIFE cohort with forthcoming WGS data (N≈1,000).

The development of genetic risk profiles is to be based on pre-defined sets of genetic variants identified in Aims 1 and 2. The selection of genetic variants for inclusion in genetic risk profiles will be similar to the process described above for clinical prediction model creation. AUC and C-statistics will be used to evaluate the discriminatory and predictive power of the final clinical and genetic risk profile prediction models in the test dataset comprised of SJLIFE participants. If possible, external validation of clinical and genetic risk profile prediction models will be conducted in the SJLIFE cohort with forthcoming WGS data.

Proposed Figures and Tables

Table 1: Demographic and treatment exposure characteristics of childhood cancer survivors in CCSS and SJLIFE

	Discovery		**Replication / Validation**	
	CCSS (N=)		**SJLIFE (N=)**	
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 therapies				
Alkylating score				
0				
1				
2				
3				
Methotrexate (any)				
Glucocorticoid				
Dexamethasone (any)				
Prednisone (any)				
Cranial irradiation (any)				
Pelvic irradiation (any)				

Table 2: Age at first fracture, total number of reported fractures, and reported history of low-trauma fractures among CCSS and SJLIFE participants

	Discovery		Replication / Validation	
	CCSS (N=)		SJLIFE (N=)	
	No.	%	No.	%
Age of first fracture, years				
0-9				
10-17				
18-35				
36-50				
>50				
Total number of reported fractures				
0				
1				
2				
≥3				
Report of low-trauma fracture(s)				
No low-trauma fracture				
At least one low-trauma fracture				
1				
≥2				

Table 3 (Aim 1): Top single genetic variant association results ($P<5x10^{-5}$) for fracture outcome in discovery and replication cohorts

Genetic variant						CCSS (Discovery, N=)				SJLIFE (Replication, N=)				Joint (N=)			
Chr.	Pos.	Nearest gene	RSID	Ref.	Alt.	Freq.	RR / Beta	SE	P	Freq.	RR / Beta	SE	P	Freq.	RR / Beta	SE	P

Abbreviations: Chromosome (Chr.), genomic position (Pos.), SNP identifier (RSID, if available), reference allele (Ref.), alternative allele (Alt.), alternative allele frequency in sample (Freq.), standard error (SE).

Table 4 (Aim 2): Top genetic variant-treatment interactions ($P<5x10^{-5}$) associated with fracture outcome, stratified by cancer treatment exposure

Genetic variant						CCSS (Discovery, N=)				SJLIFE (Replication, N=)				Joint (N=)			
Chr.	Pos.	Nearest gene	RSID	Ref.	Alt.	Interaction P	TX type	TX RR/Beta (P)	No TX RR/Beta (P)	Interaction P	TX type	TX RR/Beta (P)	No TX RR/Beta (P)	Interaction P	TX type	TX RR/Beta (P)	No TX RR/Beta (P)

Abbreviations: Chromosome (Chr.), genomic position (Pos.), SNP identifier (RSID, if available), reference allele (Ref.), alternative allele (Alt.), treatment (TX)

Table 6 (Aim 3): Fracture risk groups by clinical and genetic risk profile models, with corresponding model discrimination and predictive power

	CCSS (Training, N=)					SJLIFE (Validation, N=)				
Models	**# Events**	**RR (vs. low clinical risk)**	**95% CI**	**AUC**	**C-statistic**	**# Events**	**RR (vs. low clinical risk)**	**95% CI**	**AUC**	**C-statistic**
Clinical model										
Low risk		1.0	Referent				1.0	Referent		
Moderate risk										
High risk										
Simple genetic risk profile[a]										
Low risk										
Moderate risk										
High risk										
Treatment-genetic risk profile[b]										
Low risk										
Moderate risk										
High risk										
Complex genetic risk profile[c]										
Low risk										
Moderate risk										
High risk										

Abbreviations: Area under the curve (AUC), concordance (C).

a. Profile only includes common/low-frequency single genetic variants (accounting for relevant clinical covariates).
b. Profile also includes single genetic variant-treatment interactions (accounting for relevant clinical covariates).
c. Profile also includes complex genetic variants (accounting for relevant clinical covariates).

References:

1. Howlader, N., Noone, A., Krapcho, M., et al. (2015). SEER Cancer Statistics Review, 1975–2012, National Cancer Institute. Bethesda. Journal.
2. Gurney, J.G., Kaste, S.C., Liu, W., et al. (2014). Bone mineral density among long-term survivors of childhood acute lymphoblastic leukemia: results from the St. Jude Lifetime Cohort Study. Pediatr Blood Cancer *61*, 1270-1276.
3. Kaste, S.C., Rai, S.N., Fleming, K., et al. (2006). Changes in bone mineral density in survivors of childhood acute lymphoblastic leukemia. Pediatr Blood Cancer *46*, 77-87.
4. Wasilewski-Masker, K., Kaste, S.C., Hudson, M.M., et al. (2008). Bone mineral density deficits in survivors of childhood cancer: long-term follow-up guidelines and review of the literature. Pediatrics *121*, e705-713.
5. Chemaitilly, W., and Sklar, C.A. (2010). Endocrine complications in long-term survivors of childhood cancers. Endocr Relat Cancer *17*, R141-R159.
6. Weinstein, R.S. (2011). Glucocorticoid-induced bone disease. N. Engl. J. Med. *365*, 62-70.
7. Davies, J.H., Evans, B.A., Jenney, M.E., et al. (2005). Skeletal morbidity in childhood acute lymphoblastic leukaemia. Clin Endocrinol (Oxf) *63*, 1-9.
8. Steinbuch, M., Youket, T.E., and Cohen, S. (2004). Oral glucocorticoid use is associated with an increased risk of fracture. Osteoporos Int *15*, 323-328.
9. Davies, J.H., Evans, B.A., Jenney, M.E., et al. (2003). Effects of chemotherapeutic agents on the function of primary human osteoblast-like cells derived from children. J Clin Endocrinol Metab *88*, 6088-6097.
10. van der Sluis, I.M., van den Heuvel-Eibrink, M.M., Hählen, K., et al. (2002). Altered bone mineral density and body composition, and increased fracture risk in childhood acute lymphoblastic leukemia. The Journal of pediatrics *141*, 204-210.
11. Wilson, C.L., and Ness, K.K. (2013). Bone mineral density deficits and fractures in survivors of childhood cancer. Current osteoporosis reports *11*, 329-337.
12. Hudson, M.M., Ness, K.K., Gurney, J.G., et al. (2013). Clinical ascertainment of health outcomes among adults treated for childhood cancer. JAMA *309*, 2371-2381.
13. Kanis, J.A. (2008). Assessment of osteoporosis at the primary health-care level. Technical report. World Health Organization Collaborating Center for Metabolic Bone Diseases. University of Sheffield, UK.
14. Wilson, C.L., Dilley, K., Ness, K.K., et al. (2012). Fractures among long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. Cancer *118*, 5920-5928.
15. Estrada, K., Styrkarsdottir, U., Evangelou, E., et al. (2012). Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. Nat. Genet. *44*, 491-501.
16. Medina-Gomez, C., Kemp, J.P., Estrada, K., et al. (2012). Meta-analysis of genome-wide scans for total body BMD in children and adults reveals allelic heterogeneity and age-specific effects at the WNT16 locus. PLoS Genet *8*, e1002718.
17. Rivadeneira, F., Styrkársdottir, U., Estrada, K., et al. (2009). Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. Nat. Genet. *41*, 1199-1206.
18. Timpson, N.J., Tobias, J.H., Richards, J.B., et al. (2009). Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. Hum. Mol. Genet. *18*, 1510-1517.
19. Styrkarsdottir, U., Halldorsson, B.V., Gretarsdottir, S., et al. (2008). Multiple genetic loci for bone mineral density and fractures. N. Engl. J. Med. *358*, 2355-2365.

20. Im, C., Ness, K.K., Kaste, S.C., et al. (2018). Genome-wide search for higher order epistasis as modifiers of treatment effects on bone mineral density in childhood cancer survivors. Eur. J. Hum. Genet. *26*, 275-286.
21. Georgiou, K.R., Scherer, M.A., Fan, C.M., et al. (2012). Methotrexate chemotherapy reduces osteogenesis but increases adipogenic potential in the bone marrow. J Cell Physiol *227*, 909-918.
22. Chow, E.J., Chen, Y., Kremer, L.C., et al. (2015). Individual prediction of heart failure among childhood cancer survivors. J Clin Oncol *33*, 394.