January 2016

Residential Proximity To Manufacturing Facilities And Risk Of Thyroid Cancer

Nathan Kloczko
Yale University, nathan.kloczko@yale.edu

Recommended Citation

Kloczko, Nathan, "Residential Proximity To Manufacturing Facilities And Risk Of Thyroid Cancer" (2016). *Public Health Theses*. 1152.
http://elischolar.library.yale.edu/ysphtdl/1152

Residential proximity to industrial manufacturing facilities and risk of thyroid cancer

Nathan Kloczko

April 2016

Research Advisor and First Reader: Nicole Deziel, Assistant Professor,
Environmental Health Sciences
Second Reader: Yawei Zhang, Associate Professor, Environmental
Health Sciences

Abstract

In the past two decades, thyroid cancer incidence has increased at a rate faster than any other malignancy. Environmental contamination has been a suspected risk factor in this trend. One potential source of environmental exposures is pollution from industrial manufacturing facilities. We investigated whether proximity to industrial facilities was associated with an increased risk of thyroid cancer in an exploratory analysis within an existing population-based case-control study in Connecticut using a novel data resource.

Complete residential histories of 408 thyroid cancer cases and 470 controls were collected and geocoded. Manufacturing facility addresses were gathered from a novel source of publicly available data, the Connecticut Point Source Inventory, and categorized by 2-digit Standard Industrial Classification (SIC) code from the years 1990-2009. Binary proximity exposure metrics were created and were defined as ever having lived within 5 km or 2 km of a facility in any manufacturing sector. Additional facility-specific binary markers were created that defined if an individual had ever lived within 5 km or 2 km of each specific SIC explored in this study (20-38).

Additionally, for each participant, a cumulative inverse distance-weighted (IDW) metric was calculated for all facilities within 5 km of all residences lived in between 1990 and 2009. SIC-specific IDW metrics were also created. Odds ratios (OR) and confidence intervals (CI) detailing the relationship between each proximity metric and thyroid cancer were calculated using logistic regression, adjusting for potential confounders. Cumulative IDW was evaluated as a continuous variable, a natural-log-transformed variable, and as a categorical variable.

In both unadjusted and adjusted analyses, nonsignificant elevated risk estimates were observed among individuals who lived within either 5 km (adjusted OR = 1.15, 95% CI: 0.86 – 1.54) or 2 km of chemical facilities (SIC 28, adjusted OR = 1.20, 95% CI: 0.84 – 1.73). After adjustment, ever having lived within 5 km of lumber and wood products facilities (SIC 24, OR = 0.64, 95% CI: 0.43 – 0.94) or within 2 km of transportation equipment facilities (SIC 37, OR = 0.54, 95% CI: 0.31 – 0.96) was associated with decreased risk of thyroid cancer. All other associations were null. Analyses of the cumulative IDW exposure data revealed similar results.

The results of this exploratory analysis do not appear to support a link between residential proximity to manufacturing facilities and increased risk of thyroid cancer. However, this study

was limited by crude proximity metrics and imprecise recall of residential histories, which potentially led to exposure misclassification. Future studies could refine exposure metrics by incorporating emissions and meteorological data or by refining the accuracy of the residential addresses.

Introduction

In recent years, the incidence of thyroid cancer has increased at a faster rate than any other malignancy, especially in women.[1] In 1988, the incidence rate of thyroid cancer among women was 6.88 cases/100,000 individuals; from there it doubled to 13.42 in 2002 and is rapidly approaching another doubling, with rates of 21.67 in 2012.[1] Though it is not often a deadly disease (death rates < 0.5/100,000 from 1975-2012 for both genders, > 90% survival after 20 years),[1] there are significant costs associated with the surveillance, diagnosis, and treatment of these cancers. Additionally, individuals with differentiated thyroid cancer have a greater chance of developing secondary malignancies and a lower quality of life.[2,3]

The etiology of thyroid cancer remains largely unclear. Though some of this increased incidence is likely partially explained by increasing diagnostic capabilities, increased access to care, and overdiagnosis, recent evidence has suggested that environmental exposures may constitute between 30-50% of new incident cases.[4,5] Radiation, diet, and chemical exposures are just a few of the many ways thyroid function can be modulated.[6] For example, exposure to radiation through environmental disasters such as Chernobyl or through diagnostic radiography such as computed tomography scanning and nuclear medicine imaging have been shown to increase an individual’s odds of thyroid cancer.[5,7] Additionally, studies have demonstrated that women consuming high levels of nitrites from contaminated water and processed meats have an increased risk of thyroid cancer.[8]

In addition to the aforementioned exposures, an increasing range of industrial chemicals frequently present in the environment has been shown to interfere with typical thyroid function. Chemicals within the family of polyhalogenated aromatic hydrocarbons (PHAHs), such as polychlorinated biphenyls (PCBs), organochlorine pesticides (OCs), dioxins, and polybrominated diphenyl ethers (PBDEs) have all been observed to disrupt thyroid function, increasing the likelihood of autoimmune thyroid disease and thyroid cancer.[4,6,9] Each of these are persistent organic pollutants (POPs): stable, lipophilic compounds that tend to bioaccumulate in the environment through time. Though many of the chemicals in these families have been phased out of production due to their toxicity, many chemosimilar products have been introduced into the market as replacements.[10] The health consequences of these replacements have not yet been fully realized.

The general mode of action of these chemicals is to disrupt the physiological processes and hormone production of the thyroid.[4] For example, PBDEs are able to bind to thyroid hormone receptors and inhibit the binding of thyroid hormones to transport proteins.[6] Dioxins, OCs, and PBDEs are able to induce hepatic uridine diphosphate glucuronyltransferases, which glucuronidates thyroxine, a major thyroid hormone.[4,6] This conjugation increases clearance from the body, decreasing thyroxine's half-life. These hormonal changes may potentially lead to chronic thyroid stimulation and ultimately tumorigenesis.[8]

One potential environmental source of these thyroid disrupting chemicals is from industrial facilities. Previous studies have observed that PCB and dioxin concentrations in carpet dust increased for residences in close proximity to industrial facilities.[11,12] Additionally, excess thyroid cancer risk was observed in a community of 5,000 near an unintentional industrial release of the organochlorine pesticide hexachlorobenzene.[9] Proximity to industrial facilities has been previously used as a surrogate for exposure.[13-16]

The objective of this study was to explore whether residential proximity to industrial manufacturing facilities was associated with increased odds of thyroid cancer within a population-based case-control study in Connecticut. An existing case-control set that had collected historical information on participant residences was used in conjunction with industrial facility locations from the Connecticut Point Source Inventory, provided by the Connecticut Department of Energy and Environmental Protection. Industries with *a priori* interest included those that potentially emit PHAHs or other xenobiotics with known thyroid antagonism.

Methods

Study Population

The study population for this research is comprised of individuals enrolled in a population-based case-control study conducted by Zhang et al. to investigate the impact of diagnostic radiation on DNA repair capacity and thyroid cancer risk.[5] From 2010-2011, 462 Connecticut residents newly diagnosed with histologically confirmed papillary, follicular, medullary, or anaplastic thyroid cancer were enrolled (66% of the 701 eligible thyroid cancer cases) and completed in-person interviews. These interviews collected information detailing demographic information, diagnostic radiation use, occupational and residential histories, as well as other potential risk factors for thyroid cancer. Participants were between the ages of 21 and 84, had no previous cancer diagnoses (aside from nonmelanoma skin cancer), and were alive at the time of the interview. Cases were identified through the Yale Cancer Center's Rapid Case Ascertainment Shared Resources, a component of the Connecticut Tumor Registry. Since Connecticut public health code requires the reporting of all cancer cases, this tumor registry has high case coverage through the state. 498 controls were recruited by random-digit dialing Connecticut residents. The participation rate for controls was 62%. Controls were frequency matched to cases by age ± 5 years.

Residential Locations

Complete residential histories for cases and controls were collected as part of the interviewer-administered questionnaire. Participants were asked to provide complete addresses (street number and name, town/city, state, and country if applicable) of each home they lived in from birth to the date of the questionnaire, including year moved and age at move. Addresses were cleaned to correct spelling errors. Residential addresses were geocoded using ArcGIS (Version 10.2, ESRI, Redlands, WA). Information for 5,155 addresses were collected from the 960 participants. Initially, all residences with complete address information (n = 648) were geocoded. Of these, 618 (95%) street addresses were assigned coordinates accurate to the exact parcel in the first round. Coordinates for the final 30 addresses were found using interactive geocoding processes.

If the resident's complete address was not recorded during the interview, coordinates were assigned at the finest geographic resolution possible. Reported residences that only contained street segment information (n=1,197) were assigned a latitude and longitude that corresponded to the street segment midpoint using ArcGIS. If the location reported by the study participant was a military base, college, or village/neighborhood (n = 188), an appropriate ZIP was used. For reported addresses that contained only city (n = 2,681), county (n = 16), or state (n = 147) information, the geographic centroid of the appropriate region was calculated via the 'Feature to Point' and 'Add XY Coordinates' tools in ArcGIS. Since there were a significant number of incomplete addresses, the quality of each address was marked for later sensitivity analyses, as previously described.[17] The area of the matched geographic resolution was calculated. Zip code, town, county, and state areas were calculated as the total area of the polygon in ArcGIS, while street segment areas were calculated as the area within a 10 m setback on each side of the road (Supplemental Table 1).[17] Parcel area was not collected in the study questionnaire, and was unable to be calculated. After restrictions, 1,955 addresses reported between 1990 and 2009 were included in the analyses (Table 2).

Industrial Facility Locations and Classification

Facility information was obtained from the Connecticut Department of Energy and Environment's (DEEP's) Point Source Inventory, a component of the Connecticut Emissions Inventory.[18] All facilities within Connecticut that have a Title V permit under the Clean Air Act are required to report to DEEP facility information and yearly emissions of a number of criteria air pollutants. DEEP assembles this information in a database containing the facility address, quantity of pollutant emitted, and facility information for the years 1967-2015. All historical data prior to 1990 was incomplete and updated infrequently, with methodological yearly collection beginning in 1990.[19] To avoid unknown exposure classification error, industrial facilities were only included from 1990 onward. Latitude and longitude were present in the database, but were assumed to have some positional error, because many sites had been manually geocoded with topographical maps. To improve the accuracy of the facility coordinates, the addresses were geocoded using ArcGIS following an approach similar to that of the residential facilities. Geocoding using complete address information appropriately matched 76% of the facilities (n = 15,221). Sites that could not be matched via geocoding were assigned the existing latitude/longitude coordinates within the database. Facilities that had known inaccurate coordinates (e.g. for the Connecticut Department of Public Health headquarters) and could not be geocoded were removed from the analysis (n = 280). Manufacturing facilities were classified according to their 2-digit primary Standard Industrial Classification (SIC) number.[20] This number, defined by the Occupational Safety & Health Administration, identifies each facility by its primary manufacturing sector. SIC numbers 20-38 were considered in this study. Distances between each residence and each active facility within a 5 km buffer region of the residence for each year during 1990-2009 were calculated.

Geographic Exposure Assessment

A variety of exposure metrics were constructed to explore proximity and thyroid cancer risk. First, a binary exposure was created that determined whether an individual had ever lived within 5 km or 2 km of any manufacturing facility. Figure 1 provides a visual estimation of the binary 2 km exposure metric, using data from the year 2000 as an example. Second, binary metrics were created for ever living within 5 km or 2 km of each of the 18 specific manufacturing sectors. Third, a cumulative inverse distance weighted (IDW) facility count was calculated for each participant (Equation 1) using all residences and facilities within 5 km of the home from 1990-2009.

$$IDWFacilityCount = \sum_{i=1}^{n} \sum_{j=20}^{38} \left(\frac{1}{d_{ij}} \right)$$

IDW Facility Count is the IDW result from manufacturing facilities within a 5 km radius of participant residence for a given SIC in a given study year. d_{ij} is the distance to each given facility (i) of each given SIC (j) to the participant residence in a given year, and n is the total number of facilities within 5 km in a given year. IDW Facility Count was then summed across the years 1990-2009 to provide a cumulative IDW facility count. If a study participant reported living in more than one residence in a given year, it was assumed they spent half the year in each residence, because the date of the move was not recorded.

Statistical Analyses

Since accurate facility data was reported from 1990 onwards, initial restrictions were limited to resident addresses from 1990 until 2009, a reference date 1-2 years before diagnosis. We then conducted analyses on two subsets of the parent study population. First, we included all participants with a geocoding accuracy of town centroid or better for at least 70% of the years included in the study analysis, from 1990-2009 (432 cases, 470 controls). We also created a subgroup that had 100% of addresses geocoded to an accuracy of town centroid or better during the study period (408 cases, 436 controls). Analyses presented in the present study were constructed using this second subgroup, with 100% of cases, in an attempt to optimize accurate exposure classification. Many of the state and county centroids were located in areas with a high number of facilities, which would overestimate exposures for those participants. Therefore, using the 70% accuracy group would increase nondifferential misclassification, driving any observed odds estimates to the null. Table 1 compares demographic characteristics of the population used in this study with the parent population. Another subgroup was created that only explored exposures from 1990-1999, to test the impact of a longer latency period. SICs 21 and 31 were excluded from all analyses, because no cases ever lived within 5 km of a facility of either sector. Cumulative IDW exposure was evaluated as a continuous variable, a natural log-transformed variable, and as a categorical variable (separated into exposure quartiles based on control exposures).

Univariate analyses were conducted on select demographic variables and cumulative inverse IDW exposure to ascertain distributions for categorical analyses. Bivariate analyses were used to compare exposure distributions for the binary exposure metrics. P values were obtained from a t-test for univariate analyses and a χ^2 test for bivariate analyses.

Unconditional logistic regression was used to determine odds ratios (ORs) and 95% confidence intervals (CIs) for all individuals with non-missing variables. To control for potential confounding, ORs and 95% CIs were generated via multivariate logistic regression while adjusting for age, BMI (<25, 25-29.9, ≥30), gender, race, education, prior alcohol use, smoking, family history of thyroid disease, family history of thyroid cancer, and previous exposure to diagnostic radiation. Backwards, stepwise elimination was used to determine covariate significance: race, education, and smoking status were eliminated as nonsignificant variables from the final model. Previous exposure to diagnostic radiation was retained despite nonsignificance ($p = 0.13$), because it is an established risk factor for thyroid cancer within this population. All statistical tests were conducted at $\alpha = 0.05$, with a $p < 0.05$. All statistical analyses were conducted using SAS 9.3 (SAS Institute Inc., Cary, NC).

Results

Demographic Characteristics

In this study, 462 cases and 498 controls were analyzed to determine the risk of thyroid cancer due to residential proximity to industrial facilities. After excluding individuals who did not have address recall to the city resolution or better for 100% of the study years, 408 cases and 436 controls remained in the final analyses. This subset was similar to the parent population with respect to both demographic characteristics and collected risk factors. Table 1 describes several demographic characteristics and the distribution of important risk factors for both the case and control groups of the entire population and study population. Within this study population, a number of characteristics significantly differed between cases and controls. Cases were more likely to be female (p = <0.001), younger (p = 0.001), and less educated (p = 0.009). They were also more likely to have a family history of thyroid cancer (p = 0.002) or thyroid disease ($p < 0.001$). Cases were more likely to be obese (p = 0.001), but less likely to have a history of alcohol consumption ($p < 0.001$). No significant changes between the two groups were seen for race, smoking history, or diagnostic radiation exposure.

Geographic Recall and Exposure to Industrial Facilities

Residential recall was statistically similar between cases and controls in the restricted study population, during the study period (Table 2, p = 0.41). 23% of cases and 23% of controls reported complete street addresses for their residence during the study period, while 45% of cases and 48% of controls could only recall their address to the city level (Table 2). ZIP code information was not collected in the interview process, which may explain why the majority of addresses were recorded to the city level. Some differences were observed between cases and controls when considering their lifetime residential history (Supplemental Table 1).

Over the course of the entire study, 88% of both cases and controls lived within 5 km of at least one manufacturing facility during the study period (Table 3). The number of facilities within 5 km of a residence in a single year ranged from 0 to 110. Residence within 2 km (58% of cases and 59% of controls, Table 3) or 1km (32% of cases and 35% of controls, data not shown) of any facility was less common, though still fairly frequent.

Cases and controls were most likely to have ever lived within both 5 km (cases: n = 265, 65%; controls: n = 274, 63%) and 2 km (cases: n = 137, 34%; controls: n = 138, 32%) of fabricated metal products facilities (SIC 34, excludes machinery and transportation equipment). The second most common manufacturing sector for cases to have ever lived with 5 km of was electronic and other electrical equipment and components (SIC 37, excludes computer equipment; n = 215, 53% of cases), whereas the second most common for controls was rubber and miscellaneous plastics products (SIC 30, n = 214, 49%). Complete exposure information can be seen in Table 3.

Neither cases nor controls had significantly different exposures to any manufacturing sectors for our binary metrics. Only one sector was borderline significant—more cases lived within 2 km of chemical facilities (SIC 28, p = 0.06), which may have contributed to the elevated risk estimates seen for that sector.

Exposure to Industrial Facilities and Risk of Thyroid Cancer

Several risk models were constructed using different exposure metrics to determine the impact residential proximity to industrial manufacturing facilities has on thyroid cancer. In unadjusted analyses (Supplemental Table 2), no significant changes in risk were seen from ever having lived within 5 km (OR = 0.99) or 2 km (OR = 0.96) of any manufacturing facility as compared to never having lived within 5km of a facility. Increased risk estimates of thyroid cancer were seen in individuals that had ever lived within 2 km of textile mill product facilities (SIC 22, OR = 1.52, 95% CI: 0.81 – 2.82), chemical facilities (chemicals and allied products; SIC 28, OR = 1.39, 95% CI: 0.98 – 1.96), and paper facilities (SIC 26, paper and allied products; OR = 1.21, 95% CI: 0.70 – 2.09), as compared to participants that were not exposed. Interestingly, decreased odds of thyroid cancer were observed in individuals that had ever lived within 2 km of transportation equipment facilities (SIC 37, OR = 0.63, 95% CI: 0.37 – 1.07), lumber and wood products, except furniture (SIC 24, OR = 0.70, 95% CI: 0.35 – 1.40), and printing and publishing facilities (SIC 27, OR = 0.78, 95% CI: 0.49 – 1.26), as compared to participants that had not lived near those facilities.

After adjusting for age, BMI, gender, prior alcohol use, family history of thyroid disease, family history of thyroid cancer, and previous exposure to diagnostic radiation, the odds of thyroid cancer were lower if a study participant had ever lived within 5 km (OR = 0.87, 95% CI: 0.56 – 1.34) or 2 km (OR = 0.77, 95% CI: 0.58 – 1.04) of any manufacturing facility, compared to those who never lived near a facility (Table 3). Odds remained elevated for chemical facilities for both 5 km (SIC 28, OR = 1.15, 95% CI: 0.86 – 1.54) and 2 km (OR = 1.20, 95% CI 0.84 – 1.73) metrics. Odds were significantly lower for individuals that ever versus never lived within 5 km of lumber and wood products (SIC 24, OR = 0.64, 95% CI: 0.43 – 0.94). Additionally, odds were significantly lower for residents that had ever lived within 2 km of transportation equipment facilities (SIC 37, OR = 0.54, 95% CI: 0.31 – 0.96), as compared to those that had never lived within 2 km of those facilities.

There did not appear to be any association with cumulative IDW facility count and risk of thyroid cancer (OR = 1.00, Table 4). No significant changes were seen when log-transforming the cumulative exposure metric (OR = 0.98). No trends were observed when the cumulative IDW facility count metric was categorized into exposure quartiles. Nonsignificant elevated odds were observed in the third quartile during unadjusted analyses (OR = 1.20, 95% CI: 0.82 – 1.74), but this observation was diminished after adjustment (OR = 1.05). Also, after adjusting, individuals in the highest exposure quartile had a nonsignificant decreased risk of thyroid cancer (OR = 0.76, 95% CI: 0.51 – 1.15).

During explorations of the SIC-specific cumulative IDW exposure metric, no significant differences from the binary metrics were observed (data not shown). Additionally, analyses exploring the curtailed study period (1990-1999) to test the effects of a longer latency period did not reveal any significantly altered results (data not shown).

Discussion

In these exploratory analyses, we observed no significant associations between proximity to industrial facilities and increased risk of thyroid cancer. The majority of this study population had ever lived within 2 km of any manufacturing facility through the course of this study. Both cases and controls were most likely to have ever lived within 5 km or 2 km of a fabricated metal products facility (SIC 34). Both cases and controls were also highly likely to have lived within 5 km of electronic and other electrical equipment facilities (SIC 36), rubber and miscellaneous plastics facilities (SIC 30), and industrial and commercial machinery and computer equipment facilities (SIC 35). Elevated odds of thyroid cancer were observed for individuals living within either 5 km or 2 km of chemical facilities in both unadjusted and adjusted analyses. Exposure to a variety of manufacturing sectors appeared to confer a decreased risk of thyroid cancer, though only risk estimates for exposure to lumber and wood product facilities (SIC 24) appeared to remain consistently decreased across different exposure metrics. A ten-calendar-year latency period did not substantially change risk estimates across any metrics.

Individuals that had ever lived within 5 km of lumber and wood product facilities (SIC 24) had a statistically significant decreased risk of thyroid cancer after adjustment, a relationship which persisted in both the log-transformed cumulative IDW and longer latency analyses. Lumber and wood products manufacturing facilities were thought to contribute to thyroid cancer risk in *a priori* hypotheses due to their chemical emissions. A 2009 study by De Roos et al. demonstrated an increased risk of non-Hodgkin's lymphoma (NHL) for individuals within the closest proximity those facilities (SIC 24, ≤ 0.5 mile, OR = 2.22, 95% CI: 0.4 - 11.8).[13] De Roos also found significantly elevated risks for NHL in individuals living within close proximity to chemical (SIC 28), petroleum (SIC 29), plastics (SIC 30), and primary metal industries (SIC 33). In the present study, risk estimates for chemical facilities were consistently elevated across binary metrics, but none of the other manufacturing sectors mentioned above had consistently elevated risk estimates.

Research is still emerging on how environmental exposures may impact thyroid cancer rates. Studies of the population exposed to radioactive iodine associated with the post-Chernobyl fallout have demonstrated marked increases in childhood thyroid cancers in the area.[7,21] Similarly, diagnostic radiation has been recently implicated as risk factor, observed in the same cohort used in this study.[5] However, a geospatial study that investigated radon levels in Pennsylvania counties found no significant association between cumulative radon levels and thyroid cancer incidence.[22] This may be attributable to the coarse resolution they were using—radon levels can vary dramatically across a single county based on the underlying geology, which may not be an effective predictor of risk. Another recent study failed to demonstrate an association between serum PBDE levels and thyroid cancer, despite the established observations between PBDEs and thyroid activity.[8] PBDEs are one of the major PHAHs residents were assumed to be exposed to through close proximity to manufacturing facilities; a finding of no association between PBDE exposures and thyroid cancer may explain some of the null results observed in this study.

Strengths and Limitations

This study demonstrated a novel use of a unique publicly available data resource from the Connecticut Department of Energy and Environmental Protection. With this statewide database, we were able to investigate exposures to multiple industrial sources simultaneously. Further, the collection of residential histories allowed us to track our population over time, improving our exposure estimates. The specific population used in the current study uses only individuals with recall to the town level or better for the entire study period, which aimed to limit exposure misclassification.

It is important to qualify the observed results with the inherent limitations in this study. There is a certain amount of uncertainty built into the exposure metrics. Due to the relatively imprecise recall resolution of residential address (Table 2) as compared to similar studies,[13,16] it was necessary to use town, zip code, and county centroids to assign coordinates for all study participants for every year. This potentially added exposure misclassification into the study, biasing the results toward the null. This study did not incorporate any meteorological effects into its exposure matrix creation, and therefore assumed an even distribution of pollution in a circular buffer around the facility. Additionally, actual exposure quantities were not considered, since none of the reported emission quantities in the CT Point Source Inventory reflected chemicals of relevance.

Multiplicity must be taken into account in the interpretation of this data as well. Though the objective of this study was intended to be hypothesis-generating in nature, it is important to consider that, due to the number of comparisons made, the significantly decreased odds seen in individuals that lived in close proximity to lumber and wood facilities (SIC 24) could have occurred due to random chance alone, considering the choice for $\alpha = 0.05$.

Future research could refine exposure metrics, through the incorporation of emissions data, either directly, or through a proxy emission. Exposure metrics could also be refined by improving resident address information, potentially through the linkage to existing databases that may contain historical resident information, such as a real estate database. It may also be beneficial to consider a different range of SICs. Previous research has shown that landfills and waste incinerators may be additional potential sources of PHAHs.[11-12,23] Applying the methodology explained here to include those types of facilities may help explain how alternative environmental sources of these pollutants impact thyroid cancer risk.

Conclusion

This research explored the impact residential proximity to industrial facilities has on the risk of thyroid cancer. We adapted existing methodologies and applied them to a new source of publically available data in Connecticut. Using inverse distance between geocoded residence and geocoded facility coordinates, we constructed a variety of exposure metrics, a novel approach to determine the impact of environmental exposures on thyroid cancer. Observed associations were generally null, though elevated risk estimates were seen in individuals who had ever lived within

5 or 2 km of chemical product manufacturing facilities, and significantly decreased risk estimates were consistently observed across metrics for individuals living in close proximity to lumber or wood product facilities. Our framework could be applied in future studies, which should aim to refine the exposure metrics used by improving resident coordinate accuracy, including emissions data, and incorporating additional pollution sources.

Acknowledgements and funding sources:

This study was funded through the Yale Cancer Center & American Cancer Society Institutional Research Grant IRG-58-012-57. Tremendous thanks to Nicole Deziel (Yale School of Public Health, New Haven, CT) for her advice and guidance through this process in her roles as thesis advisor and first reader. Additional thanks to Yawei Zhang (Yale School of Public Health, New Haven, CT) for her advice on exposure metrics, and her roles second reader and as primary investigator of the case-control study used here. We also thank all individuals associated with the data collection in the case-control study, especially Huang Huang (Yale School of Public Health) for her assistance with the variable coding and distribution, Lisa A McKay (Yale Center for Perinatal, Pediatric and Environmental Epidemiology, New Haven, CT) for her assistance in developing the GIS address locator used in this study, and Chris Mulcahy (Connecticut Department of Energy and Environmental Protection, Hartford, CT) for his assistance with providing and interpreting the state data.

References

1 Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2012/, based on November 2014 SEER data submission, posted to the SEER web site, April 2015.
2 Brown AP, Chen J, Hitchcock YJ, Szabo A, Shrieve DC, Tward JD. The risk of second primary malignancies up to three decades after the treatment of differentiated thyroid cancer. *J Clin Endocrinol Metab*. 2008; 93:504-515.
3 Husson O, Haak HR, Buffart LM et al. Health-related quality of life and disease and disease specific symptoms in long-term thyroid cancer survivors: A study from the population-based PROFILES registry. *Acta Oncologica*. 2013; 52(2):249-258.
4 Brent GA. Environmental Exposures and Autoimmune Thyroid Disease. *Thyroid*. 2010. 20(7):755-761.
5 Zhang Y, Chen Y, Huang H, Sandler J, Dai M, Ma S, Udelsman R. Diagnostic radiography exposure increases the risk for thyroid microcarcinoma: a population-based case-control study. *European Journal of Cancer Prevention*. 2015. 24:439-446.
6 Marcello MA, Malandrino P, Almeida JFM, Martins MB, Cunha LL, ZBufalo Ne, Pellegriti G, and Ward LS. The influence of the environment on the development of thyroid tumors: a new appraisal. *Endocrine-Related Cancer*. 2014. 21: T235-T254.
7 Sinnott B, Ron E, Schneider A. Exposing the thyroid to radiation: A review of its current extent, risks, and implications. *Endocrine Reviews*. 2010; 31(5):756-773.
8 Aschebrook-Kilfoy B, Shu XO, Gao YT, Ji BT, Yang G, Li HL, Rothman N, Chow WH, Zheng W, Ward MH. Thyroid cancer risk and dietary nitrate and nitrite intake in the Shanghai Women's Health Study. *Int J Cancer*. 2013. 132(4): 897-904.

---

9 Grimalt JO, Sunyer J, Moreno V, Amaral OC, Sala M, Rosell A, Anto JM, Albaiges J. Risk excess of soft-tissue sarcoma and thyroid cancer in a community exposed to airborne organochlorinated compound mixtures with a high hexachlorobenzene content. *Int J Cancer*. 1994. 56: 200-203.

10 Dodson RE, Perovich LJ, Covaci A, Van den Eede N, Ionas AC, Dirtu AC, Brody JG, Rudel RA. After the PBDE phase-out: A broad suite of flame retardants in repeat house dust samples from California. *Environ Sci Technol*. 2012; 46:13056-13066.

11 DellaValle CT, Wheeler DC, Deziel NC, De Roos AJ, Cerhan JR, Cozen W, Severson RK, Flory AR, Locke SJ, Colt JS, Hartge P, Ward MH. Environmental determinants of polychlorinated biphenyl concentrations in residual carpet dust. *Environ Sci Technol.* 2013; 47(18):10405-10414.

12 Deziel NC, Nuckols JR, Colt JS, De Roos AJ, Pronk A, Gourley C, Severson RK, Cozen W, Cerhan JR, Hartge P, Ward MH. Determinants of polychlorinated dizbenzo-*p*-dioxins and polychlorinated dibenzofurans in house dust samples from four areas of the United States. *Science of the Total Environment*. 2012; 433:516-522.

13 De Roos AJ, Davis S, Colt JS, Blair A, Airola M, Severson RK, Cozen W, Cerhan JR, Hartge P, Nuckols JR, Ward MH. Residential proximity to industrial facilities and risk of non-Hodgkin lymphoma. *Environmental Research*. 2009. 110: 70-78.

14 Langlois PH, Brender JD, Suarez L, Zhang FB, Mistry JH, Scheuerle A, Moody K. Maternal residential proximity to waste sites and industrial facilities and conotruncal heart defects in offspring. *Paediatric and Perinatal Epidemiology*. 2009; 23:321-331.

15 Brender JD, Maantay JA, Chakraborty J. Residential proximity to environmental hazards and adverse health outcomes. *Am J Public Health*. 2011; 101:S37-S52.

16 Kheifets L, Crespi CM, Hooper C, Oksuzyan S, Cockburn M, Ly T, Mezei G. Epidemiologic study of residential proximity to transmission lines and childhood cancer in California: description of design, epidemiologic methods and study population. *Journal of Exposure Science and Environmental Epidemiology*. 2015; 25:43-52.

17 Goldberg DW, Cockburn MG. Improving geocode accuracy with candidate selection criteria. *Transactions in GIS*. 2010; 14(s1):149-176.

18 Connecticut Department of Energy and Environmental Protection. Emissions Inventory. <http://www.ct.gov/deep/cwp/view.asp?Q=432056> 2016. (accessed November 13, 2015).

19 Connecticut Department of Energy and Environmental Protection. 2011 Periodic Emissions Inventory. <http://www.ct.gov/deep/lib/deep/air/inven/pei2011final.pdf> 2016. (Accessed Feb 1, 2016).

20 Department of Labor, Occupational Safety and Health Administration (OSHA). <https://www.osha.gov/pls/imis/sic_manual.html> 2008. (Accessed Feb 1, 2016).

21 Pacini F, Vorontsova T, Demidchik EP et al. Post-Chernobyl thyroid carcinoma in Belarus children and adolescents: Comparison with naturally occuring thyroid carcinoma in Italy and France. *Journal of Clinical Endocrinology and Metabolism*. 1997; 82(11):3563-3569.

22 Goyal N, Camacho F, Mangano J, Goldenberg D. Evaluating for a geospatial relationship between radon levels and thyroid cancer in Pennsylvania. *The Laryngoscope*. 2015; 125:E45-E49.

23 Frederiksen M, Vorkamp K, Thomsen M, Knudsen LE. Human internal and external exposure to PBDEs—A review of levels and sources. *Int J Hyg Environ Health*. 2009; 212:109-134.

Table 1. Selected characteristics of cases and controls in the Connecticut Thyroid Cancer Case-Control study (Zhang et al, 2015) and in the current study of residential proximity to industrial facilities. Frequency (percentage), except where noted.

Characteristic	Connecticut Thyroid Cancer Case-Control study			Study of residential proximity to industrial facilities[a]		
	Cases (n = 462)	Controls (n = 498)	p^b	Cases (n = 408)	Controls (n = 436)	p^b
Gender						
Female	375 (81.2)	344 (69.1)	*<0.0001*	332 (81.4)	295 (67.7)	*<0.0001*
Male	87 (18.8)	154 (30.9)		76 (18.6)	141 (32.3)	
Age (years)						
Mean Age (standard dev)	51.22 (12.3)	54.15 (13.1)	*0.0004*	52.04 (12.2)	55.1 (13.0)	*0.0005*
<40	86 (18.6)	64 (12.9)	*0.0017*	66 (16.2)	54 (12.4)	*0.0027*
40-49	115 (24.9)	123 (24.7)		101 (24.8)	91 (20.9)	
50-59	149 (32.3)	139 (27.9)		137 (33.6)	128 (29.4)	
≥60	112 (24.2)	172 (34.5)		104 (25.5)	163 (37.4)	
Education						
No High School	5 (1.1)	4 (0.8)	*0.0102*	5 (1.2)	3 (0.7)	*0.0089*
Any high school	124 (26.8)	84 (16.9)		117 (28.7)	81 (18.6)	
Trade School/College	216 (46.8)	261 (52.4)		190 (46.6)	228 (52.3)	
Graduate School	100 (21.6)	130 (26.1)		85 (20.8)	113 (25.9)	
Other	13 (2.8)	13 (2.6)		11 (2.7)	11 (2.5)	
Missing	4 (0.9)	6 (1.2)		0 (0.0)	0 (0.0)	
Race						
White	415 (89.8)	450 (90.4)	*0.3259*	374 (91.7)	398 (91.3)	*0.6560*
Black	18 (3.9)	25 (5.0)		17 (4.2)	23 (5.3)	
Other	29 (6.3)	22 (4.4)		17 (4.2)	15 (3.4)	
Family history of cancer						
Thyroid Cancer	74 (16.0)	48 (9.6)	*0.0111*	66 (16.2)	37 (8.5)	*0.0019*
Other Cancer	245 (53.0)	291 (58.4)		215 (52.7)	264 (60.6)	
No	143 (31.0)	159 (31.9)		127 (31.1)	135 (31.0)	
Family History of Disease						
Thyroid Disease	115 (24.9)	76 (15.3)	*<0.0001*	95 (23.3)	64 (14.7)	*0.0006*
Other Disease	7 (1.5)	1 (0.2)		6 (1.5)	1 (0.2)	
No	340 (73.6)	421 (84.5)		307 (75.2)	371 (85.1)	
Body mass index (kg/m^2)						
<25	145 (31.4)	203 (40.8)	*0.0003*	126 (30.9)	173 (39.7)	*0.0002*
25.0-29.9	146 (31.6)	168 (33.7)		124 (30.4)	151 (34.6)	
≥30	166 (35.9)	118 (23.7)		157 (38.5)	108 (24.8)	
Missing	5 (1.1)	9 (1.8)		1 (0.2)	4 (0.9)	
Smoking						
Yes	141 (30.5)	172 (34.5)	*0.1844*	132 (32.4)	156 (35.8)	*0.2940*
No	321 (69.5)	326 (65.5)		276 (67.6)	280 (64.2)	
Alcohol consumption						
Yes	188 (40.7)	267 (53.6)	*<0.001*	173 (42.4)	237 (54.4)	*0.0005*
No	274 (59.3)	231 (46.4)		235 (57.6)	199 (45.6)	
Any Diagnostic Radiation						
Yes	423 (91.6)	440 (88.4)	*0.0997*	377 (92.4)	391 (89.7)	*0.1673*
No	39 (8.4)	58 (11.6)		31 (7.6)	45 (10.3)	

[a] Analysis included participants with 100% of reported address information from 1990-2009 reported to the city resolution or finer.

[b] p values generated from χ^2-test or t-test (age).

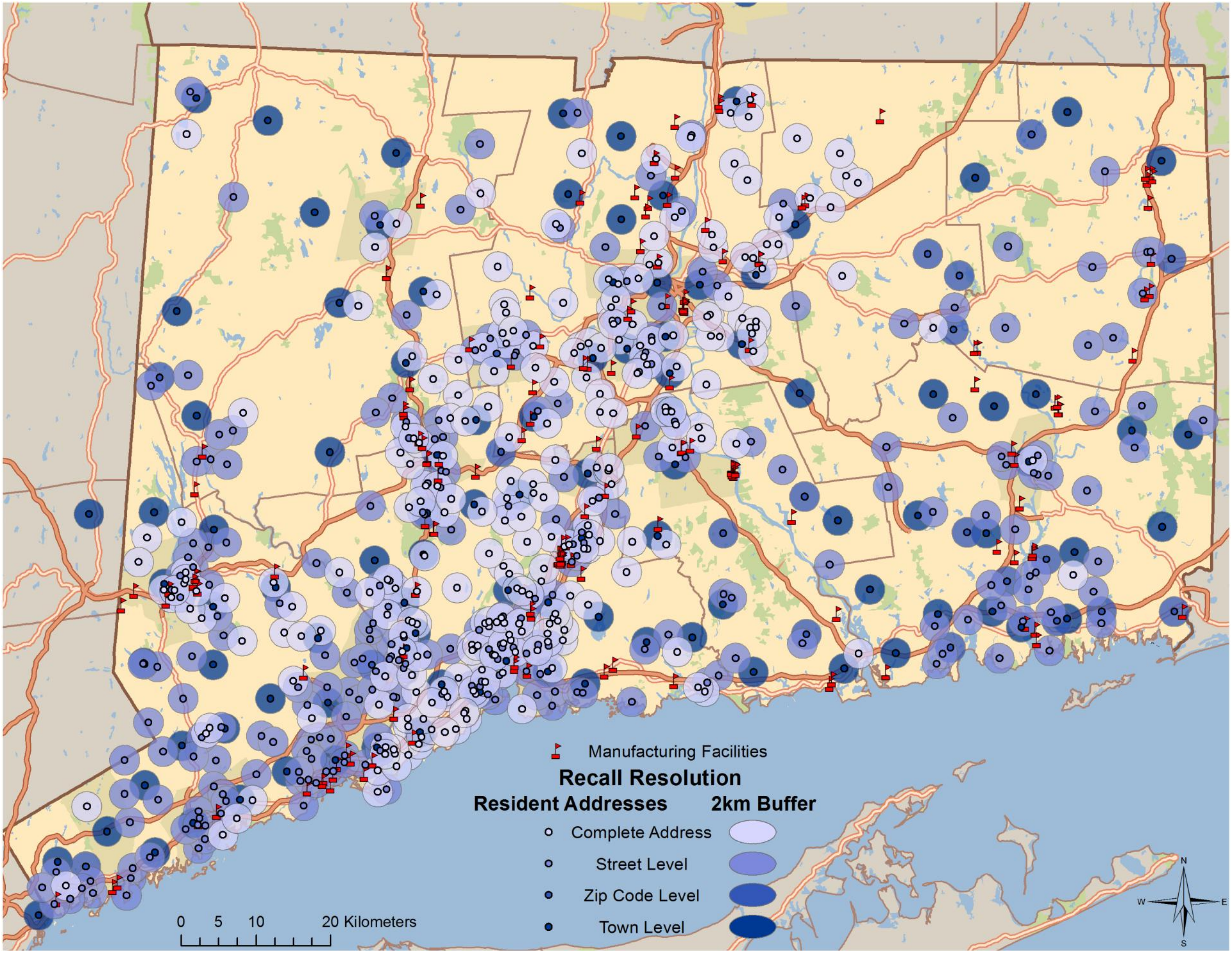

Figure 1. An example of the exposure metric used in this study. Resident addresses were geocoded to the greatest resolution possible with the available information. In this example, an exposure would be considered for all individuals who had a manufacturing facility within 2 km of their residence in the year 2000.

Table 2. Accuracy levels for geocoding of residential addresses reported in the study population (100% of reported address geocoded to city resolution or better between the years 1990 and 2009. Frequency (percentage).

Geographic Resolution	Case (n = 974)	Control (n = 981)
Parcel	220 (22.6)	226 (23.0)
Street Name/Segment	282 (29.0)	251 (25.6)
Zip Code	35 (3.6)	37 (3.8)
City	437 (44.9)	467 (47.6)

p [a] = 0.4057

[a] p value generated from χ^2-test.

Table 3. Associations of thyroid cancer with residential proximity to industrial facilities from 1990-2009, using binary metrics. Significance at $p < 0.05$ shown in bold.[a]

	Ever Lived Within 5 km of a Facility				Ever Lived Within 2 km of a Facility			
Industry Sector: SIC	Case (n = 408)	Control = (436)	p[b]	OR (95% CI)[c]	Case (n = 408)	Control = (436)	p[b]	OR (95% CI)[c]
Any Manufacturing SIC	358 (87.7)	383 (87.8)	0.97	0.87 (0.56-1.34)	236 (57.8)	257 (58.9)	0.75	0.77 (0.58-1.04)
Food and Kindred Products:20	122 (29.9)	132 (30.3)	0.91	0.88 (0.65-1.21)	26 (6.4)	28 (6.4)	0.98	0.84 (0.47-1.53)
Textile Mill Products: 22	104 (25.5)	93 (21.3)	0.15	1.07 (0.76-1.50)	25 (6.1)	18 (4.1)	0.19	1.05 (0.55-2.02)
Apparel and other Finished Products: 23	45 (11.0)	50 (11.5)	0.84	0.84 (0.54-1.33)	16 (3.9)	16 (3.7)	0.85	1.10 (0.52-2.34)
Lumber and Wood Products, except Furniture: 24	59 (14.5)	78 (17.9)	0.18	**0.64 (0.43-0.94)**	14 (3.4)	21 (4.8)	0.31	0.57 (0.27-1.18)
Furniture and Fixtures: 25	61 (15.0)	72 (16.5)	0.53	0.87 (0.59-1.29)	15 (3.7)	19 (4.4)	0.61	0.74 (0.36-1.53)
Paper and Allied Products: 26	142 (34.8)	156 (35.8)	0.77	0.86 (0.64-1.17)	29 (7.1)	26 (6.0)	0.50	1.17 (0.65-2.10)
Printing, Publishing, and Allied Industries: 27	194 (47.5)	184 (42.2)	0.12	1.09 (0.82-1.45)	33 (8.1)	44 (10.1)	0.31	0.63 (0.38-1.04)
Chemicals and Allied Products: 28	191 (46.8)	185 (42.4)	0.20	1.15 (0.86-1.54)	89 (21.8)	73 (16.7)	0.06	1.20 (0.84-1.73)
Petroleum Refining and Related Industries: 29	181 (44.4)	193 (44.3)	0.98	0.89 (0.67-1.19)	41 (10.0)	44 (10.1)	0.98	0.84 (0.52-1.36)
Rubber and Miscellaneous Plastics Products: 30	206 (50.5)	214 (49.1)	0.68	0.92 (0.69-1.23)	76 (18.6)	74 (17.0)	0.53	0.96 (0.66-1.39)
Stone, Clay, Glass, and Concrete Products: 32	171 (41.9)	182 (41.7)	0.96	0.89 (0.66-1.18)	42 (10.3)	49 (11.2)	0.66	0.73 (0.46-1.16)
Primary Metal Industries: 33	194 (47.5)	188 (43.1)	0.20	1.06 (0.80-1.42)	76 (18.6)	79 (18.1)	0.85	0.78 (0.53-1.13)
Fabricated Metal Products, except Machinery and Transportation Equipment:34	265 (65.0)	274 (62.8)	0.52	1.03 (0.76-1.38)	137 (33.6)	138 (31.7)	0.55	0.93 (0.69-1.27)
Industrial and Commercial Machinery and Computer Equipment: 35	203 (49.8)	196 (45.0)	0.16	1.11 (0.84-1.48)	63 (15.4)	78 (17.9)	0.34	0.69 (0.47-1.02)
Electronic and other Electrical Equipment and Components, except Computer Equipment: 36	215 (52.7)	212 (48.6)	0.24	1.09 (0.82-1.45)	58 (14.2)	65 (14.9)	0.78	0.78 (0.52-1.17)
Transportation Equipment: 37	167 (40.9)	190 (43.6)	0.44	0.82 (0.61-1.10)	23 (5.6)	38 (8.7)	0.08	**0.54 (0.31-0.96)**
Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks: 38	143 (35.0)	152 (34.9)	0.95	0.97 (0.72-1.30)	39 (9.6)	41 (9.4)	0.94	0.96 (0.60-1.55)

[a] Analysis included participants with 100% of reported address information from 1990-2009 reported to the city resolution or finer.

[b] p values generated from χ^2-test.

[c] Odds Ratio and 95% Confidence intervals generated from unconditional logistic regression models adjusting for age, BMI, gender, prior alcohol use, family history of thyroid disease, family history of thyroid cancer, and previous exposure to diagnostic radiation.

Table 4. Associations of thyroid cancer with residential proximity to industrial facilities from 1990-2009, using continuous metrics.[a]

Cumulative IDW Quartile[b]	Cumulative IDW Range	Unadjusted OR (95% CI)	Adjusted OR (95% CI)[c]
1	0 - 9.26	1.00	1.00
2	9.31 - 45.20	0.89 (0.60-1.32)	0.81 (0.54-1.22)
3	42.21 - 138.36	1.20 (0.82-1.74)	1.05 (0.71-1.56)
4	140.57 - 1398.77	0.95 (0.65-1.40)	0.76 (0.51-1.15)

	OR (95% CI)	OR (95% CI)[c]
Cumulative IDW Exposure[d]	1.00 (1.00-1.00)	1.00 (1.00-1.00)
ln(Cumulative IDW Exposure)	1.00 (0.96-1.05)	0.98 (0.89-1.08)

[a] Analysis included participants with 100% of reported address information from 1990-2009 reported to the city resolution or finer.

[b] Inverse Distance Weighted Quartiles generated from control exposure data.

[c] Odds Ratio and 95% Confidence intervals generated from unconditional logistic regression models adjusting for age, BMI, gender, prior alcohol use, family history of thyroid disease, family history of thyroid cancer, and previous exposure to diagnostic radiation.

[d] See Equation 1.