# Childhood and adolescent factors and thyroid cancer incidence in adult women in the Sister Study cohort

**Running title: Early-life factors and thyroid cancer incidence**

Thi-Van-Trinh Tran Ph.D.1, Katie O’brien Ph.D.2, Rebecca Troisi Ph.D.3, Dale Sandler Ph.D.2, Cari M. Kitahara Ph.D.1\*

1 Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 9609 Medical Center Drive, Room 7E-456, Bethesda, Maryland 20892-9778, USA.

2 Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.

3 Trans-Divisional Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 9609 Medical Center Drive, Room 7E-456, Bethesda, Maryland 20892-9778, USA.

\* Corresponding author

**Keywords**: Thyroid neoplasms, incidence, childhood exposures, adolescent exposures, risk factors, observational study

**Word count**: 2623

# Abstract (309 words)

**Background:** Thyroid cancer ranks as the fifth most common cancer in women, with 90% of cases categorized as differentiated thyroid cancer (DTC). DTC often affects individuals at younger ages, particularly females starting from adolescence, with reasons remained unclear. Therefore, our study aimed to investigate the association between childhood and adolescent factors and the incidence of DTC in adult females.

**Methods**: From the U.S. nationwide cohort Sister Study, we examined data of 47913 cancer-free women at baseline (2003–2009). Using Cox regression models, we assessed the association between baseline self-reported growth, reproductive, lifestyle, and socioeconomic factors during childhood and adolescence and self-reported DTC incidence during the follow-up. We adjusted our models for attained age (timescale), race/ethnicity, BMI, smoking status, personal history of benign thyroid disease, educational level, household income, and Area Deprivation Index.

**Results**: At baseline, the median age at baseline was 55.4 (interquartile range: 48.9‑62.1). Over the follow-up (median: 13.1 years), there were 239 DTC cases reported, of which, 72% had available medical records. Factors associated with a higher incidence of DTC included a taller height and at age 10 (hazard ratio [HR] = 1.41, 95% confidence interval [CI] 1.06-1.89), a lighter (HR=1.28, 95%CI 0.96-1.71) or heavier (HR=1.37, 95%CI 0.97-1.91) weight during adolescence compared to peers, and experiences of not having enough to eat during childhood (HR=1.67, 95%CI 1.15-2.43). Conversely, a lower incidence was observed among those with a household education level of bachelor’s degree or higher at age 13 compared to those with a high school or GED or less (HR=0.75, 95%CI 0.55-1.03). These associations were independent of baseline BMI and other socioeconomic factors during childhood and adulthood, such as household income, and educational level. No significant associations were found for other childhood and adolescent exposures.

**Conclusions**: Our study suggests that experiences during childhood and adolescence may contribute to the development of female thyroid cancer later in life.

# Introduction

Thyroid cancer, primarily differentiated thyroid cancer (DTC), ranks as the 13th most common cancer diagnosis in general and the fifth most common in women globally.1 Despite decades of extensive research, the understanding on thyroid cancer risk factors remains limited beyond obesity and radiation exposure during childhood.2

Emerging evidence suggests that early-life factors may influence thyroid cancer development. While most benign and malignant thyroid nodules typically grow slowly, thyroid cancer is often diagnosed at relatively young ages compared to other cancers.2 A significant gender disparity in thyroid cancer incidence emerging as early as adolescence, with women being more likely than men to develop the disease. Furthermore, radiation exposure at a younger age has a more pronounced effect on thyroid cancer risk, while accumulating data demonstrate the link between obesity and height in early life and thyroid cancer risk, regardless of adult anthropometric status. Experimental studies have suggested several mechanisms to explain these observations, including heightened susceptibility of the thyroid gland, and hormonal changes, both external (e.g. oral contraception) and internal (e.g., cortisol, insulin, growth hormone, and the sex steroids) during childhood and adolescence. However, previous epidemiologic studies, predominantly focusing on adult populations, have reported inconsistent findings on the association between these hormonal characteristics and thyroid cancer risk.

Owing to limited longitudinal studies with comprehensive data on early-life factors and long-term follow-up for thyroid cancer,3 the roles of growth, reproductive, and lifestyle factors during childhood and adolescence on thyroid cancer development are unclear. Additionally, the impact of socioeconomic status during childhood and adolescence is poorly understood, despite its known association with adult socioeconomic disadvantages and unhealthy behaviors in adulthood, both of which have been shown to influence the incidence of thyroid cancer.

To address these gaps, this study aims to investigate the associations of a wide range of childhood and adolescent factors - including growth, reproductive, lifestyle, and socioeconomic factors - and thyroid cancer incidence, using data from the Sister Study, a large U.S. prospective nationwide cohort.

# Methods

**Study population**

The study design, data collection, and outcome measurements of the Sister Study cohort have been previously described. Briefly, the Sister Study is a nationwide prospective cohort of 50,884 women aged 35–74 in the U.S. (including Puerto Rico), enrolled between 2003–2009. All participants had a sister with breast cancer but were breast cancer-free themselves at baseline. Baseline data, including information on early-life, demographic, medical, and lifestyle factors, were collected with self-administered questionnaires and computer-assisted telephone interviews. Anthropometric measurements and biospecimens were obtained via in-person home visits. Participants are recontacted every 2-3 years for health and lifestyle updates, with response rates consistently exceeding 85%. All participants provided written informed consent, and the National Institutes of Health's institutional review board approved the study.

We excluded individuals who withdrew from the study (n=5) and those with a history of invasive cancer (n=2911), or chemotherapy or radiotherapy for cancer (n=55) before baseline. The study population comprised 47,913 individuals (Figure 1). Follow-up time was accounted from baseline to the date of the first diagnosis of invasive cancer (excluding non-melanoma cancer), death, loss of follow-up, or mid-September 2021, whichever occurred first, unless otherwise specified.

**Outcome definition:**

By the end of follow-up in mid-September 2021 (data release 11.1), 252 women had reported a diagnosis of thyroid cancer. Among these, 188 cases (74.6%) were confirmed through medical records/pathology reports (N=187) or National Death Index/death certification/next of kin (n=1). We defined DTC cases (n=239) as all self-reported cases excluding those with poorly differentiated thyroid carcinoma (confirmed with pathology reports; n=5), anaplastic thyroid carcinoma (histology code: 8021; n=1), medullary thyroid carcinoma (histology codes: 8346, 8347, 8510, n=5), and indeterminate (histology code: 8265, 9084, n=2). We further classified DTC cases with confirmed histology codes (n=174) into papillary thyroid carcinomas (histology codes: 8050, 8260, 8340-8344, 8350, 8450-8460, n=164), follicular thyroid carcinomas (histology codes: 8290, 8330-8335, n=7), and unspecified carcinomas and neoplasms (histology code: 8000, 8010, n=3).

**Exposure definition:**

Data on childhood and adolescent exposures were ascertained through baseline questionnaires. Growth and reproductive factors included weight at age 10 vs. peers, height at age 10 vs. peers, teenage weight vs. peers, age reached full adult height, age at breast development, age at menarche, start age for hormonal birth control. Lifestyle factors encompassed physical activity before age 20 (MET hours/week), age started drinking regularly, alcohol consumption under 20, age started smoking, smoking before 20, and total years of secondhand smoking under 18. Socioeconomic status during childhood and adolescence was determined by family income while growing up, ever not having enough to eat during childhood, highest household education level at age 13, household composition at age 13, and childhood residence. For all of these, we considered responses of “definitely” and “probably” as exposed and categorized “probably not” and “definitely not” as unexposed. Missing data were classified using missing indicator variables. Add details for age at menarche, thelarche (>21yo 🡪 improbable 🡪 missing value)

**Covariates**

In-person measurements of weight and height at baseline were used to calculate body mass index (BMI, weight in kilograms divided by the square of height in meters). Information on other characteristics, including age, self-identified race/ethnicity (Non-Hispanic White, Non-Hispanic Black, Hispanic, Non-Hispanic all other races, Unknown), personal attained educational level, personal history of benign thyroid disease, and household annual income were assessed during the computer-assisted telephone interview. Individuals were considered to have a history of benign thyroid disease if they reported a diagnosis of hyperthyroidism, hypothyroidism, non-toxic goiter/thyroid nodules, or any other thyroid conditions, or if they ever used thyroid hormone substitutes, antithyroid, or iodine/iodide medication.

**Statistical analysis**

We used Cox proportional hazards models to estimate hazard ratios (HRs) and 95% confidence intervals (CIs), with attained age as the time scale and self-reported race/ethnicity as covariates. Fully-adjusted models included the potential confounders described earlier. We assessed proportional hazards assumptions with plots of scaled Schoenfeld residuals against attained age, and formal testing included introducing an interaction term between exposures and attained age. No evidence of violation was found.

We considered health- and medical surveillance-related factors during adulthood as potential mediators for the associations between childhood and adolescent factors and adult thyroid cancer. Therefore, we conducted stratified analyses by baseline examiner-measured BMI, personal educational level, and household annual income.

The impact of carcinogenic exposures, especially when occurring during the early-life periods, may vary between early- and late-onset cancers, defined as being diagnosed before and after 50 years of age. Therefore, we conducted a sensitivity analysis, accounting for early and late-onset cancers separately. For the late-onset analysis, we restricted the study population to individuals of 50 years of age or more at baseline. For the early-onset analysis, we included only individuals under 50 at baseline, with the 50th birthday added as an additional endpoint for follow-up. We also performed sensitivity analyses restricted to medically confirmed cases, and papillary thyroid carcinomas, separately. We conducted complete case analysis by excluding individuals with missing data on a model-specific basis and accounted for competing risks, i.e., death and diagnosis of invasive cancer other than thyroid and non-melanoma cancers, using Fine-Gray models. Lastly, we calculated E-values for both the observed association estimates and the limit of the confidence interval closest to the null. The E-value is defined as the minimum strength of association that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain the observed associations.

Data analyses were conducted using SAS 9.4 and R version 4.3.1.

# Results

The median age at baseline was 55.4 years (IQR: 48.9-62.1). Table 1 presents the baseline descriptive statistics. Most women were non-Hispanic White (n=39,863, 83.4%), and 61.7% (n=29487) were categorized as having a BMI over 25. Over half of the participants held a bachelor's degree or higher (n=24,408, 51.1%), while 33.7% reported a household annual income of $100,000 or more (n=16,108).

Table 2 shows multivariable-adjusted HRs for childhood and adolescent factors and DTC incidence. During the follow-up (median 13.1 years, interquartile range, IQR 11.5-15), 239 cases of incident DTC were reported. Being taller than peers at age 10 (HR=1.41; 95%CI 1.06–1.89) and lighter (HR=1.37, 95%CI 0.97-1.91) or heavier than peers during adolescence (HR=1.28, 95%CI 0.96-1.71) were associated with higher DTC incidence. Women who ever used hormonal birth control had a lower incidence of DTC, with no variation in risk estimates between those starting before (HR=0.72, 95%CI 0.50-1.05) or after 20 years of age (HR=0.69, 95%CI 0.48-1.00). Ever not having enough to eat during childhood (HR=1.67, 95%CI 1.15–2.43) was associated with a higher DTC incidence. Conversely, women from household where the highest education level at age 13 was a bachelor’s degree or higher had a lower DTC incidence (HR=0.75, 95%CI 0.55-1.03) compared to those from households with high school education or GED or less. Other growth, reproductive, lifestyle, and socioeconomic factors examined were not associated with DTC incidence.

When simultaneously adjusting for height at age 10 vs. peers, teenage weight vs. peers, highest household educational level at age 13, and ever not having enough to eat during childhood, all these factors remained associated with DTC incidence. Further adjustment for other childhood socioeconomic factors showed a more pronounced association for ever not having enough to eat during childhood (HR=1.92, 95%CI 1.26-2.90) (Table 3). Upon stratification by baseline socioeconomic status and BMI categories, the associations for taller height at age 10 vs. peers, lighter weight during adolescence vs. peers, ever not having enough to eat during childhood, and highest household educational level at age 13 remained consistent (p-interactions>0.05, Supplementary Figure 1), except for the higher DTC incidence associated with being lighter than peers during adolescence, which varied according to the baseline BMI (p-interaction=0.01).

Because of the notable interaction between teenage weight vs. peers and baseline BMI categories, we examined the joint associations of these factors and DTC incidence (Figure 2, Supplementary table 1). The reference group was women with a baseline BMI under 25 who reported having same weight as peers during adolescence. Among women with a baseline BMI under 25, both being lighter (HR=2.70, 95%CI 1.61-4.50) or heavier (HR=2.34, 95%CI 1.06-5.13) than peers during adolescence were associated with higher DTC incidence. For those with a baseline BMI over 25, we observed increased DTC incidence regardless of teenage weight perception, except for those with a BMI over 30 who reported being lighter than peers during adolescence (HR=1.55, 95%CI 0.70-3.41), based on 9 cases. We repeated these analyses with weight compared to peers at age 10 instead of during adolescence and similar results were observed (Supplementary table 2).

The E-values for the associations between thyroid cancer incidence and being taller than peers at age 10, ever not having enough to eat during childhood, and having a bachelor’s degree or higher as the highest household education level at age 13 were 2.18, 2.73, and 1.99, respectively. The association for height at age 10 vs. peers, ever not having enough to eat during childhood, and highest household education level at age 13 appeared to be stronger in early-onset cancers. Results for the analyses on teenage weight vs. peers did not vary according to early- or late-onset status. The finding interpretation did not change when considering competing risk, medically confirmed cases, or papillary histology only (Supplementary table 3).

# Discussion

To our knowledge, this is the first longitudinal study conducting comprehensive analyses of childhood and adolescent factors on DTC incidence beyond anthropometric factors, and ionizing radiation exposure. We found that DTC incidence was associated with being taller than peers at age 10, being lighter or heavier than peers during adolescence, used of hormonal birth control, and ever not having enough to eat during childhood. Conversely, having higher household levels of education at the age of 13 was associated with a lower DTC incidence. These findings provide support to the independent influence of early-life factors on DTC incidence.

Our findings, consistent with previous research <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8971972/>, showed a positive association between heavier weight in early life and higher adult DTC incidence (Cari pooled cohort), even among women with a baseline BMI under 25. Additionally, we observed that the joint effects of heavier weight in childhood and adolescence and overweight and obesity in adulthood do not demonstrate a stronger association with DTC incidence compared to being overweight or obese solely in adulthood, indicating a limited cumulative effect of adiposity over time. Furthermore, we found an unexpected higher risk of DTC among women with a lighter weight than peers during childhood and adolescence, which attenuated among women with a baseline BMI over 30. Taken together, these findings suggest the influence of both excess and low weight during childhood and adolescence on adult DTC incidence and the existence of distinct underlying causes beyond adiposity-related pathways.

During childhood, ever not having enough to eat and the highest household levels of education at the age of 13 were independently associated with DTC incidence, even when adjusting for adult potential and well-known risk factors. As no notable associations were found for other childhood socioeconomic factors in this study and the association for ever not having enough to eat during childhood was stronger when considering childhood socioeconomic factors, socioeconomic status might not be the primary explanatory factor. Alternatively, these findings may serve as indicators of the influence of early-life nutrition. While some dietary factors, including iodine ([3](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4575832/#R3)), goitrogens ([4](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4575832/#R4)), fat ([5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4575832/#R5)), and flavonoids ([6](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4575832/#R6)), have been suggested to be associated with thyroid cancer risk, (for ref: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4575832/#S7title>), understanding the impact of dietary factors on DTC incidence remains limited. Given the interplay of diet, lifestyle, and socioeconomic factors, future research should adopt a life course framework, utilizing longitudinal data with accurately measured both childhood and adulthood exposures to confirm these findings, disentangle confounding effects, and estimate the impact of individual risk factors on DTC risk.

One paragraph for Biological mechanisms

Strengths of this study include the large sample size, long follow-up, and wide range of childhood and adolescent exposures. Our analysis also included an evaluation of potential factors that could influence the likelihood of incidental detection of thyroid cancer, such as baseline BMI, household annual income, and educational levels; the strength of the associations did not differ substantially by these factors. Thus, overdiagnosis is unlikely to be the main explanatory factor in our study. However, the current study has several limitations. As the Sister Study enrolled exclusively women with a sister diagnosed with a breast cancer, the results may not be generalizable to men or to women without a family history of breast cancer. The data, collected between 1930s and 1970s, may not reflect modern context, considering that the lifestyles and environmental exposures during early life and young adulthood changed substantially over time. For example, the earlier timing of the study data is likely to be less exposed to obesogenic diet, lifestyle, and environmental factors that have become more commonplace since the 1980s (<https://www.niddk.nih.gov/health-information/health-statistics/overweight-obesity>). Owing to the reliance on personal recollection on childhood and adolescent exposures as the main source of data, assessment of early-life exposures may be prone to recall and misclassification bias. Although we found high E-values of the analyses concerning maternal complications and birth weight, which indicate that additional unmeasured confounding associated with early-life factors and thyroid cancer incidence by a HR of at least 2 to 3-fold each would be necessary to explain the observed associations, the absence of information on diet during childhood and adolescence is an important limitation. Lastly, we did not have information on exposure to ionizing radiation in childhood of thyroid cancer, but it is unlikely to be a strong confounders as it is not associated with childhood and adolescent exposures.

# Conclusion

In this study, certain childhood and adolescent exposures were associated with DTC incidence, particularly early-onset cancers. These results support the impact of early-life exposures on DTC incidence later in life, and offer further insights into understanding the DTC early age at diagnosis and the gender disparity in DTC incidence.

# References

# Author Contribution statement

T.V.T.T. contributed to the analysis, interpretation, drafting of the article, and approval of the final version; C.M.K. contributed to the concept, interpretation, critical review of the manuscript, and approval of the final version; K.O., R.T., and D.S. contributed to the interpretation, critical review of the manuscript, and approval of the final version.

All authors have no disclosure to report.

# Funding statement

This research was funded by the Intramural Research Program of the National Cancer Institute, and the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health (Z1AES044005 to D.P.S.).

Correspondence to:

Cari M. Kitahara, PhD, MHS, 9609 Medical Center Drive, Rm 7E-456, Bethesda, MD 20892, kitaharac@mail.nih.gov