Association between thyroid function and heart rate monitored by wearable devices in patients with hypothyroidism

Kim, Ki Hun; Lee, Juhui; Ahn, Chang Ho; Yu, Hyeong Won; Choi, June Young; Lee, Ho Young; Lee, Won Woo; Moon, Jae Hoon

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Association between Thyroid Function and Heart Rate Monitored by Wearable Devices in Patients with Hypothyroidism

Ki-Hun Kim[1,2,3], Juhui Lee[4], Chang Ho Ahn[5], Hyeong Won Yu[6], June Young Choi[6], Ho-Young Lee[7], Won Woo Lee[7], Jae Hoon Moon[4,5]

[1]Department of Industrial Engineering, Pusan National University, Busan, Korea; [2]Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherland; [3]Department of Industrial Engineering, Ulsan National Institute of Science and Technology; [4]Thyroscope Inc., Ulsan; Departments of [5]Internal Medicine, [6]Surgery, [7]Nuclear Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Korea

Background: Heart rate (HR) monitored by a wearable device (WD) has demonstrated its clinical feasibility for thyrotoxicosis subjects. However, the association of HR monitored by wearables with hypothyroidism has not been examined. We assessed the association between serum thyroid hormone concentration and three WD-HR parameters in hypothyroid subjects.

Methods: Forty-four subjects scheduled for radioactive iodine therapy (RAI Tx) after thyroid cancer surgery were included. Thirty subjects were prepared for RAI Tx by thyroid hormone withdrawal (hypothyroidism group) and 14 subjects by recombinant human thyrotropin (control group). Three WD-HR parameters were calculated from the HR data collected during rest, during sleep, and from 2:00 AM to 6:00 AM, respectively. We analyzed the changes in conventionally measured resting HR (On-site rHR) and WD-HR parameters relative to thyroid hormone levels.

Results: Serum free thyroxine (T4) levels, On-site rHR, and WD-HR parameters were lower in the hypothyroid group than in the control group at the time of RAI Tx. WD-HR parameters also reflected minute changes in free T4 levels. A decrease in On-site rHR and WD-HR parameters by one standard deviation (On-site rHR, approximately 12 bpm; WD-HR parameters, approximately 8 bpm) was associated with a 0.2 ng/dL decrease in free T4 levels ($P<0.01$) and a 2-fold increase of the odds ratio of hypothyroidism ($P<0.01$). WD-HR parameters displayed a better goodness-of-fit measure (lower quasi-information criterion value) than On-site rHR in predicting the hypothyroidism.

Conclusion: This study identified WD-HR parameters as informative and easy-to-measure biomarkers to predict hypothyroidism.

Keywords: Wearable electronic devices; Hypothyroidism; Heart rate

INTRODUCTION

Hypothyroidism is a state of deficiency of thyroid hormone, which is essential for maintaining normal cellular function in the human body. Primary hypothyroidism is the most common cause of hypothyroidism and accounts for 99% of all cases

[1,2]. The prevalence of overt hypothyroidism in the general population has been reported to be up to 2% and that of subclinical hypothyroidism as up to 10% [3,4]. Because thyroid hormone acts ubiquitously, hypothyroidism results in various symptoms and signs, including fatigue, shortness of breath, mood lability, weight gain, constipation, cold intolerance, dry skin, and edema [5]. However, these symptoms and signs are relatively nonspecific, and the accuracy of various symptom scoring systems in identifying hypothyroid patients is low [6]. Therefore, a thyroid function test (TFT) performed in the laboratory using patients' blood samples is the gold standard for diagnosing hypothyroidism. Due to the nonspecific symptoms and the requirement for a blood test, diagnosing hypothyroidism is often delayed until patients reach a severe status and visit the clinic. Therefore, a simple, easy-to-measure parameter to predict thyroid status would be useful for the timely management of hypothyroidism.

Recently developed wearable devices (WDs) collect accurate data about physical activity levels and heart rate (HR) in real time [7-9]. We previously demonstrated that a resting HR (rHR) calculated from these data reflected the severity of thyrotoxicosis in the subjects examined [10,11]. This rHR can be continuously calculated from the real-time WD data, and hence enables continuous and accurate monitoring of thyroid function [10]. The results of our previous study imply the possibility that such HR parameters might also represent the hypothyroidism severity. However, while palpitation is one of the most common symptoms of thyrotoxicosis, bradycardia is not a major symptom in hypothyroid patients, and some patients actually experience palpitation in the hypothyroid state [6]. Therefore, it is difficult to guarantee that hypothyroid patients have the similar association between HR data from wearables and thyroid function observed in thyrotoxic patients.

In this study, we examined if and how well HR and activity data collected by commercial WDs reflected thyroid function across the clinical course of thyroid hormone withdrawal in patients undergoing thyroid hormone replacement therapy after total thyroidectomy due to thyroid cancer.

METHODS

Study design and participants

This was a single-center, prospective observational study. Subjects were recruited from the outpatient clinic at Seoul National University Bundang Hospital (SNUBH) from August 2019 to July 2020. Subjects aged 20 to 60 years who had undergone total thyroidectomy due to differentiated thyroid cancer (DTC) and who were scheduled for radioactive iodine therapy (RAI Tx) were eligible to participate. The study excluded subjects who had other comorbidity affecting their HR, were taking medicine that affected HR, or could not use the WD and its mobile application. A total of 48 patients were screened for participation, and 45 were finally enrolled. Among them, one participant withdrew because of a skin reaction to the device on the wrist. All participants were informed about the study and provided written informed consent. The study was approved by the SNUBH Institutional Review Board (IRB #B-1906-547-304) and registered at ClinicalTrials.gov (trial registration #NCT04332770).

Procedures

The study design is illustrated in Fig. 1. At the first on-site visit, potential candidates were provided with a device and brief instructions for its use. They were instructed to wear the device as much as possible throughout the day and during sleep, and were told that if they did not wear the device or synchronize it with the application for more than 5 consecutive days, they would be withdrawn from the study due to poor adherence. Apart from wearing the device and using the mobile application, there was no other intervention for subjects undergoing conventional management for follow-up and RAI Tx for DTC. Subjects in the hypothyroidism group (*n*=30) changed their levothyroxine (LT4) medication to liothyronine (LT3) 4 weeks prior to RAI Tx, and stopped LT3 2 weeks prior to RAI Tx. After this preparation period, they entered a hypothyroid state at the time of RAI Tx. In the control group (*n*=14), subjects received recombinant human thyrotropin (rhTSH) injections instead of discontinuing thyroid hormone and did not experience a hypothyroid state. Because rhTSH is only covered by health insurance for low-dose RAI Tx (≤30 mCi of [I^{131}]) in Korea, the control group underwent low-dose RAI Tx, whereas the hypothyroidism group underwent high-dose RAI Tx (≥100 mCi of [I^{131}]). The dose of [I^{131}] was determined by an endocrinologist (J.H.M.) and two nuclear medicine specialists (W.W.L. and H.Y.L.), according to the surgical pathology and clinical information. Patients started LT4 medication at a dose of 2.0 μg/kg of body weight on the day after total thyroidectomy. A LT4 dose of 12.5 μg/day was increased or decreased as necessary to reach the appropriate target range of serum thyroid stimulating hormone (TSH) concentration. All patients underwent treatment according to the Korean Thyroid Association guidelines [12].
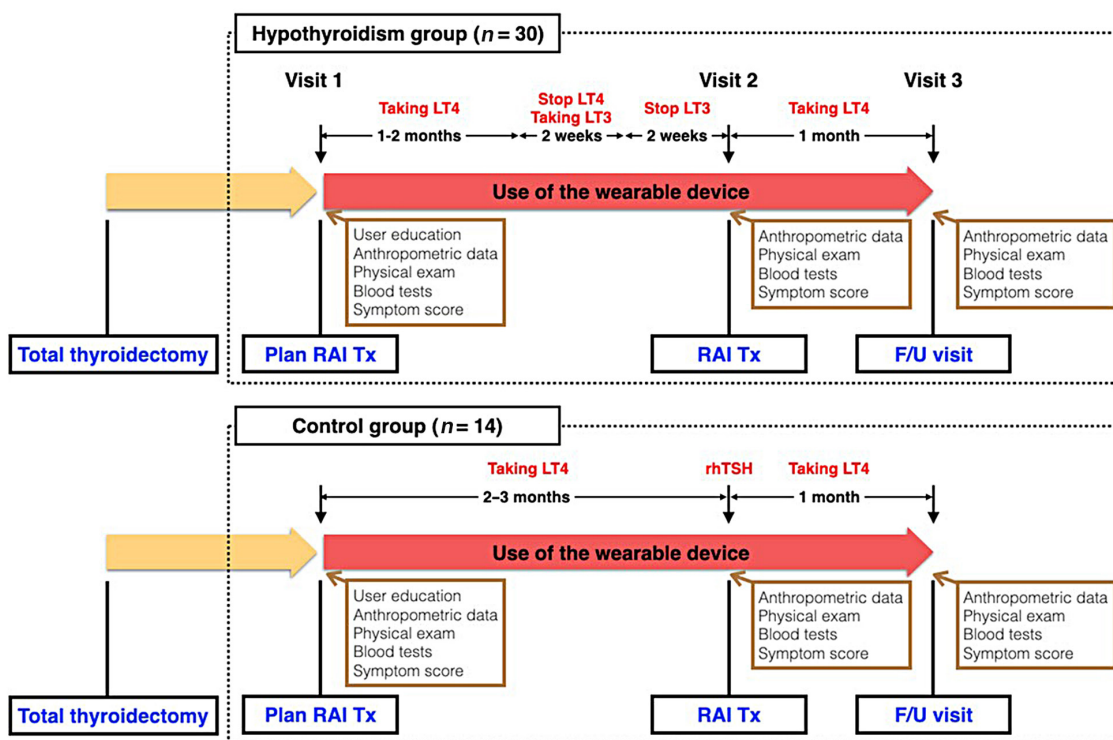
Fig. 1. Study design and flow. Blood tests included a thyroid function test and other biochemical tests. LT4, levothyroxine; LT3, liothyronine; RAI Tx, radioactive iodine (I^{131}) therapy; F/U, follow-up; rhTSH, recombinant human thyrotropin.

Wearable devices and applications

We used the Fitbit Charge 2 (Fitbit, San Francisco, CA, USA) and the Fitbit application for iOS (Apple, Cupertino, CA, USA) or Android (Google, Mountain View, CA, USA). The firmware version for Fitbit Charge 2 was 22.58.0 at the end of the study, and the latest version was maintained continuously over the study period. Activity and HR data were collected by the 3-axis accelerometer and photoplethysmography sensor of the device, respectively. During the study period, each participant's Fitbit account information, including identification and password, were shared with the researchers, allowing them to access the online Fitbit accounts to monitor their synchronization status between the Fitbit data server and the WD. After the end of the study, the participant's device account was terminated, and each participant created a new account to use for his/her device.

Zulewski's clinical score for hypothyroidism

To assess participants' clinical status of hypothyroidism at each clinic visit, the overseeing endocrinologist (J.H.M.) evaluated the patients using the Zulewski's clinical score for hypothyroidism [13]. The score consists of 12 items, including seven symptoms (diminished sweating, hoarseness, paresthesia, dry skin,

constipation, impairment of hearing, and weight increase) and five physical signs (slow movements, delayed ankle reflex, coarse skin, periorbital puffiness, and cold skin) of hypothyroidism. Each item was scored as 0 (absent) or 1 (present), and item scores were totaled to obtain an overall score, which could range from 0 to 12 points.

HR parameters

The American Heart Association defines rHR as the heart beats per minute (bpm) pumping the lowest amount of blood someone needs when they are in a resting position. The rHR is known to be altered by emotional state, medication, or comorbidity. Based on this fact, four different HR parameters (On-site rHR, rHR from wearable device [WD-rHR], HR during sleep from wearable device [WD-sleepHR], and HR from 2:00 AM to 6:00 AM from wearable device [WD-2to6HR]) were defined as HR measures expected to represent the hypothyroid state. The parameters were based on different measurement methods and calculating algorithms as follows. On-site rHR is the HR measured manually on the right wrist in a seated position after at least 10 minutes of resting on each visit day. The other three HR parameters (WD-HR parameters) were obtained from the

HR log generated by the WD. WD-rHR is the mean of daily rHR values for the past 5 days before a visit day. A daily rHR value is the median of the HR data extracted from the 1-day HR log within time windows that indicate the absence of physical activity during the preceding ≥15 minutes. Detailed HR and activity data were downloaded from the Fitbit database [14]. WD-sleepHR and WD-2to6HR are the mean of daily sleepHR values and daily 2to6HR values, respectively, for the previous 5 days before a visit day. A daily sleepHR value is the median of HR data collected when the participant slept within 24 hours from midnight, while a daily 2to6HR value is the median of HR data extracted from 2:00 AM to 6:00 AM, regardless of sleeping state or physical activity. Whether the user was sleeping or not was determined by the WD's own algorithm. Please note that only for visit 1, WD-HR parameters were calculated with the HR data collected for the 5 days after visit 1 because the HR data collection started at the day of visit 1.

Anthropometric and biochemical measurements

We measured the subjects' height and weight while wearing light clothing and without shoes to the nearest 0.1 cm and 0.1 kg, respectively. Body mass index (BMI) was calculated by determining the ratio between weight and the square of the height and expressed in kg/m^2. Right arm blood pressure was measured with the subject in a seated position after at least 10 minutes of resting. Serum levels of blood urea nitrogen, creatinine, and glucose were measured by automated standard laboratory methods (Hitachi 747, Hitachi, Tokyo, Japan). Serum total protein, albumin, total bilirubin, alkaline phosphatase, aspartate aminotransferase, and alanine aminotransferase were measured with an autoanalyzer (TBA-200FR, Toshiba, Tokyo, Japan). For TFT, concentrations of serum free thyroxine (T4; DiaSorin, Saluggia, Italy) and TSH (CIS Bio International, Gif-sur-Yvette, France) were measured using immunoradiometric assays. The reference ranges for free T4 and TSH were 0.89–1.79 ng/dL and 0.3–4.0 mIU/L, respectively. Hypothyroidism was defined based on the results of the TFT, as low free T4. TSH was not used for defining hypothyroidism because subjects in the control group used rhTSH, which increased serum TSH levels, regardless of the peripheral action of thyroid hormone.

Data analysis

Data were expressed as the mean±standard deviation (SD) or median (interquartile range). To compare variables in the hypothyroidism and control groups, we used Student's t test or the Mann-Whitney U test for continuous variables and Fisher's ex-

act probability test for categorical variables. The paired t test and the Wilcoxon signed-rank test were used to analyze changes in parameters during the study period. The linear and binary logistic regression models of generalized estimating equations (GEE) [15,16] were fit to the data of all subjects to evaluate the association of each HR parameters with serum free T4 levels and hypothyroidism, respectively. Each regression model had serum free T4 or hypothyroidism as its dependent variable and a standardized variable of HR parameter as a predictor. The regression models also included age, gender, and BMI as covariates. The quasi-information criterion (QIC) [17] of each regression model

Table 1. Baseline Characteristics before Total Thyroidectomy

Characteristic	Hypothyroidism (n=30)	Control (n=14)	P value
Age, yr	38.1±8.4	42.4±7.4	0.112[a]
Sex			
Male	15 (50.0)	2 (14.3)	0.044[b]
Female	15 (50.0)	12 (85.7)	
Body mass index, kg/m^2	25.2±4.6	23.2±4.8	0.204[a]
Systolic blood pressure, mm Hg	127.4±12.5	122.7±9.4	0.221[a]
Diastolic blood pressure, mm Hg	77.9±10.2	72.8±10.2	0.130[a]
On-site rHR, bpm	74.0±9.5	73.4±9.9	0.850[a]
Thyroid function test			
Free T4, ng/dL	1.32±0.24	1.37±0.14	0.470[a]
TSH, mIU/L	1.91 (1.27)	1.44 (0.95)	0.332[c]
Glucose, mg/dL	97.9±22.9	93.7±7.2	0.514[a]
BUN, mg/dL	14.5±10.8	12.1±3.9	0.434[a]
Creatinine, mg/dL	0.74±0.18	0.64±0.09	0.049[a]
Total cholesterol, mg/dL	181.3±32.3	172.1±32.9	0.389[a]
Total protein, g/dL	7.3±0.4	7.4±0.6	0.637[a]
Albumin, g/dL	4.4±0.3	4.3±0.4	0.885[a]
AST, mg/dL	23.5±8.1	23.4±4.6	0.940[a]
ALT, mg/dL	24.9±16.7	22.9±9.9	0.685[a]
WBC, no/mm^3	6,838±1,966	5,049±2,384	0.023[a]
Hemoglobin, mg/dL	13.8±1.8	13.2±1.0	0.190[a]
Platelet, no/mm^3	283±56	281±61	0.913[a]

Values are expressed as mean±standard deviation, number (%), or median (interquartile range).

rHR, resting heart rate; bpm, beats per minute; T4, thyroxine; TSH, thyroid stimulating hormone; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ALT, alanine aminotransferase; WBC, white blood cell.

[a]Derived from Student's t test; [b]Derived from Fisher exact probability test; [c]Derived from Mann-Whitney U test.

was also computed as a goodness-of-fit measure. In the binary logistic model of GEE, the standardized variable resulted from not only standardizing the HR parameter to compare the odds ratios (ORs), but also multiplying the standardized HR parameter by '–1' to intuitively interpret the ORs (i.e., how much the probability ratio of hypothyroidism to euthyroidism increases per 1 SD 'decrease' in each parameter). A two-tailed $P<0.05$ was considered statistically significant. All statistical analyses and data preparation were performed using Python SciPy 1.0 version 3.8.5 (Python Software Foundation, Beaverton, OR, USA) [18] and SAS/STAT PROC GENMOD procedure (SAS University Edition, SAS Institute Inc., Cary, NC, USA).

RESULTS

Baseline characteristics

Participants' baseline characteristics are summarized in Table 1. There were no differences in age and BMI between the hypothyroidism and the control groups. Males were more frequent in the hypothyroidism group. Blood pressure, On-site rHR, TFT, and other anthropometric and biochemical parameters did not differ between the hypothyroidism and control groups, except for creatinine levels and leukocyte count.

Changes in TFT and associated parameters across visits

The TFT results and associated parameters for each visit are presented in Table 2. In the hypothyroidism group, free T4 and TSH levels were significantly changed according to the visits, with decreased free T4 and increased TSH levels in visit 2 (off LT4) compared with visit 1 or 3 (on LT4). All HR parameters displayed the same trends as free T4 levels in the hypothyroidism group. In the control group, although free T4 levels were not significantly different between visits, free T4 levels were slightly decreased at visit 2 compared with visit 1 (1.69 ± 0.20 vs. 1.56 ± 0.18, $P=0.072$). Along with this trend, all WD-HR parameters significantly decreased at visit 2 compared with visit 1. Although On-site rHR had a lower numerical mean value at visit 2 than at visit 1, this difference was not statistically significant. Fig. 2 presents HR parameters were altered consistently with free T4 levels while Zulewski's clinical scores were not.

Table 2. Change of TFT Results and Associated Parameters during the Study Period

TFT results and associated parameters	Visit 1	Visit 2	Visit 3	*P* value (Visit 1 vs. 2)[a]	*P* value (Visit 2 vs. 3)[b]
Hypothyroidism group, *n*	24	30	23	-	-
Free T4, ng/dL	1.73±0.25	0.08±0.06	1.58±0.31	<0.001[c]	<0.001[c]
TSH, mIU/L	0.25 (1.54)	84.1 (49.3)	0.17 (1.45)	<0.001[d]	<0.001[d]
Zulewski's clinical score	1 (2)	1 (2)	1 (2)	0.040[d]	0.360[d]
On-site rHR, bpm	85.5±10.7	73.0±10.8	83.4±11.7	<0.001[c]	0.022[c]
WD-rHR, bpm	68.9±8.9	60.2±6.4	68.0±6.9	<0.001[c]	<0.001[c]
WD-sleepHR, bpm	62.8±9.7	55.3±6.1	62.8±7.3	<0.001[c]	<0.001[c]
WD-2to6HR, bpm	64.6±8.8	56.7±6.7	63.7±6.6	<0.001[c]	<0.001[c]
Control group, *n*	12	14	8	-	-
Free T4, ng/dL	1.69±0.20	1.56±0.18	1.68±0.33	0.072[c]	0.538[c]
TSH, mIU/L	0.26 (0.61)	157.63 (50.61)	0.11 (0.36)	0.002[d]	0.012[d]
Zulewski's clinical score	2 (2)	2 (1)	1.5 (2)	0.589[d]	0.705[d]
On-site rHR, bpm	85.3±14.4	80.0±12.5	83.9±11.6	0.312[c]	0.302[c]
WD-rHR, bpm	67.3±7.2	63.6±6.8	67.0±6.1	0.019[c]	0.062[c]
WD-sleepHR, bpm	63.6±6.5	60.2±6.9	63.7±6.5	0.004[c]	0.022[c]
WD-2to6HR, bpm	63.8±6.2	60.5±7.2	63.1±6.4	0.017[c]	0.069[c]

Values are expressed as mean±standard deviation or median (interquartile range).

TFT, thyroid function test; T4, thyroxine; TSH, thyroid stimulating hormone; rHR, resting heart rate; bpm, beats per minute; WD-rHR, rHR from wearable device; WD-sleepHR, HR during sleep from wearable device; WD-2to6HR, HR from 2:00 AM to 6:00 AM from wearable device.

[a]*P* value of statistical equality tests between Visit 1 and 2; [b]*P* value of statistical equality tests between Visit 2 and 3; [c]Derived from paired *t* test; [d]Derived from Wilcoxon signed-rank test.
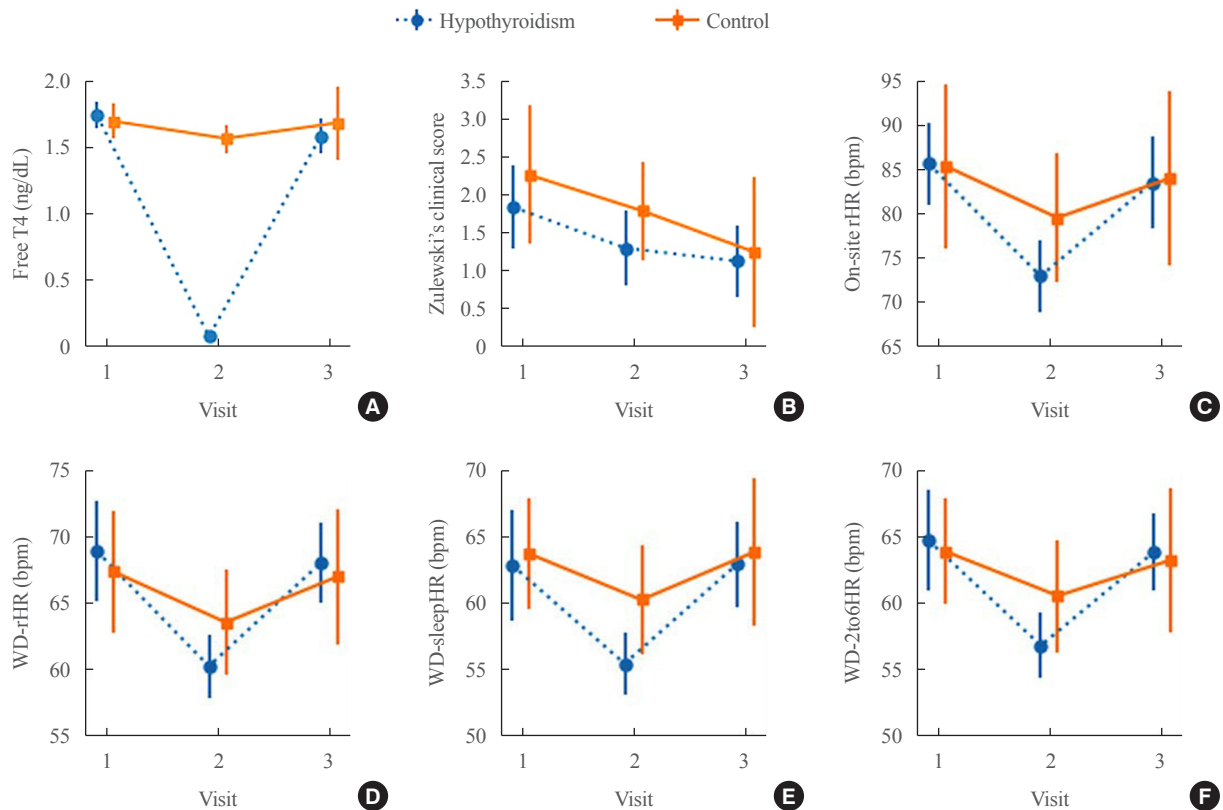
Fig. 2. Change of (A) serum free thyroxine (T4) levels, (B) Zulewski's clinical score, (C) On-site resting heart rate (rHR), (D) rHR from wearable device (WD-rHR), (E) HR during sleep from wearable device (WD-sleepHR), and (F) HR from 2:00 AM to 6:00 AM from wearable device (WD-2to6HR) during the study period. Error bars represent 95% confidence interval of the means. The symbols indicate two groups.

Association between serum free T4 levels and each HR parameter

The linear regression models of GEE were fit to the data of all subjects to verify the relationship between serum free T4 and each HR parameter (Table 3). The descriptive statistics of the data on serum free T4 and the HR parameters are presented in Supplemental Table S1. The dependent variable and predictor of each regression model were serum free T4 and the standardized variable of each HR parameter, respectively. The linear regression models additionally included age, gender, and BMI as covariates. The standardized variables had a mean of 0 and SD of 1, which led to their comparable regression coefficients. All HR parameters had statistically significant coefficients for serum free T4 levels. The coefficients were similar, ranging from 0.2 to 0.25. The maximum QIC difference between the regression models was 0.45, which was small compared with the QIC difference threshold of 2 set by an earlier study [19], in which the QIC difference <2 between two models indicates they have similar goodness of fit.

OR of each HR parameter for hypothyroidism

The relationship between hypothyroidism and each associated parameter was estimated by the binary logistic regression models of GEE (Table 4). Age, gender, and BMI were covariates in these models. All parameters had significant coefficients for predicting hypothyroidism. However, On-site rHR displayed a relatively lower OR (1.88) for hypothyroidism with a 1 SD decrease compared with the ORs of the WD-HR parameters, which were similar (2.24 to 2.57). The models including WD-HR parameters also produced similar QIC values, ranging from 119.76 to 120.69. Compared with these QIC values, the model with On-site rHR had a substantially larger QIC value of 123.52, based on the QIC difference threshold of 2 [19].

DISCUSSION

This study demonstrated that WD-HR parameters were significantly associated with thyroid function in hypothyroidism, as reported previously for thyrotoxicosis [10]. In the hypothyroid-

Table 3. Results for the Linear Regression of Free T4 Levels on Each HR Parameter

HR parameter[a]	Coefficient	SE	95% CI	*P* value	QIC
On-site rHR	0.213	0.059	0.098–0.327	<0.001	112.722
WD-rHR	0.205	0.051	0.105–0.305	<0.001	112.620
WD-sleepHR	0.247	0.056	0.137–0.357	<0.001	112.272
WD-2to6HR	0.201	0.057	0.089–0.313	<0.001	112.414

Each regression model had serum free T4 as its dependent variable and a HR parameter as a predictor. Age, gender, and body mass index were included as covariates.
T4, thyroxine; HR, heart rate; SE, standard error; CI, confidence interval; QIC, quasi-information criterion; rHR, resting heart rate; WD-rHR, rHR from wearable device; WD-sleepHR, HR during sleep from wearable device; WD-2to6HR, HR from 2:00 AM to 6:00 AM from wearable device.
[a]HR parameters were standardized (mean, 0; SD, 1).

Table 4. Results for the Binary Logistic Regression of Hypothyroidism on Each HR Parameter

HR parameter[a]	Coefficient (OR)	SE	95% CI	*P* value	QIC
On-site rHR	0.630 (1.878)	0.238	0.164–1.096	0.008	123.517
WD-rHR	0.942 (2.565)	0.280	0.393–1.491	<0.001	120.124
WD-sleepHR	0.859 (2.358)	0.241	0.387–1.331	<0.001	119.764
WD-2to6HR	0.806 (2.239)	0.255	0.306–1.306	0.002	120.691

Each regression model had hypothyroidism as its dependent variable and a HR parameter as a predictor. Age, gender, and body mass index were included as covariates.
HR, heart rate; OR, odds ratio; SE, standard error; CI, confidence interval; QIC, quasi-information criterion; rHR, resting heart rate; WD-rHR, rHR from wearable device; WD-sleepHR, HR during sleep from wearable device; WD-2to6HR, HR from 2:00 AM to 6:00 AM from wearable device.
[a]HR parameters were standardized (mean, 0; SD, 1) and multiplied by –1 to see the OR of hypothyroidism when a parameter decreases by its one standard deviation.

ism group, the subjects rapidly entered a hypothyroid state for the 2 weeks prior to RAI Tx. WD-HR parameters and On-site rHR reflected this sudden change in thyroid function. In the control group, although comparable free T4 levels were predicted across visits, an unexpected trend for altered free T4 levels between visits 1 and 2 was observed and WD-HR parameters displayed the same trend. Because the mean dose of LT4 at each visit were not significantly different (data not shown), other factors including drug compliance or absorption rate might have produced the observed free T4 levels in the control group. A comparison of free T4 levels and HR parameters measured at the visit for regular follow-up (on LT4) and the visit for RAI Tx (off LT4 in the hypothyroidism group; rhTSH in the control group) in each group (Supplemental Table S2) revealed that the change in WD-HR parameters was associated with the change in free T4 levels in a significant dose-dependent manner, with large changes in free T4 levels and HR parameters in the hypothyroid group and smaller changes in the control group. In this analysis, On-site rHR did not significantly change relative to the slight change in free T4 levels in the control group, suggesting

that WD-HR parameters are more sensitive than the conventionally measured rHR. In both groups, Zulewski's clinical score did not change consistently with free T4 levels during the study period, suggesting that a scoring system using clinical symptoms and signs cannot reflect rapidly altered thyroid function in hypothyroidism because subjective symptoms may not be clear in rapidly progressing hypothyroidism.

This study demonstrated that the HR parameters had all the statistically significant and similar levels for the linear association with serum free T4 levels in the linear regression models. All models also had similar goodness-of-fit values, revealing the subtle differences between their QIC values. Unlike the linear regression analysis, the binary logistic regression results suggested that the WD-HR parameters were more appropriate for evaluating the hypothyroidism risk than On-site rHR. WD-HR parameters had larger coefficients and ORs and lower QIC values for their models than On-site rHR. In addition to this superior fitness, WD-HR parameters are more convenient than On-site rHR measurement, which requires patients to maintain a resting position without activity for 10 to 15 minutes and to manually

measure their HR. This procedure makes it challenging to measure rHR manually, several times daily. In contrast, WD-HR parameters only require patients to wear a WD. In particular, WD-sleepHR and WD-2to6HR do not require the device to be worn all day to collect the necessary data, and WD-2to6HR does not even require the activity data needed for WD-rHR and WD-sleepHR. Thus, WD-HR parameters are more informative of the hypothyroidism risk and easier to measure than other assessment tools or conventional methods of measuring rHR.

Highly detailed, longitudinal measurements of physical indices can be collected from WDs. Such WD data are expected to provide more accurate and objective information than subjective symptoms. However, the HR data from WDs can still have some variability caused by various factors, such as body position or emotional state even though we can exclude the effect of physical activity using the activity data. This consideration led to the creation of WD-sleepHR and WD-2to6HR that use the HR data extracted within time windows when most participants are likely to be in a resting position with minimal physical activity. In addition, the calculation algorithms for WD-HR parameters could offset this variability by integrating continuous and high-definition data. WD-rHR, WD-sleepHR, and WD-2to6HR were calculated as the 5-day mean of daily rHR, daily sleepHR, and daily 2to6HR, respectively, from the HR data. This calculation smoothed out the variability in daily repeated short-term factors such as emotional state or body position. In contrast, long-term, continuous factors, such as disease status, still influenced WD-HR parameters. Therefore, the present study demonstrates that these parameters accurately represent the disease state of hypothyroidism. Furthermore, the accuracy of WD-HR parameters for predicting thyroid function will improve with other biosignals associated with thyroid function such as sweating, tremor, or respiration rate.

This study has significant clinical implications. Many hypothyroid patients taking LT4 at a stable dose display a stable clinical course and maintain their TFT results within the normal range. Nevertheless, they are required to visit the hospital every few months to have blood tests and evaluate whether the current dose of the drug is appropriate, as there is no other way to evaluate their thyroid function status. However, based on the results of this study, if thyroid function status can be monitored with WDs, the expense of frequent hospital visits and blood tests can be reduced in well-controlled patients. On the other hand, even if taking a stable dose of LT4, patients' serum thyroid hormone levels can vary due to changes in LT4 demand or absorption rate associated with body weight change, drug compliance, or

comorbidity in the gastrointestinal tract. Without appropriate action, patients in this situation will experience symptoms of thyroid dysfunction until their next hospital visit. Thyroid function monitoring by wearable-derived biosignals could detect the change in thyroid status and let the patients know that their symptoms are due to thyroid dysfunction and visit the hospital. This monitoring could also detect when the patients who undergo RAI Tx due to hyperthyroidism enter hypothyroid status and let them start LT4 replacement therapy in a timely manner. This study was conducted with commercially available WDs and applications. Thus, the WD-HR parameters can be immediately applied to patients and serve as a ready-to-use algorithm for a hypothyroidism management service. In addition, the results demonstrate that WD-rHR and WD-sleepHR, which consider the users' physical activity and/or body position, and the WD-2to6HR parameter, collected in the early morning hours, were all similarly associated with serum free T4 levels even though WD-2to6HR did not use the physical activity data. This result suggests that the convenience of the data collection could be further improved in actual clinical application.

This study has some limitations to be considered when interpreting the results. First, younger people, who are generally more familiar with using WDs and mobile applications than older people, were more likely to participate in the study. The age of the participants in this study ranged from 21 to 54 years. Thus, these results may not be generalizable to people who are unfamiliar with using smart devices. More elderly patients will use thyroid function monitoring solutions using wearable biosignals as the elderly get used to smart devices and the wearables. Therefore, future research is required to validate the clinical feasibility of the study results for elderly patients. Second, the gender ratio was different between two groups. Subjects in the hypothyroidism group might have a more advanced stage of DTC because they underwent high-dose RAI Tx. This difference might have resulted this gender difference, as men tend to have more aggressive disease at diagnosis [20]. However, this difference was not problematic for the HR parameters that resulted from the mean of HR, which was known to be similar in gender groups [21]. Third, this single-center study had a small sample size, with an imbalance between the hypothyroidism and control group sizes. Further investigations that include more participants and study sites are warranted to confirm the results of this study. Fourth, subjects in the hypothyroidism group showed severe hypothyroidism with TSH of more than 30 mIU/L because we performed TFT on the day of RAI Tx. Therefore, this study could not completely guarantee the validity of the

study results for early or mild hypothyroidism. However, TSH would increase gradually from the day of stopping LT3 to the day of RAI Tx. In addition, in Supplemental Fig. S1, the WD-HR parameters of the hypothyroidism group started to decrease from the beginning of discontinuing LT3 (i.e., early or mild hypothyroidism). This result implies the applicability of WD-HR parameters to the prediction of early or mild hypothyroidism.

In conclusion, our results indicate that WD-HR parameters have excellent potential for accurately predicting the thyroid function status in patients with hypothyroidism and the serum free T4 levels even in well-controlled patients. These parameters can be measured easily and may be useful for the timely management of hypothyroidism. This study expands the spectrum of clinical applications of wearable-based biosignals to overall thyroid dysfunction, including hypothyroidism and thyrotoxicosis.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

AUTHOR CONTRIBUTIONS

Conception or design: J.H.M. Acquisition, analysis, or interpretation of data: K.H.K., J.L., C.H.A., H.W.Y., J.Y.C., H.Y.L., W.W.L., J.H.M. Drafting the work or revising: K.H.K., J.H.M. Final approval of the manuscript: K.H.K., J.L., C.H.A., H.W.Y., J.Y.C., H.Y.L., W.W.L., J.H.M.

ORCID

Ki-Hun Kim *https://orcid.org/0000-0001-9882-0389*
Jae Hoon Moon *https://orcid.org/0000-0001-6327-0575*

REFERENCES

1. Chaker L, Bianco AC, Jonklaas J, Peeters RP. Hypothyroidism. Lancet 2017;390:1550-62.
2. Taylor PN, Albrecht D, Scholz A, Gutierrez-Buey G, Lazarus JH, Dayan CM, et al. Global epidemiology of hyperthyroidism and hypothyroidism. Nat Rev Endocrinol 2018;14:301-16.
3. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. Arch Intern Med 2000;160:526-34.
4. Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). J Clin Endocrinol Metab 2002;87:489-99.
5. Braverman LE, Cooper DS, Kopp PA. Werner & Ingbar's the thyroid: a fundamental and clinical text. 11th ed. Philadelphia: Wolters Kluwer; 2021.
6. Carle A, Pedersen IB, Knudsen N, Perrild H, Ovesen L, Laurberg P. Hypothyroid symptoms and the likelihood of overt thyroid failure: a population-based case-control study. Eur J Endocrinol 2014;171:593-602.
7. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. JMIR Mhealth Uhealth 2017;5:e34.
8. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2017;2:104-6.
9. Al-Kaisey AM, Koshy AN, Ha FJ, Spencer R, Toner L, Sajeev JK, et al. Accuracy of wrist-worn heart rate monitors for rate control assessment in atrial fibrillation. Int J Cardiol 2020;300:161-4.
10. Lee JE, Lee DH, Oh TJ, Kim KM, Choi SH, Lim S, et al. Clinical feasibility of monitoring resting heart rate using a wearable activity tracker in patients with thyrotoxicosis: prospective longitudinal observational study. JMIR Mhealth Uhealth 2018;6:e159.
11. Moon JH, Steinhubl SR. Digital medicine in thyroidology: a new era of managing thyroid disease. Endocrinol Metab (Seoul) 2019;34:124-31.
12. Yi KH, Lee EK, Kang HC, Koh Y, Kim SW, Kim IJ, et al. 2016 Revised Korean Thyroid Association management guidelines for patients with thyroid nodules and thyroid cancer. Int J Thyroidol 2016;9:59-126.

13. Zulewski H, Muller B, Exer P, Miserez AR, Staub JJ. Estimation of tissue hypothyroidism by a new clinical score: evaluation of patients with various grades of hypothyroidism and controls. J Clin Endocrinol Metab 1997;82:771-6.

14. Lee JE, Lee DH, Oh TJ, Kim KM, Choi SH, Lim S, et al. Clinical feasibility of continuously monitored data for heart rate, physical activity, and sleeping by wearable activity trackers in patients with thyrotoxicosis: protocol for a prospective longitudinal observational study. JMIR Res Protoc 2018;7:e49.

15. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.

16. Yang W, Liao S. A study of quadratic inference functions with alternative weighting matrices. Commun Stat Simul Comput 2017;46:994-1007.

17. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics 2001;57:120-5.

18. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020;17: 261-72.

19. Burnham KP, Anderson DR. A practical information-theoretic approach. New York: Springer-Verlag New York; 2002.

20. Rahbari R, Zhang L, Kebebew E. Thyroid cancer gender disparity. Future Oncol 2010;6:1771-9.

21. Ryan SM, Goldberger AL, Pincus SM, Mietus J, Lipsitz LA. Gender- and age-related differences in heart rate dynamics: are women more complex than men? J Am Coll Cardiol 1994;24:1700-7.