Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Serum thyroid hormone concentrations and breast cancer risk: an extended cohort study with time-varying exposure analysis

Chang Ik Yoon
The Catholic University of Korea

Hyung Suk Kim
Hanyang University Guri Hospital, Hanyang University College of Medicine

Soong June Bae
Yonsei University College of Medicine

Jang-Hee Lee
Ewha Womans University Mokdong Hospital, Ewha Womans University School of Medicine

Eun Young Kim
Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine

Chan Heun Park
Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine

Yong Lai Park
Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine

Ji-Sup Yun
Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine

Yun Tae Kim
Kangbuk Samsung Hospital

Kwan Ho Lee
kwanho.lee@samsung.com

Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine

---

Research Article

Keywords: breast cancer, thyroid-stimulating hormone, free thyroxine, cohort study, time-varying Cox regression, cancer epidemiology

Posted Date: October 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10103031/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

Single-baseline categorization of biochemical exposures that fluctuate within individuals can misclassify cumulative exposure, yet most cohort studies of thyroid function and breast cancer rely on a single measurement and self-reported outcomes. We examined serum thyroid-stimulating hormone (TSH) and free thyroxine (FT4) in relation to breast cancer risk in a large Korean cohort, comparing static baseline with pre-specified time-varying exposure modeling, and tested whether previously reported associations from this cohort were reproduced.

Methods

In 162,702 women aged ≥ 20 years in the Kangbuk Samsung Health Study (2011–2024), incident breast cancer (ICD-10 C50 or D05) was ascertained by Korea Central Cancer Registry linkage. We fit progressively adjusted static Cox models and time-varying Cox models that updated TSH and FT4 at repeated visits, with pre-specified subgroup (menopausal status, body mass index) and sensitivity analyses.

Results

Over a mean 6.94 years, 1,841 incident breast cancers occurred. For euthyroid-range TSH, static and time-varying models gave concordant inverse estimates (highest vs lowest tertile: fully adjusted hazard ratio [HR] 0.93, 95% confidence interval [CI] 0.81–1.07; time-varying 0.90, 0.80–1.02). The inverse trend was stronger in postmenopausal women (HR 0.69, 0.47–1.02) and for invasive-only disease (HR 0.88, 0.76–1.01). A previously reported positive elevated-FT4 association was not reproduced and was unstable across static and time-varying models (0.62 vs 1.05; 12 events).

Conclusion

Euthyroid TSH showed a consistent but attenuated inverse association with breast cancer risk, whereas a prior elevated-FT4 signal was not replicated. Repeated-measurement designs matter most for episodic biochemical exposures; subtype-stratified replication is needed to test the postmenopausal TSH signal.

1. Introduction

Breast cancer is the most commonly diagnosed cancer worldwide and the leading cause of cancer mortality in women, with incidence rising in East Asian populations including South Korea [1, 2]. Although established risk factors—age, reproductive history, family history, body mass index, and exogenous hormone exposure—account for a substantial fraction of inter-individual variation in risk [3],

additional modifiable hormonal markers that could refine risk stratification remain under active investigation.

Thyroid hormones have a long-standing biological rationale as candidates in mammary carcinogenesis. Triiodothyronine binds the thyroid hormone receptor on breast epithelial cells, promoting proliferation [4], and acts as a partial estrogen-receptor agonist through downstream nuclear estrogen-receptor phosphorylation [5]; thyroid-stimulating hormone (TSH), by contrast, has been variably reported to exert direct or indirect effects on breast tissue through downstream changes in circulating thyroid hormone levels [6]. Epidemiological studies have produced inconsistent findings: some have observed positive associations between elevated free thyroxine (FT4) or hyperthyroidism and breast cancer risk, particularly in postmenopausal women [7, 8]; others have reported inverse associations with TSH within the euthyroid range [9, 10]; and still others have found no meaningful association [11, 12]. Heterogeneity in study populations, exposure ascertainment, follow-up duration, and outcome verification likely contributes to these discrepancies [6].

In a previous report from the Kangbuk Samsung Health Study (KSHS) [10], we examined 62,546 Korean women aged ≥ 40 years across 2002–2018, identifying 834 incident breast cancers. We reported that the highest tertile of TSH within the euthyroid reference range was associated with a 32% lower breast cancer hazard relative to the lowest tertile (adjusted hazard ratio [HR] 0.68, 95% confidence interval [CI] 0.55–0.84), and that abnormally elevated FT4 (> 1.7 ng/dL) was associated with a near-doubling of breast cancer risk (aHR 1.98, 95% CI 1.02–3.83). Although biologically plausible and statistically robust within that sample, those findings were subject to several methodological limitations: cancer outcomes were principally ascertained through self-report; only a single baseline measurement of TSH and FT4 was used despite well-documented intra-individual variability over time; the analytic age range excluded women under 40 years of age, limiting generalizability to younger Korean women; and follow-up extended only through 2018.

To address these limitations and to test whether the previously reported associations are replicated in a substantially larger and more thoroughly characterized cohort, we conducted the present analysis with five pre-specified extensions: (i) the present analytic cohort comprises KSHS participants undergoing health screening between January 2011 and December 2024; this recruitment window overlaps that of our prior report (2002–2018) during 2011–2018, so that a subset of participants is shared between the two analyses (see § 2.1); (ii) breast cancer ascertainment was performed by direct linkage to the Korea Central Cancer Registry (KCCR), eliminating reliance on self-report; (iii) the eligible age range was expanded to all women aged 20 years or older to capture earlier-life thyroid function and breast cancer risk relations; (iv) time-varying Cox proportional hazards modeling using repeated KSHS measurements of TSH and FT4 was added as a pre-specified analytic strategy to mitigate exposure misclassification due to single-baseline measurement; and (v) the primary outcome was defined as incident invasive breast cancer or ductal carcinoma in situ (ICD-10 C50 or D05), with restriction to invasive disease (C50 only) pre-specified as a sensitivity analysis. We hypothesized that the inverse association between euthyroid TSH and breast cancer risk and the positive association with elevated FT4 would be replicated,

and that time-varying analyses would refine static effect estimates by addressing exposure misclassification due to single-baseline measurement.

2. Materials and Methods

2.1 Study design and population

This retrospective cohort study used data from the Kangbuk Samsung Health Study (KSHS), an ongoing cohort of adults undergoing comprehensive health screening at the Kangbuk Samsung Hospital Total Healthcare Centers in Seoul and Suwon, South Korea [13]. The present analysis draws on KSHS participants enrolled between January 2011 and December 2024, extending the temporal coverage of our prior KSHS report [10] (which spanned 2002–2018). Because the two recruitment windows overlap during 2011–2018 and draw on the same source population, a subset of participants is shared between the two analyses; the present study is therefore best regarded as an extension and partial re-analysis of the prior cohort rather than a fully independent replication. The broadened eligibility (age ≥ 20 vs ≥ 40 years), registry-linked outcome ascertainment, and time-varying exposure modeling nonetheless distinguish it substantially from the earlier report.

Eligible participants were women aged 20 years or older with paired serum TSH and FT4 measurements obtained between January 1, 2011, and December 31, 2024. Exclusion criteria, applied sequentially, were: any prior history of cancer at baseline; current use of thyroid medications (levothyroxine, antithyroid agents, or iodine preparations); follow-up duration shorter than 6 months; and missing TSH or FT4 at baseline. Cohort entry was defined as the first eligible visit. Of 192,160 women in the source population, 166,916 women remained after sequential exclusions. For the primary analyses based on the three-category thyroid function classification, an additional 4,214 women whose thyroid function could not be unambiguously classified (euthyroid TSH with abnormal FT4, or discordant TSH/FT4) were excluded, yielding an analytic cohort of 162,702 women with 1,841 incident breast cancers under the primary outcome definition. The cohort selection process is summarized in Supplementary Fig. S0 (CONSORT-style flow diagram).

2.2 Outcome ascertainment

Incident breast cancer was ascertained by direct linkage to the KCCR through December 31, 2024 [14]. Linkage was performed using national unique identifiers, providing complete incident-case ascertainment without dependence on self-report; prior history of cancer at baseline, applied as an exclusion criterion (§ 2.1), was ascertained from the KSHS self-administered baseline questionnaire and supplemented by KCCR-linkage exclusion of any participant with a registry-recorded cancer diagnosis prior to cohort entry. The pre-specified primary outcome was incident invasive breast cancer or ductal carcinoma in situ (ICD-10 C50 or D05), consistent with the inclusive breast cancer outcome definition commonly applied in Korean cohort epidemiology [15]. Restriction to invasive breast cancer alone (ICD-10 C50 only) was pre-specified as a sensitivity analysis. Because the 6-month landmark exclusion was applied with respect to the outcome under analysis, the invasive-only (C50) sensitivity cohort comprised

162,758 women, marginally larger than the primary C50 + D05 cohort (n = 162,702): 57 women whose qualifying diagnosis was ductal carcinoma in situ within 6 months of baseline were excluded under the primary definition but retained under the invasive-only definition (net difference of 56 after thyroid-function reclassification). Person-time was accrued from cohort entry to the earliest of breast cancer diagnosis, death, or December 31, 2024.

2.3 Exposure assessment

Serum TSH and FT4 were measured by electrochemiluminescent immunoassay (ECLIA) on Cobas e801 analyzers (Roche Diagnostics) using laboratory reference ranges previously validated in KSHS [13] (TSH 0.25–5.0 μIU/mL; FT4 0.93–1.7 ng/dL). Thyroid function status at baseline was classified into three primary categories: euthyroid (TSH and FT4 both within reference range); hypothyroid (TSH > 5.0 μIU/mL, encompassing both subclinical and overt forms); and hyperthyroid (TSH < 0.25 μIU/mL). A five-category classification—euthyroid, subclinical hypothyroidism, overt hypothyroidism, subclinical hyperthyroidism, and overt hyperthyroidism—was examined as a sensitivity analysis. Within euthyroid participants, TSH was further partitioned into cohort-derived tertiles (T1 ≤ 1.52, T2 1.52–2.42, T3 ≥ 2.42 μIU/mL), reproducing the analytic strategy of our prior report [10]. Elevated FT4 was defined dichotomously as FT4 > 1.7 ng/dL. Because participants with discordant TSH–FT4 patterns—including euthyroid-range TSH accompanied by FT4 > 1.7 ng/dL—were excluded from the analytic cohort (§ 2.1), within this cohort the elevated-FT4 indicator coincided with the overt-hyperthyroidism category of the five-category classification; the two therefore comprise the same participants and show matching event counts within each outcome definition.

For time-varying analyses, baseline TSH and FT4 values were updated with subsequent KSHS measurements at follow-up visits, using last-observation-carried-forward and a 6-month exposure lag (12-month lag examined as a sensitivity analysis).

2.4 Covariates

Covariates were selected to match those used in our prior report [10] to enable direct quantitative comparison. Demographic and lifestyle covariates included age (years), examination center (Seoul/Suwon), examination year, body mass index (BMI, kg/m²), smoking status (never/former/current), alcohol intake (g/day), and vigorous physical activity. Reproductive covariates included menopausal status, parity, age at menarche, and exogenous female hormone use. First-degree family history of breast cancer was also included. Examination-center distributions of TSH and FT4 are summarized in Supplementary Table S5.

2.5 Statistical analysis

Baseline characteristics were summarized by breast cancer event status, using means with standard deviations (SDs) for continuous variables and counts with percentages for categorical variables; group differences were assessed by t-test or one-way ANOVA for continuous variables and by χ^2 test for

categorical variables. Analyses were restricted to participants with complete data on the exposure and all pre-specified covariates (complete-case analysis); multiple imputation was not performed.

The primary analyses fit three nested Cox proportional hazards models. Model 1 adjusted for age, examination center, and examination year. Model 2 additionally adjusted for body mass index, menopausal status, parity, age at menarche, hormone use, and family history of breast cancer. Model 3 further adjusted for smoking status, alcohol intake, and physical activity (the fully adjusted model, reported as primary for static analyses). Hazard ratios (HRs) with 95% confidence intervals (CIs) were reported for the three-category thyroid function classification (euthyroid as reference); TSH tertiles within the euthyroid range (T1 as reference); continuous TSH and FT4 expressed per 1-SD increment within the euthyroid range; and the binary elevated-FT4 indicator (FT4 > 1.7 ng/dL; FT4 ≤ 1.7 ng/dL as reference). Restricted cubic splines (RCS) with knot placement matching our prior report [10] were used to examine non-linear dose-response relations of TSH and FT4 with breast cancer risk across both the euthyroid and full measurement ranges [16]; tests for non-linearity were based on the Wald statistic for the non-linear spline terms.

Time-varying Cox models were fit using counting-process data [17], in which TSH and FT4 exposures were updated at each subsequent KSHS visit using last-observation-carried-forward with a 6-month exposure lag to mitigate reverse causation. Static-vs-time-varying contrasts were presented under harmonized Model 2 adjustment to match the most comprehensive covariate set with complete data across repeated measurements. Pre-specified subgroup analyses examined effect modification by menopausal status (premenopausal vs postmenopausal) and body mass index (< 23, 23–24.9, ≥ 25 kg/m²) for the TSH tertile association, with multiplicative interaction tested by likelihood ratio tests.

Pre-specified sensitivity analyses included: (i) restriction of the outcome to invasive breast cancer alone (ICD-10 C50); (ii) the five-category thyroid function exposure classification, with Firth penalized Cox regression [18] applied to categories with fewer than 10 events to address sparse-data bias; (iii) extension of the time-varying exposure lag from 6 to 12 months; and (iv) inclusion of participants on thyroid medication who had been excluded a priori. The proportional hazards assumption was assessed by Schoenfeld residual tests [19, 20] for each exposure–covariate combination.

Analyses were performed in R version 4.3 with the *survival* [17] and *rms* [16] packages, and verified in Stata version 15. Two-sided *p* values < .05 were considered statistically significant. Reporting follows the STROBE guideline [21]. Artificial-intelligence–based tools were used solely to assist with English-language copy editing of the manuscript; they were not used for the collection, analysis, or interpretation of data or for the production of any figures or results.

2.6 Ethics

The study protocol was approved by the Institutional Review Board of Kangbuk Samsung Hospital (KBSMC 2026-03-023). Because the analyses used de-identified data from existing health-screening and registry records, the requirement for individual informed consent was waived. The study was conducted

in accordance with the Declaration of Helsinki and applicable Korean regulations on the protection of personal health information.

3. Results

3.1 Cohort characteristics and follow-up

Among 166,916 eligible women, 162,702 were classifiable into one of three thyroid function categories at baseline; 4,214 women with indeterminate classification—euthyroid TSH with abnormal FT4 or discordant TSH/FT4—were excluded from primary analyses. The analytic cohort contributed 1,128,401 person-years of follow-up (mean 6.94 years per participant), during which 1,841 incident breast cancers (invasive [C50]: 1,347; ductal carcinoma in situ [D05]: 494) were ascertained through direct linkage to the Korea Central Cancer Registry, yielding an overall incidence rate of 1.63 per 1,000 person-years.

Compared with non-cases, women who developed breast cancer were older at baseline (mean 40.5 vs 38.4 years, $p < .001$), more likely to be parous (88.1% vs 74.1%, $p < .001$), and more likely to report a first-degree family history of breast cancer (5.3% vs 3.0%, $p < .001$); body mass index and exogenous hormone use did not differ ($p = .733$ and .302, respectively), and although smoking status and alcohol intake reached nominal statistical significance ($p = .006$ and .013), the absolute between-group differences were small. Mean baseline TSH (2.40 vs 2.45 μIU/mL) and FT4 (1.23 vs 1.23 ng/dL) were similar between groups. In the analytic cohort, 92.2% were euthyroid, 6.3% had biochemical hypothyroidism (subclinical or overt), and 1.5% had biochemical hyperthyroidism. Baseline characteristics by event status are presented in Table 1.

Table 1
Baseline characteristics by breast cancer event status

Characteristic	No breast cancer	Breast cancer	p value
n	160,861	1,841	—
Age, years (mean ± SD)	38.42 ± 10.39	40.46 ± 7.69	< .001
Body mass index, kg/m² (mean ± SD)	21.94 ± 3.23	21.91 ± 3.07	.733
Waist circumference, cm (mean ± SD)	76.16 ± 8.64	76.47 ± 8.29	.169
TSH, μIU/mL (mean ± SD)	2.45 ± 2.86	2.40 ± 2.08	.423
Free T4, ng/dL (mean ± SD)	1.23 ± 0.23	1.23 ± 0.30	.856
Examination center, n (%)			< .001
Seoul	69,538 (43.2)	997 (54.2)	
Suwon	91,323 (56.8)	844 (45.8)	
Smoking status, n (%)			.006
Never	126,136 (89.8)	1,268 (87.4)	
Former	10,421 (7.4)	139 (9.6)	
Current	3,844 (2.7)	44 (3.0)	
Alcohol intake, g/day (mean ± SD)	4.81 ± 8.83	4.24 ± 7.16	.013
Postmenopausal, n (%)	23,846 (14.9)	208 (11.4)	< .001
Parous, n (%)	113,792 (74.1)	1,550 (88.1)	< .001
Female hormone use, n (%)	2,685 (1.7)	37 (2.0)	.302
Family history of breast cancer, n (%)	4,764 (3.0)	98 (5.3)	< .001
Thyroid function category, n (%)			.209
Euthyroid	148,296 (92.2)	1,717 (93.3)	
Hypothyroid	10,121 (6.3)	102 (5.5)	
Hyperthyroid	2,444 (1.5)	22 (1.2)	

Note: Analytic cohort: n = 162,702 women (no breast cancer 160,861; breast cancer 1,841). Primary outcome: incident invasive breast cancer or ductal carcinoma in situ (ICD-10 C50 or D05). Continuous variables compared by t-test; categorical variables by χ^2 test. Missing data handled by complete-case analysis; participants with missing data on any listed variable were excluded from the corresponding analysis. Percentages for categorical variables are computed among participants with non-missing data for each variable and therefore use variable-specific denominators rather than the column total.

3.2 Thyroid function and breast cancer risk

In fully adjusted Cox proportional hazards models (Model 3; Table 2), neither biochemical hypothyroidism (HR 0.83, 95% CI 0.64–1.07; p = .157) nor biochemical hyperthyroidism (HR 0.68, 95% CI 0.39–1.17; p = .163) was significantly associated with breast cancer risk relative to euthyroidism, although both point estimates were directionally inverse. Within the euthyroid range, the HR for the highest versus lowest TSH tertile was 0.93 (95% CI 0.81–1.07; p = .320) and continuous TSH per 1-SD increment yielded HR 0.95 (95% CI 0.80–1.12; p = .506). Continuous FT4 per 1-SD was null (HR 0.97, 95% CI 0.88–1.07; p = .588); among the 1,197 women with elevated FT4 (>1.7 ng/dL; 12 events), the HR was 0.62 (95% CI 0.28–1.39; p = .247). Restricted cubic spline analyses showed no evidence of non-linear dose-response for TSH or FT4 across either the euthyroid or full measurement ranges (all p values for non-linearity > .30; Supplementary Fig. S1).

Table 2 Hazard ratios for breast cancer by thyroid function exposure, by progressive adjustment (Models 1–3)

Exposure	Category	N	Events	IR / 1,000 PY	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)
A. Thyroid function (3-cat)	Euthyroid	150,013	1,717	1.65	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
	Hypothyroid	10,223	102	1.46	0.85 (0.66–1.09)	0.86 (0.69–1.06)	0.83 (0.64–1.07)
	Hyperthyroid	2,466	22	1.21	0.71 (0.43–1.17)	0.69 (0.44–1.09)	0.68 (0.39–1.17)
B. TSH tertile (Euthyroid)	T1 (≤1.52)	50,506	580	1.66	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
	T2 (1.52–2.42)	49,888	573	1.65	0.96 (0.85–1.09)	0.96 (0.85–1.09)	0.96 (0.83–1.10)
	T3 (≥2.42)	49,619	564	1.64	0.95 (0.84–1.07)	0.95 (0.84–1.07)	0.93 (0.81–1.07)
C. Continuous (Euthyroid)	TSH per 1 SD	150,013	1,717	—	0.96 (0.83–1.13)	0.95 (0.81–1.12)	0.95 (0.80–1.12)
	FT4 per 1 SD	150,013	1,717	—	0.97 (0.88–1.07)	0.97 (0.88–1.07)	0.97 (0.88–1.07)
D. Elevated FT4	FT4 ≤ 1.7 ng/dL	161,505	1,829	1.63	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
	FT4 > 1.7 ng/dL	1,197	12	1.33	0.63 (0.28–1.40)	0.62 (0.28–1.39)	0.62 (0.28–1.39)

Note: Model 1: adjusted for age, examination center, and examination year. Model 2: Model 1 + body mass index, menopausal status, parity, age at menarche, exogenous female hormone use, and family history of breast cancer. Model 3: Model 2 + smoking status, alcohol intake, and vigorous physical activity (fully adjusted; reported as primary in the main text). N indicates the number of participants in each stratum at baseline.

Abbreviations: CI, confidence interval; FT4, free thyroxine; HR, hazard ratio; IR, incidence rate; PY, person-years; SD, standard deviation; TSH, thyroid-stimulating hormone.

Time-varying Cox models, in which TSH and FT4 categories were updated at each repeated KSHS visit with a 6-month exposure lag, yielded estimates that diverged from static baseline models in informative

ways (Table 3; harmonized Model 2 adjustment). For the highest versus lowest TSH tertile within the euthyroid range, the static Model 2 HR of 0.95 became 0.90 (95% CI 0.80–1.02; p = .096) under time-varying modeling, and continuous TSH per 1-SD became 0.89 (95% CI 0.77–1.02; p = .094)—shifts consistent with mild attenuation of the static estimate toward the null. In contrast, the suppressed-TSH and elevated-FT4 categories showed directional change under time-varying modeling: the static HR for hyperthyroidism of 0.69 became 1.04 (95% CI 0.72–1.50; p = .840), and the static HR for FT4 > 1.7 ng/dL of 0.62 became 1.05 (95% CI 0.62–1.77; p = .863). These patterns, although limited by small numbers of events in the relevant strata, are consistent with single-baseline elevated-FT4 or suppressed-TSH status reflecting transient measurement events rather than sustained exposure (see Discussion §4.2).

Table 3 Static baseline versus time-varying Cox models (harmonized Model 2 adjustment)

	Static HR (95% CI)	p value	Time-varying HR (95% CI)	p value
A. Thyroid function (3-cat)				
Euthyroid	1.00 (Ref)		1.00 (Ref)	
Hypothyroid	0.86 (0.69–1.06)	.147	0.96 (0.78–1.17)	.653
Hyperthyroid	0.69 (0.44–1.09)	.111	1.04 (0.72–1.50)	.840
B. TSH tertile (Euthyroid)				
T1 (≤1.52)	1.00 (Ref)		1.00 (Ref)	
T2 (1.52–2.42)	0.96 (0.85–1.09)	.542	0.99 (0.87–1.11)	.825
T3 (≥2.42)	0.95 (0.84–1.07)	.403	0.90 (0.80–1.02)	.096
C. Continuous (Euthyroid)				
TSH per 1 SD	0.95 (0.80–1.12)	.506	0.89 (0.77–1.02)	.094
FT4 per 1 SD	0.97 (0.88–1.07)	.588	1.01 (0.93–1.10)	.743
D. Elevated FT4				
FT4 > 1.7 vs ≤ 1.7 ng/dL	0.62 (0.28–1.39)	.247	1.05 (0.62–1.77)	.863

Note: Both columns use Model 2 covariate adjustment (age, examination center, examination year, body mass index, menopausal status, parity, age at menarche, hormone use, family history of breast cancer). Time-varying Cox models updated TSH and FT4 exposure at each KSHS follow-up visit using last-observation-carried-forward with a 6-month exposure lag and a counting-process data structure.

Abbreviations: CI, confidence interval; FT4, free thyroxine; HR, hazard ratio; SD, standard deviation; TSH, thyroid-stimulating hormone.

3.3 Subgroup analyses by menopausal status and body mass index

In pre-specified subgroup analyses (Table 4), the TSH tertile association was confined to postmenopausal women. The HR for the highest versus lowest TSH tertile was 0.98 (95% CI 0.86–1.11; p = .739) among premenopausal women and 0.69 (95% CI 0.47–1.02; p = .061) among postmenopausal women; the multiplicative interaction yielded a likelihood ratio p value of .086 (T3 × postmenopausal). No effect modification by BMI was observed (all interaction p > .40).

Table 4 Subgroup analyses of the TSH tertile association by menopausal status and body mass index

Stratum	TSH tertile	N	Events	HR (95% CI)	p	p interaction
Premenopausal	T1 (≤1.52)	43,456	510	1.00 (Ref)		
	T2 (1.52–2.42)	42,461	506	0.98 (0.87–1.12)	.803	
	T3 (≥2.42)	41,394	497	0.98 (0.86–1.11)	.739	
Postmenopausal	T1 (≤1.52)	6,718	67	1.00 (Ref)		
	T2 (1.52–2.42)	7,023	60	0.79 (0.54–1.14)	.207	.266
	T3 (≥2.42)	7,813	63	0.69 (0.47–1.02)	.061	.086
BMI <23 kg/m²	T3 (≥2.42)	33,724	397	0.96 (0.84–1.11)	.603	
BMI 23–24.9 kg/m²	T3 (≥2.42)	7,714	83	0.94 (0.68–1.29)	.690	.867
BMI ≥25 kg/m²	T3 (≥2.42)	7,984	81	0.90 (0.65–1.25)	.539	.708

Note: Subgroup analyses adjusted for age, examination center, examination year, body mass index (where applicable), menopausal status (where applicable), parity, age at menarche, hormone use, and family history of breast cancer. p interaction tested by likelihood ratio test for the multiplicative interaction term (TSH tertile × stratifier). BMI subgroup HRs show T3 versus T1 within each BMI stratum.

Abbreviations: BMI, body mass index; CI, confidence interval; HR, hazard ratio; TSH, thyroid-stimulating hormone.

3.4 Sensitivity analyses

When the outcome was restricted to invasive breast cancer alone (ICD-10 C50; 1,347 events), the TSH tertile association strengthened relative to the C50+D05 primary analysis: the HR for the highest versus lowest TSH tertile within the euthyroid range was 0.88 (95% CI 0.76–1.01; p = .068), and the postmenopausal subgroup estimate was HR 0.65 (95% CI 0.42–0.99; p = .049) (Supplementary Table S1). The systematic attenuation observed under the C50+D05 primary outcome relative to the C50-only sensitivity outcome is consistent with non-differential dilution of the principal signal by ductal carcinoma in situ events. Under a five-category exposure classification, inverse signals at the functional extremes were directionally consistent under both outcome definitions (Supplementary Table S2).

Principal findings were robust across additional sensitivity analyses (Supplementary Tables S3–S5). Extending the time-varying exposure lag from 6 to 12 months yielded near-identical estimates (T3 vs T1 HR 0.91 [95% CI 0.80–1.03] vs primary 0.90); including 1,887 women on thyroid medication who had been excluded a priori did not alter the principal TSH tertile association (HR 0.95 [0.84–1.07]); and Schoenfeld residual tests indicated that the proportional hazards assumption was met for the principal exposure variables (thyroid 3-category χ^2 = 0.11, p = .945; TSH tertile χ^2 = 0.27, p = .875). Variable-specific tests indicated proportional hazards violation for age and parity (and, in the TSH-tertile model, family history of breast cancer); consequently, the global test of proportional hazards was significant (see §4.4 Limitations).

4. Discussion

4.1 Principal findings

In this extended Korean cohort analysis using direct cancer-registry linkage and time-varying exposure modeling, the previously reported inverse association between euthyroid TSH and breast cancer risk [10] was attenuated relative to our prior Kangbuk Samsung Health Study (KSHS) report, reaching only directional consistency without conventional statistical significance under the primary outcome definition that incorporates both invasive disease and ductal carcinoma in situ (T3 vs T1 HR 0.93, 95% CI 0.81–1.07). Of our two a priori hypotheses (§1), the inverse TSH association was supported in attenuated form across three convergent lines of evidence described below, whereas the positive elevated-FT4 association was not replicated and is reinterpreted in §4.2 in light of the time-varying analysis.

First, time-varying Cox modeling [17]—pre-specified as a principal analytic strategy to address single-baseline exposure misclassification—shifted the static TSH tertile estimate from HR 0.95 to 0.90 (p = .096), with a parallel shift for continuous TSH (static 0.95 → time-varying 0.89, p = .094). These shifts, while not crossing conventional significance thresholds, are consistent with mild attenuation of static estimates toward the null by non-differential exposure misclassification, and the time-varying point estimates approximate the magnitude reported in the prior literature [10]. Second, the inverse association was substantially more pronounced among postmenopausal women (T3 vs T1 HR 0.69, 95% CI 0.47–1.02; multiplicative interaction p = .086 for T3 × postmenopausal). Although the interaction

confidence interval included the null, the point estimate is directionally and quantitatively similar to the corresponding postmenopausal estimate in our prior report [10] (aHR 0.52).

Third, restricting the outcome to invasive breast cancer alone (ICD-10 C50)—the outcome definition used in the prior KSHS analysis [10]—yielded an inverse estimate of greater magnitude than under the primary outcome (T3 vs T1 HR 0.88, 95% CI 0.76–1.01); the upper confidence bound included the null but the point estimate was consistent in direction and approximate magnitude with the prior report. The attenuation observed under the expanded outcome definition is most plausibly explained by non-differential dilution of an invasive-specific signal by ductal carcinoma in situ events, which carry greater detection-related variability and weaker biological linkage to thyroid hormone–driven proliferation.

In contrast to the inverse TSH findings, the previously reported positive association between elevated FT4 (>1.7 ng/dL) and breast cancer risk [10] was not replicated under either static or time-varying modeling in the present cohort. The methodological interpretation of this discordance—including the directional change observed under time-varying modeling—is addressed in §4.2.

4.2 Static vs time-varying modeling and the interpretation of episodic exposures

A further observation in the present analysis is the directional change observed under time-varying modeling for the suppressed-TSH and elevated-FT4 categories: the static HR for hyperthyroidism (TSH < 0.25 μIU/mL) of 0.69 shifted to 1.04 under time-varying modeling, and the static HR for FT4 > 1.7 ng/dL of 0.62 shifted to 1.05. We interpret these shifts cautiously: the relevant strata are modest in size (n = 2,466 with 22 events for hyperthyroidism; n = 1,197 with 12 events for elevated FT4), confidence intervals are wide and all include the null, and the directional change is therefore compatible with sampling variability alone. With this caveat, the parallel direction of the shift across both episodic exposures is consistent with the possibility that single-baseline measurement instability contributes to apparent risk gradients in this exposure range, though the present data are not sufficient to establish this as a definitive finding.

Under this interpretation, single-baseline measurement of TSH or FT4 captures only a transient snapshot of a fluctuating biological signal; participants measured during a transient excursion below the TSH reference range or above the FT4 reference range may regress toward euthyroidism on subsequent measurements. Time-varying modeling [17] reclassifies such participants at each repeated visit and can reveal that an apparent risk gradient associated with baseline categorization reflected, at least in part, the measurement event rather than a stable causal exposure. Conversely, biochemical exposures whose static and time-varying estimates agree closely—such as euthyroid-range TSH in the present analysis—are likely to be capturing more stable inter-individual differences.

In our prior KSHS report [10], we used a single baseline measurement to define both euthyroid-range TSH tertiles and the elevated-FT4 indicator, and reported a positive association between elevated FT4 and

breast cancer risk (aHR 1.98, 95% CI 1.02–3.83). With access to repeated measurements in the present extended cohort, the time-varying analysis suggests that our earlier elevated-FT4 finding may have, in part, reflected transient measurement events rather than sustained thyroid hormone exposure. The relevant strata remain small in both reports and a definitive conclusion is precluded, but the directional shift under time-varying modeling is consistent with this interpretation and underscores the value of repeated biomarker measurements when assessing episodic exposures. Similar cautions may apply to the broader literature on elevated FT4 or biochemical hyperthyroidism and breast cancer risk, which has been characterized by inconsistent findings across cohorts using single-measurement designs [6–8, 11, 12]. Analyses of episodic or extreme thyroid function exposures may benefit, when feasible, from incorporating repeated measurements as an analytic safeguard against this potential source of exposure misclassification.

4.3 Postmenopausal-specific signal and the role of estrogen receptor pathways

The persistence of an inverse TSH–breast cancer association specifically in postmenopausal women, observed under both the C50+D05 primary outcome and the C50-only sensitivity outcome, has biological plausibility. Thyroid hormone receptors share substantial structural homology with the estrogen receptor (ER), and triiodothyronine can act as a partial ER agonist in breast tissue, promoting proliferation in ER-positive mammary epithelium [4, 5, 22]. In the postmenopausal state, the decline in ovarian estradiol production renders the breast more dependent on alternative hormonal influences, including thyroid hormone–mediated stimulation; higher TSH—reflecting lower circulating thyroid hormone within the euthyroid range—would in this setting be expected to reduce proliferative drive on breast tissue. This interpretation generates a testable prediction: the postmenopausal inverse association should be confined to, or more pronounced in, ER-positive tumors. We were unable to test this directly because tumor estrogen receptor status was not available through the KCCR linkage in the present cohort, and subtype-specific replication in cohorts with linked tumor pathology is therefore a priority for future investigation.

Two considerations qualify the postmenopausal interpretation. First, postmenopausal women contributed only 11.3% of all incident cases (208 of 1,841 events; of these, 190 occurred within the euthyroid range that underlies the TSH-tertile subgroup estimate), reflecting the relatively young mean age (38.4 years) of the screening population; the postmenopausal estimate is therefore imprecise, and the multiplicative interaction *p* value was .086. Second, postmenopausal status was operationalized through self-reported cessation of menses for ≥12 months rather than through biochemical confirmation; misclassification at the menopausal transition cannot be excluded. Both considerations argue for treating the postmenopausal finding as suggestive rather than definitive.

4.4 Strengths and limitations

This study has several strengths. The sample size (162,702 women, 1,841 events) is among the largest reported for a single-cohort analysis of thyroid function and breast cancer risk in an Asian population [6], providing precision sufficient to detect hazard ratios of modest magnitude. Outcome ascertainment by direct cancer-registry linkage virtually eliminates dependence on self-report and minimizes differential misclassification by exposure status. The repeated thyroid function measurements available in KSHS [13] permitted formal time-varying analysis, which is uncommon in this literature. The use of harmonized covariates and analytic specifications matched to the prior KSHS report [10] enabled direct quantitative comparison of estimates and informed interpretation of the observed effect-size attenuation. Because the recruitment windows overlap during 2011–2018 and a subset of participants is common to both reports (§2.1), this comparison reflects an extension and partial re-analysis of the prior cohort rather than replication in a fully independent sample; the attenuation should be interpreted with this partial dependence in mind.

Several limitations warrant consideration. First, tumor-subtype information (estrogen receptor, progesterone receptor, HER2 status) was not available through the KCCR linkage in our cohort, precluding direct testing of the ER-mediated mechanism proposed in §4.3. Subtype-specific replication using cohorts with linked tumor pathology is needed to test this mechanism directly. Second, Schoenfeld residual tests [19, 20] indicated that the proportional hazards assumption was met for the principal exposure variables (TSH tertile $\chi^2 = 0.27$, $p = .875$; thyroid 3-category $\chi^2 = 0.11$, $p = .945$) but was violated for the age and parity covariates (and, in the TSH-tertile model, for family history of breast cancer; Supplementary Table S6); the principal hazard ratio estimates for the exposure variables are therefore considered unbiased, although residual time-dependent confounding by age and parity cannot be entirely excluded. Third, the cohort consists of Korean women undergoing voluntary comprehensive health screening, and generalizability to less health-conscious subgroups or to non-Korean populations is uncertain. Fourth, thyroid autoantibodies (anti-thyroid peroxidase, anti-thyroglobulin) were not systematically measured and could not be incorporated as effect modifiers, although autoantibody-positive autoimmune thyroid disease may exhibit distinct breast cancer risk patterns [8]. Fifth, dietary iodine intake—an environmental modulator of thyroid function with traditionally high consumption levels in Korean populations—was not measured. Sixth, missing-data handling was by complete-case analysis rather than multiple imputation; a multiple-imputation sensitivity analysis would have strengthened robustness assessment. Seventh, follow-up duration is summarized using the mean (6.94 years); given the relatively narrow span of recruitment dates (2011–2024), the distribution is unlikely to be materially skewed, but the median with interquartile range is not reported here. Finally, residual confounding by unmeasured factors—including obesity-related inflammation, insulin resistance, and reproductive-hormonal variation beyond menopausal status and parity—cannot be excluded.

5. Conclusion

In an extended Korean cohort of 162,702 women with cancer-registry-linked outcomes and pre-specified time-varying exposure modeling, the previously reported inverse association between euthyroid TSH and breast cancer risk [10] was attenuated under the primary outcome definition (incident invasive plus in

situ disease) but remained directionally consistent across three independent analyses: time-varying modeling of repeated exposure measurements, restriction to invasive-only outcome, and postmenopausal subgroup analysis. The previously reported positive association between elevated FT4 and breast cancer risk was not replicated under time-varying analysis. Taken together, these findings suggest that prior reports of episodic thyroid function exposures and breast cancer risk may have been partly driven by exposure misclassification inherent in single-baseline measurements, and that incorporation of repeated biomarker measurements is an analytic priority for future cohort studies of fluctuating biochemical exposures. The postmenopausal-specific inverse signal is biologically plausible under a thyroid hormone–estrogen receptor crosstalk mechanism but requires subtype-specific replication for confirmation.

Abbreviations

BMI, body mass index; CI, confidence interval; DCIS, ductal carcinoma in situ; ECLIA, electrochemiluminescent immunoassay; ER, estrogen receptor; FT4, free thyroxine; HR, hazard ratio; ICD-10, International Classification of Diseases, 10th revision; IQR, interquartile range; KCCR, Korea Central Cancer Registry; KSHS, Kangbuk Samsung Health Study; LRT, likelihood ratio test; PH, proportional hazards; RCS, restricted cubic spline; SD, standard deviation; TSH, thyroid-stimulating hormone.

Declarations

Acknowledgments

The authors thank the staff of the Kangbuk Samsung Hospital Total Healthcare Centers and the Korea Central Cancer Registry for their contributions to data collection and linkage.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

C.I. Yoon and H.S. Kim contributed equally as co-first authors and were jointly responsible for conceptualization, methodology, data curation, formal analysis, and writing of the original draft. Y.T. Kim contributed to methodology, formal analysis, software, and review and editing of the manuscript. S.J. Bae and J.-H. Lee contributed to investigation and review and editing. E.Y. Kim, C.H. Park, Y.L. Park, and J.-S. Yun contributed to data curation and resources. K.H. Lee (corresponding author) contributed to

conceptualization, supervision, project administration, and review and editing. All authors read and approved the final manuscript.

Data Availability

The datasets generated and/or analysed during the current study are not publicly available due to institutional privacy regulations governing health-screening and cancer-registry records but are available from the corresponding author (K.H. Lee; kwanho.lee@samsung.com) on reasonable request.

Ethics Approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institutional Review Board of Kangbuk Samsung Hospital (No. KBSMC 2026-03-023).

Consent to Participate

Because the analyses used de-identified data from existing health-screening and cancer-registry records, the requirement for individual informed consent was waived by the Institutional Review Board of Kangbuk Samsung Hospital.

Consent to Publish

Not applicable. The manuscript does not contain any individual person's identifiable data in any form.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249. https://doi.org/10.3322/caac.21660
2. Park EH, Jung KW, Park NJ, Kang MJ, Yun EH, Kim HJ, Kim JE, Kong HJ, Choi KS, Yang HK, Community of Population-Based Regional Cancer Registries (2025) Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2022. Cancer Res Treat 57:312–330. https://doi.org/10.4143/crt.2025.264
3. Britt KL, Cuzick J, Phillips KA (2020) Key steps for effective breast cancer prevention. Nat Rev Cancer 20:417–436. https://doi.org/10.1038/s41568-020-0266-x
4. Hall LC, Salazar EP, Kane SR, Liu N (2008) Effects of thyroid hormones on human breast cancer cell proliferation. J Steroid Biochem Mol Biol 109:57–66. https://doi.org/10.1016/j.jsbmb.2007.12.008
5. Tang HY, Lin HY, Zhang S, Davis FB, Davis PJ (2004) Thyroid hormone causes mitogen-activated protein kinase-dependent phosphorylation of the nuclear estrogen receptor. Endocrinology 145:3265–3272. https://doi.org/10.1210/en.2004-0308
6. Tran TV, Kitahara CM, Leenhardt L, de Vathaire F, Boutron-Ruault MC, Journy N (2023) The effect of thyroid dysfunction on breast cancer risk: an updated meta-analysis. Endocr Relat Cancer

30:e220155. https://doi.org/10.1530/ERC-22-0155

7. Tosovic A, Bondeson AG, Bondeson L, Ericsson UB, Malm J, Manjer J (2010) Prospectively measured triiodothyronine levels are positively associated with breast cancer risk in postmenopausal women. Breast Cancer Res 12:R33. https://doi.org/10.1186/bcr2587
8. Tosovic A, Becker C, Bondeson AG, Bondeson L, Ericsson UB, Malm J, Manjer J (2012) Prospectively measured thyroid hormones and thyroid peroxidase antibodies in relation to breast cancer risk. Int J Cancer 131:2126–2133. https://doi.org/10.1002/ijc.27470
9. Kuijpens JL, Nyklíček I, Louwman MW, Weetman TA, Pop VJ, Coebergh JW (2005) Hypothyroidism might be related to breast cancer in post-menopausal women. Thyroid 15:1253–1259. https://doi.org/10.1089/thy.2005.15.1253
10. Kim EY, Chang Y, Lee KH, Yun JS, Park YL, Park CH, Ahn J, Shin H, Ryu S (2019) Serum concentration of thyroid hormones in abnormal and euthyroid ranges and breast cancer risk: a cohort study. Int J Cancer 145:3257–3266. https://doi.org/10.1002/ijc.32283
11. Søgaard M, Farkas DK, Ehrenstein V, Jørgensen JO, Dekkers OM, Sørensen HT (2016) Hypothyroidism and hyperthyroidism and breast cancer risk: a nationwide cohort study. Eur J Endocrinol 174:409–414. https://doi.org/10.1530/EJE-15-0989
12. Hellevik AI, Åsvold BO, Bjøro T, Romundstad PR, Nilsen TI, Vatten LJ (2009) Thyroid function and cancer risk: a prospective population study. Cancer Epidemiol Biomarkers Prev 18:570–574. https://doi.org/10.1158/1055-9965.EPI-08-0911
13. Zhang Y, Chang Y, Ryu S, Cho J, Lee WY, Rhee EJ, Kwon MJ, Pastor-Barriuso R, Rampal S, Han WK, Shin H, Guallar E (2014) Thyroid hormones and mortality risk in euthyroid individuals: the Kangbuk Samsung health study. J Clin Endocrinol Metab 99:2467–2476. https://doi.org/10.1210/jc.2013-3832
14. Shin HR, Won YJ, Jung KW, Kong HJ, Yim SH, Lee JK, Noh HI, Lee JK, Pisani P, Park JG, Members of the Regional Cancer Registries (2005) Nationwide cancer incidence in Korea, 1999 ~ 2001; first result using the national cancer incidence database. Cancer Res Treat 37:325–331. https://doi.org/10.4143/crt.2005.37.6.325
15. Cha CD, Park CS, Shin HC, Han J, Choi JE, Kim JH, Jung KW, Lee SB, Nam SE, Yoon TI, Kang YJ, Kim Z, Jung SY, Kim HA (2024) Korean Breast Cancer Society Breast cancer statistics in Korea, 2021. J Breast Cancer 27:351–361. https://doi.org/10.4048/jbc.2024.0213
16. Harrell FE (2015) Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis, 2nd edn. Springer, Cham
17. Therneau TM, Grambsch PM (2000) Modeling survival data: extending the Cox model. Springer, New York
18. Firth D (1993) Bias reduction of maximum likelihood estimates. Biometrika 80:27–38. https://doi.org/10.1093/biomet/80.1.27
19. Schoenfeld D (1982) Partial residuals for the proportional hazards regression model. Biometrika 69:239–241. https://doi.org/10.1093/biomet/69.1.239

20. Grambsch PM, Therneau TM (1994) Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81:515–526. https://doi.org/10.1093/biomet/81.3.515
21. von Elm E, Altman DG, Egger M, Pocock SJ, Götzsche PC, Vandenbroucke JP, STROBE Initiative (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 370:1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X
22. Hercbergs A, Mousa SA, Leinung M, Lin HY, Davis PJ (2018) Thyroid hormone in the clinic and breast cancer. Horm Cancer 9:139–143. https://doi.org/10.1007/s12672-018-0326-9

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryCCC.docx