Childhood Cancer Survivor Study
Analysis Concept Proposal

August 9, 2011

1. Title: Development and validation of an absolute risk prediction model for thyroid cancer in childhood cancer survivors

2. Investigators: This proposed study will be within the Second Malignancies Working Group. Proposed investigators include:

Statistical leads:

Stephanie Kovalchik	stephanie.kovalchik@nih.gov	NCI/Biostatistics
Ruth Pfeiffer	pfeiffer@mail.nih.gov	NCI/Biostatistics

Co- and supporting investigators:

Alice Sigurdson	sigurdsa@mail.nih.gov	NCI/Radiation
Cécile Ronckers	c.m.ronckers@uva.amc.nl	The Netherlands
Lene Veiga	veigal@mail.nih.gov	NCI/CNEN, Brazil
Parveen Bhatti	pbhatti@fcrdc.org	FHCRC, Seattle
Marilyn Stovall	mstovall@mdanderson.org	UTMDACC, Houston
Susan Smith	sasmith@mdanderson.org	UTMDACC, Houston
Rita Weathers	rweather@mdanderson.org	UTMDACC, Houston
Wendy Leisenring	wleisenr@fhcrc.org	FHCRC, Seattle
Chuck Sklar	sklarc@mskcc.org	MSKCC, New York
Sarah Donaldson	sarah2@.stanford.edu	Stanford Univ., CA
Sue Hammond	sue.hammond@nationwidechildrens.org	Ohio State U, Columbus
Peggy Tucker	tucker@mail.nih.gov	NCI/Genetics
Ann Mertens	ann.mertens@choa.org	Emory, Atlanta
Joe Neglia	jneglia@umn.edu	Univ. Minnesota
Anna Meadows	meadows@email.chop.edu	Philadelphia
Peter Inskip	inskippe@mail.nih.gov	NCI/Radiation
Leslie Robison	Les.Robison@st.jude.org	St Jude's, Memphis

3. Background and Rationale

Cancer of the thyroid gland is relatively rare, although the incidence has significantly increased in the past two decades, in particular for papillary tumors (Enewold et al, 2009). In contrast to most other solid cancers, thyroid cancer is more common in females than males and the average age at diagnosis is in the forties rather than later in life. A well-established thyroid cancer risk factor is exposure to ionizing radiation in the head and neck region at a young age. Others are iodine deficiency and a history of thyroid diseases, particularly nodular hyperplasia and thyroid adenomas (reviewed in Ron 1996).

Several studies of childhood cancer survivors that included children whose thyroid gland was exposed to high radiation doses (>10 Gy) (Tucker 1991; deVathaire 1999, Sigurdson 2005; Ronckers 2006) have found increased risk of thyroid cancer. The relationship with radiation dose has been confirmed as curvilinear (Ronckers 2006; Bhatti 2010) using data from the Childhood Cancer Survivor Study (CCSS) which is the largest cohort study to report on thyroid cancer risk. Ronckers et al (2006) used a case-control study design with 69 thyroid cancer cases and found a suggestion that age at exposure modified the relation with dose. Bhatti et al (2010) used a cohort design and based on 119 cases confirmed the curvilinear dose-response relationship. Bhatti et al (2010) further found that age at exposure modified the excess relative risk and that gender and time since exposure modified the excess absolute risk of subsequent thyroid cancer. Most recently we have found that alkylating agents also increased thyroid cancer risk, but only in the radiation dose range under 20 Gy, with a suggestive effect for an individual alkylating agent, procarbazine (Veiga submitted). Analyses are currently underway to combine all the childhood cohorts from the US and Europe in a pooled study to further elucidate the shape of the radiation dose-response and risk modifiers (Pooled International Radiation And Thyroid cancer Epidemiology Study, PIRATES-led by Drs. Veiga and Lubin) (Veiga, in preparation).

Given the high risk associated with radiation treatment in individuals exposed at young ages and the curvilinear relationship with dose (unusual for most second solid tumors), predicting absolute risk of thyroid cancer in childhood cancer survivors is complex but would be useful for physicians when they are making screening recommendations and counseling patients. Our goal is thus to build a model to predict absolute thyroid cancer risk among childhood cancer patients who have survived to adulthood when the risk of thyroid cancer continues to be elevated, as elevated thyroid cancer risk from radiation treatment is present for decades after first exposure (reviewed in Ron 1996). To our knowledge, no such model has been developed to date.

4. Specific aims/research objectives/hypotheses to be tested

While including detailed information about a patient's treatment history including absorbed thyroid dose is expected to give the most accurate risk prediction, such a model might have limited practical use if treatment information is not readily available to the clinician. We thus propose to

build two models to predict risk of any second primary thyroid cancer in individuals who were diagnosed with a childhood cancer and survived at least 5 years past that diagnosis.

1. Model 1 will include variables that were obtained through self report in the studies, such as age at first cancer diagnosis, gender, type of first primary cancer, and thyroid conditions. A variable selection procedure will be applied to identify important risk factors among the self-reported variables which will be based on multivariable analyses with possible first order interactions.

2. Model 2 will consider the additional inclusion of more detailed information about treatment history, including yes/no indicators for the receipt of radiation or chemotherapy treatment, type of chemotherapy received, and body regions irradiated. The same variable selection procedure will be applied to identify important risk factors in multivariate analyses of the treatment and self-report variables. We regard Model 2 as a "clinical" model in contrast to the self-report model of Model 1 since treatment history for the first cancer might require chart review.

3. The "gold standard model" will include dose estimates to the thyroid (from the UTMDACC radiation dosimetry group) for radiation treatment of the first cancer, irradiated body region and type of chemotherapy in addition to information on self reported risk factors. Again, variable selection will be applied to determine important variables.

4. We will validate all models in an independent cohort and compare their predictive performance.

5. Analysis Framework:

Model building

The absolute risk A* of developing thyroid cancer (TC) in the age interval (a,b) is given by

$$A^*(a,b) = \int \lambda_t(u,x)\, S^*(u-)du / S^*(a-) \qquad (1)$$

where $S^*(a) = \exp(-\int \{ \lambda_t (u ,x) + \lambda_o (u ,x) + \lambda_M(u)\} du)$.

This models the incidence of secondary primary TC, in the presence of competing risks from all other secondary primary cancers and causes of death other than cancer captured by the hazards $\lambda_o (u ,x)$ and $\lambda_M(u)$ respectively. Let T denote the age of onset of the TC outcome, age at a second cancer diagnosis other than TC or age at death due to other causes. The cause specific hazards that may depend on covariates x are defined as $\lambda_t(a,x)= \lim_{\varepsilon \to 0} P(a \leq T < a + \varepsilon, J=TC, x)/\varepsilon$. We model $\lambda_t (a,x)= \lambda_0(a)\, rr(a,x)$ as the product of the age specific hazard rate and the relative risk part for TC, rr(a,x), that includes covariates. We may also include covariates in the hazard estimation for the competing events of other second primary cancer and death. For the patient based and clinical models the relative risks will be based on an exponential model, i.e. $rr(a,x) = \exp(\beta X)$, with possible interactions of X with age, and we will use an excess relative risk-based model that includes known (estimated from UTMDACC) dose for the gold standard model. We propose to combine data from the Childhood Cancer Survivor Study (CCSS) (Bhatti et al., 2010), the Nordic study (Svahn-Tapper et al., 2006), and the Late Effects Study Group (LESG) (Tucker et al., 1991) to estimate relative risk parameters for the models to minimize variability of the estimates, after assessing heterogeneity of the estimates across the three studies. We will do this by combining the cohort partial likelihood for CCSS with a logistic regression likelihood for the case-control studies and jointly maximizing the combined likelihood. Table 1 lists the number of cases and non-cases available for each study. We will obtain semiparametric estimates of the baseline hazard rates for thyroid cancer and competing events based on the event times within the CCSS cohort. We will extend variance computations based on influence function methods (Graubard and Fears, 2005; Pfeiffer and Petracci, 2011) to the cohort setting and when relative risk parameters are estimated from cohort and case-control data combined. Variables for the models will be selected based on statistical significance and also biological relevance.

Model validation

We propose to validate the models using the French-UK CCSS (de Vathaire et al., 1999) and to compare the performance of all models in the validation data. We will compare the expected (E) and the observed (O) numbers of cases overall and in subgroups defined by risk-factor

combinations. The expected number of cases is calculated by summing the individual projected probabilities, given the baseline covariate values for each person over the time from entry into the validation cohort to end of follow-up. The 95% confidence intervals for E/O ratios are calculated using the normal approximation to the Poisson distributions: $\frac{E}{O}e^{\pm 1.96\sqrt{\frac{1}{O}}}$. An E/O ratio above 1 indicates the model overestimates the risk of cancer and E/O<1 indicates the model underestimates the risk of cancer. We will evaluate the discriminatory accuracy of the prediction models using the area under the receiver-operating characteristic curve (AUC).

After comparing the predictive performance of the models we will decide if a risk prediction tool based on Model 1 or Model 2 might be clinically useful.

Proposed tables follow:

Table 1: Description of analytic data set, cases and non-cases, and risk factor distribution for CCSS, LESG, and Nordic data.

Characteristic	**Cases, No. (%)**			**Non-cases, No. (%)**		
	CCSS	**LESG**	**NORDIC**	**CCSS**	**LESG**	**NORDIC**
N	113	22	13	11868	82	36
Gender:						
Male	34 (30.1)	8 (36.4)	3 (23.1)	6265 (52.8)	28 (34.1)	9 (25.0)
Female	79 (69.9)	14 (63.6)	10 (76.9)	5603 (47.2)	54 (65.9)	27 (75.0)
First Primary Diagnosis						
Leukemia	27 (23.9)	NA	2 (15.4)	4061 (34.2)	NA	7 (19.4)
CNS	13 (11.5)	1 (4.5)	0 (0.0)	1543 (13.0)	3 (3.7)	7 (19.4)
HD	39 (34.5)	5 (22.7)	5 (38.5)	1503 (12.7)	15 (18.3)	1 (2.8)
NHL	6 (5.3)	2 (9.1)	1 (7.7)	876 (7.4)	5 (6.1)	1 (2.8)
Kidney (Wilms)	2 (1.8)	4 (18.2)	1 (7.7)	1045 (8.8)	21 (25.6)	2 (5.6)
Neuroblastoma	8 (7.1)	7 (31.8)	0 (0.0)	805 (6.8)	27 (32.9)	1 (2.8)
Soft tissue sarcoma)	8 (7.1)	1 (4.5)	0 (0.0)	1033 (8.7)	3 (3.7)	2 (5.6)
Bone	10 (8.8)	0 (0.0)	1 (7.7)	1002 (8.4)	4 (4.9)	3 (8.3)
Age at diagnosis of first primary (years)						
<5	28 (24.8)	11 (50.0)	3 (23.1)	4848 (40.8)	51 (62.2)	12 (33.3)
5-9	24 (21.2)	5 (22.7)	2 (15.4)	2602 (21.9)	15 (18.3)	6 (16.7)
10-14	48 (42.4)	4 (18.2)	2 (15.4)	2339 (19.7)	11 (13.4)	4 (11.1)
15+	13 (11.5)	2 (9.1)	6 (46.2)	2079 (17.5)	5 (6.1)	14 (38.9)
Follow-up (years)						
5-14						
15-24						

25+						
Age at last follow-up						
Cutpoints TBD						
Treatment Era or Birth year						
Cutpoints TBD						
Visits to a physician						
None						
Number of times, other cutpoints TBD						
Number of years since last physical examination						
Smoking						
Former						
Current						
Never						
Use of thyroid hormone replacement						
No						
Yes						
History of underactive Thyroid						
No						
Yes						
Not sure						
History of overactive Thyroid						
No						
Yes						
Not sure						
History of thyroid gland enlargement						
No						
Yes						
Not sure						
Chemotherapy						
No						
Yes						
Chemotherapy						
No						
Yes						
Chemotherapy drugs						
Alkylating agents						
Anthracylines						
Bleomycin						
Platinum						
Radiation						
No						
Yes						
Radiation Dose (Gy) (categories subject to change)						
None						
>0-<5						

5-10						
>10-20						
>20-30						
>30-40						
>40						

TBD=To be determined

Table 2a: Multivariate relative risk estimates for patient-based model (factors TBD)

Variables that will be assessed during model building include: Age at first primary, birth year, gender, time since first cancer, type of first cancer, thyroid nodule, thyroid enlargement, under- or over-active thyroid, taking thyroid medication, and personal screening practices and doctor visits (may increase thyroid cancer detection).

Table 2b: Multivariate relative risk estimates for patient- based and clinician model (may include a qualitative expression of dose if such is available for the tumor and tumor location—see appendix for possible combinations)

Table 3: Examples of absolute risk estimates for select risk factor profiles

Table 4: Observed and expected thyroid cancers in cells of covariates in the validation cohort

6. Special Considerations

In 2006, in our concept proposal that was approved by the CCSS publications committee, we had described building a prediction model nicknamed by Dr. Anna Meadows as the "yardstick", with the intent that it could be used in clinical practice during survivor follow-up visits. In the intervening years cancer risk prediction modeling has gained in sophistication and accuracy, partly due to the efforts of Dr. Ruth Pfeiffer (Freedman 2009; Park 2009; Pfeiffer 2010; O'Brien 2011). We have been fortunate to engage Dr. Pfeiffer to oversee the risk prediction effort for thyroid cancer along with her post-doctoral fellow, Dr. Stephanie Kovalchik.

References

Bhatti P, Veiga LS, Ronckers CM, Sigurdson AJ, Stovall M, Smith SA, Weathers R, Leisenring W, Mertens AC, Hammond S, Friedman DL, Neglia JP, Meadows AT, Donaldson SS, Sklar CA, Robison LL, Inskip PD. Risk of second primary thyroid cancer after radiotherapy for a childhood cancer in a large cohort study: an update from the childhood cancer survivor study. Radiat Res 2010; 174:741-52.

de Vathaire F, Hardiman C, Shamsaldin A, Campbell S, Grimaud E, Hawkins M, Raquin M, Oberlin O, Diallo I, Zucker JM, Panis X, Lagrange JL, Daly-Schveitzer N, Lemerle J, Chavaudra J, Schlumberger M, Bonaiti C. Thyroid carcinomas after irradiation for a first cancer during childhood. Arch Intern Med 1999;159:2713-9.

Freedman AN, Slattery ML, Ballard-Barbash R, Willis G, Cann BJ, Pee D, Gail MH, Pfeiffer RM. Colorectal cancer risk prediction tool for white men and women without known susceptibility. J Clin Oncol. 2009 Feb 10;27(5):686-93.

Enewold L, Zhu K, Ron E, Marrogi AJ, Stojadinovic A, Peoples GE, Devesa SS. Rising thyroid cancer incidence in the United States by demographic and tumor characteristics, 1980-2005. Cancer Epidemiol Biomarkers Prev. 2009;18:784-91.

Graubard BI, Fears TR. Standard errors for attributable risk for simple and complex sample designs. Biometrics. 2005 Sep;61(3):847-55.

Pfeiffer RM, Petracci E. Variance computations for functionals of absolute risk estimates, Statistics & Probability Letters 2011, 81 (7): 807-812

Pfeiffer RM, Gail MH. Two Criteria for Evaluating Risk Prediction Models. Biometrics. 2010 Dec 14. doi: 10.1111/j.1541-0420.2010.01523.x. [Epub ahead of print]

O'Brien TR, Everhart JE, Morgan TR, Lok AS, Chung RT, Shao Y, Schiffman ML, Dotrang M, Sninsky JJ, Bonkovsky HL, Pfeiffer RM; and the HALT-C Trial Group. An IL28B Genotype-Based Clinical Prediction Model for Treatment of Chronic Hepatitis C. PLoS One. 2011;6(7):e20904. Epub 2011 Jul 8.

Park Y, Freedman AN, Gail MH, Pee D, Hollenbeck A, Schatzkin A, Pfeiffer RM. Validation of a colorectal cancer risk prediction model among white patients age 50 years and older. J Clin Oncol. 2009 Feb 10;27(5):694-8. Epub 2008 Dec 29.

Ronckers CM, Sigurdson AJ et al. Thyroid cancer in childhood cancer survivors. A detailed evaluation of radiation dose response and its modifiers. Radiat Res 2006; 166: 618-28.

Ron E. Thyroid cancer. In: Schottenfeld D & Fraumeni JF Jr. (eds). Cancer epidemiology and prevention. 1996 Oxford Press, 2nd edition.

Sigurdson AJ, Ronckers CM, Mertens AC, Stovall M, Smith SA, Liu Y, Berkow R, Sklar CA, Hammond S, Neglia JP, Meadows AT, Robison LL, Inskip PD. Second primary thyroid cancer after a first childhood malignancy: A report from the Childhood Cancer Survivor Study. Lancet 2005;365:1014-23.

Svahn-Tapper G, Garwicz S, Anderson H, Shamsaldin A, de Vathaire F, Olsen JH, Dollner H, Hertz H, Jonmundsson G, Langmark F, Lanning M, Sankila R, Tulinius H, Moller T. Radiation dose and relapse are predictors for development of second malignant solid tumors after cancer in childhood and adolescence: A population-based case-control study in five Nordic cohorts. Acta Oncol 2006;45:438-48.

Tucker MA, Jones PH, Boice JD Jr, Robison LL, Stone BJ, Stovall M, Jenkin RD, Lubin JH, Baum ES, Siegel SE, et al. Therapeutic radiation at a young age is linked to secondary thyroid cancer. The Late Effects Study Group. Cancer Res 1991;51:2885-8.

Veiga LHS, Bhatti P, Ronckers CM, Sigurdson AJ, Smith SA, Weathers R, Leisenring W, Mertens AC, Hammond S, Neglia JP, Meadows AT, Donaldson SS, Sklar SA, Friedman DL, Robison LL, Inskip PD. Chemotherapy and thyroid cancer risk: A report from the Childhood Cancer Survivor Study (submitted to Cancer Epidemiology Biomarkers and Prevention).

Appendix from 2006 proposal:

Possible selection of representative/typical patients to illustrate effects of various patient and treatment characteristics on estimated thyroid dose

Patients with XRT			Age at Diagnosis Number of patients (% age group)				
CCSS Primary Disease	% of Primary Group	Total number of patients	<5	5 to 9	10 to 14	15 to 20	Thyroid Radiation Exposure
Leukemia		2913					
Spine XRT	15%	439	243(16%)	113(15%)	66(15%)	17(8%)	high
No Spine XRT	84%	2474	1241(84%)	655(85%)	379(85%)	199(92%)	low
CNS		1153					
C-Spine XRT	39%	453	147(39%)	159(45%)	107(37%)	40(31%)	high
No C-Spine XRT	61%	700	234(61%)	194(55%)	183(63%)	89(69%)	medium
HD		1567					
Above the diaphragm	95%	1496	36(84%)	170(89%)	512(96%)	778(97%)	medium-high
No treatment above the diaphragm	5%	71	7(16%)	21(11%)	20(4%)	23(3%)	low
NHL		636					
Above the diaphragm	57%	365	39(42%)	116(58%)	115(58%)	95(65%)	medium-high
No treatment above the diaphragm	43%	271	53(58%)	83(42%)	83(42%)	52(35%)	low
Wilms/Kidney		693					
Chest treated	38%	263	157(33%)	93(51%)	11(41%)	2(29%)	medium
No chest treatment	62%	430	318(67%)	91(50%)	16(59%)	5(71%)	low
Neuroblastoma		412					
Chest or neck	47%	195	165(46%)	21(57%)	8(73%)	1(25%)	medium-high
No chest or neck	53%	217	195(54%)	16(43%)	3(27%)	3(75%)	low
Soft Tissue Sarcoma		686					
Chest or neck or face	58%	396	141(57%)	119(73%)	76(53%)	60(46%)	medium-high
No chest or neck or face	52%	290	105(43%)	45(27%)	68(47%)	72(55%)	low
Bone tumors		382					
Chest or neck or face	48%	197	13(65%)	38(49%)	73(53%)	61(42%)	medium-high
No chest or neck or face	52%	185	7(35%)	40(51%)	65(47%)	85(58%)	low
TOTAL		8442					