MAYO CLINIC

Artificial Intelligence Application in Graves Disease: Atrial Fibrillation, Heart Failure and Menstrual Changes

Jwan A. Naser, MBBS; Zachi I. Attia, PhD; Sorin V. Pislaru, MD, PhD; Marius N. Stan, MD; Patricia A. Pellikka, MD; Peter A. Noseworthy, MD; Paul A. Friedman, MD; and Grace Lin, MD

Abstract

Objective: To study the utility of artificial intelligence (AI)–enabled electrocardiograms (ECGs) in patients with Graves disease (GD) in identifying patients at high risk of atrial fibrillation (AF) and heart failure with reduced ejection fraction (HFrEF), and to study whether AI-ECG can reflect hormonal changes and the resulting menstrual changes in GD.

Patients and Methods: Patients diagnosed with GD between January 1, 2009, and December 31, 2019, were included. We considered AF diagnosed at 30 days or fewer before or any time after GD and de novo HFrEF not explained by ischemia, valve disorder, or other cardiomyopathy at/after GD diagnosis. Electrocardiograms at/after index condition were excluded. A subset analysis included females younger than 45 years of age to study the association between ECG-derived female probability and menstrual changes (shorter, lighter, or newly irregular cycles).

Results: Among 430 patients (mean age, 50±17 years; 337 (78.4%) female), independent risk factors for AF included ECG probability of AF (hazard ratio [HR], 1.5; 95% CI, 1.2 to 1.6 per 10%; $P<.001$), older age (HR, 1.05; 95% CI, 1.03 to 1.07 per year; $P<.001$), and overt hyperthyroidism (HR, 3.9; 95% CI, 1.2 to 12.7; $P=.03$). The C-statistic was 0.85 for the combined model. Among 495 patients (mean age, 52±17 years; 374 (75.6%) female), independent risk factors for HFrEF were ECG probability of low ejection fraction (HR, 1.4; 95% CI, 1.1 to 1.6 per 10%; $P=.001$) and presence of AF (HR, 8.3; 95% CI, 2.2 to 30.9; $P=.002$), and a C-statistic of 0.89 for the combined model. Lastly, of 72 females younger than 45 years, 30 had menstrual changes at time of GD and had a significantly lower AI ECG–derived female probability [median 77.3; (IQR 57.9 to 94.4)% vs. median 97.7 (IQR 92.4 to 99.5)%, $P<.001$].

Conclusion: AI-enabled ECG identifies patients at risk for GD-related AF and HFrEF and was associated with menstrual changes in women with GD.

 ■ Mayo Clin Proc. 2022;97(4):730-737

From the Department of Internal Medicine (J.A.N.), Department of Cardiovascular Medicine (J.A.N., Z.I.A., S.V.P., P.A.P., P.A.N., P.A.F., G.L.), and the Department of Endocrinology and Metabolism (M.N.S.), Mayo Clinic, Rochester, MN.

The effects of hyperthyroidism on the cardiovascular system are widely recognized.[1] In particular, its association with atrial fibrillation (AF) is well-established and occurs in approximately 10% of patients.[2,3] Another known but less common consequence is heart failure with reduced ejection fraction (HFrEF) occurring in approximately 2% of patients.[4,5] Both conditions result in significant morbidity and mortality.[2,5] Previous studies have shown that persistence of hyperthyroidism is associated with worse cardiovascular outcomes.[6] Therefore, it is pertinent to identify patients at increased risk of developing AF and HFrEF who may benefit from closer surveillance and more prompt restoration of euthyroidism. In a large population undergoing clinically indicated electrocardiograms (ECGs), artificial intelligence (AI)–enabled ECGs using a convolutional neural network can identify the signature of silent AF in ECGs obtained while patients are in sinus rhythm as well as that of low ejection

fraction (EF) for each ECG performed with areas under curve (AUCs) of 0.87 and 0.93, respectively.[7,8] Whether the existing AI platforms can identify patients at highest risk of AF and HFrEF secondary to hyperthyroidism is unknown.

Hyperthyroidism is also known to cause menstrual changes and has been linked to infertility in women.[9] This can be due to the increased production of androstenedione and testosterone induced by the excess thyroid hormone.[9,10] However, screening for sex hormone levels in all women diagnosed with hyperthyroidism to identify patients more likely to have infertility is not cost-effective and not clinically indicated. Artificial intelligence–augmented ECG can identify the signature of female with an AUC of 0.97.[11] In this model, AI-ECG gives a probability number between 0 (0%) and 1 (100%) of being female.[11] A lower number implies high probability of being male and a higher number implies high probability of being a female.[11] Whether this inexpensive tool is reflective of hormonal levels and the resulting changes in menstrual cycle and fertility associated with hyperthyroidism remains to be investigated.

In this study, we aimed to (1) determine the ability of the AI-enabled ECG platform to identify patients who are highly susceptible to develop Graves disease (GD)–related AF and HFrEF, and (2) to assess the relation of GD and the associated menstrual changes with AI-ECG–derived female probability.

PATIENTS AND METHODS

Study Population

The study was approved by the Institutional Review Board. Only patients who had previously agreed to include their data in a retrospective chart review research study were included. Consecutive patients diagnosed with GD between January 1, 2009, and December 30, 2019, who had ECG data as detailed below were included. In the analyses involving menstrual changes, women who were pregnant or on hormonal contraceptives were excluded. Incidence of AF and HFrEF as well as baseline demographic data, comorbidities, laboratory data, ECGs, echocardiography data, and data on pregnancy status, hormonal contraceptives, and menstrual changes were extracted from the electronic medical records (EMRs).

Artificial Intelligence-Enabled ECGs

Artificial intelligence–enabled ECG data were extracted using the previously trained AI-platforms.[7,8,11] Probabilities of AF and low EF expressed as percentages were obtained from ECGs performed within 2 months before and up to 2 years after GD diagnosis; when multiple ECGs were present, the earliest was considered. Electrocardiograms performed at or after the diagnosis of AF or HFrEF, respectively, were excluded. In the case of AF, only ECGs in sinus rhythm were included. Moreover, ECGs that were used in the training of the initial algorithms were excluded.[7,8,11]

To compare the female probability derived from the AI-enabled ECGs (expressed as a percentage), three intervals relative to GD diagnosis were identified: (1) before GD, defined as more than 2 months before the formal diagnosis of GD; (2) during GD, defined as within 2 months before the diagnosis of GD and before attaining euthyroidism; and (3) after GD, defined as after attaining euthyroidism. Females younger than 45 years of age were identified to study the association between AI-enabled ECG female probability with menstrual changes. Electrocardiograms used for this particular analysis were obtained during GD.

Definitions

Graves disease was diagnosed based on biochemical thyrotoxicosis with diffuse thyroid uptake on I123 scan and/or positive thyroid-stimulating hormone receptor antibodies. When clinical concern for silent thyroiditis was present, the diagnosis was confirmed by thyroid ultrasound with Doppler flow. Restoration of euthyroidism was considered upon normalization of thyroid-stimulating hormone levels. Overt hyperthyroidism was defined as free T4 level greater than 1.7 ng/dL. Menstrual changes were defined as shorter-duration, lighter-

TABLE 1. Baseline Characteristics of Patients With Graves Disease With and Without Atrial Fibrillation[a,b]

	All GD patients N=430	GD patients with AF n=43	GD patients without AF n=387	*P*
Age at GD, years	51 (36-64)	66 (58-75)	50 (35-62)	<.001
Male	93 (21.6)	14 (32.6)	79 (20.4)	.07
BMI (SD), kg/m^2	27.7 (24.1-33.3)	29.8 (25.7-34.0)	27.4 (24.0-33.1)	.48
Hypertension	176 (40.9)	32 (74.4)	144 (37.2)	<.001
Diabetes mellitus	54 (12.6)	9 (20.9)	45 (11.6)	.08
Hyperlipidemia	153 (35.6)	26 (60.5)	127 (32.8)	<.001
History of CAD	24 (5.6)	9 (20.9)	15 (3.9)	<.001
History of stroke/TIA	13 (3.0)	3 (7.0)	10 (2.6)	.11
COPD	29 (6.7)	5 (11.6)	24 (6.2)	.18
Chronic kidney disease	31 (7.2)	8 (18.6)	23 (5.9)	.002
Anemia	66 (15.3)	10 (23.2)	56 (14.5)	.13
Ever smoker	199 (46.3)	18 (41.9)	181 (46.8)	.54
Free T4 at GD, ng/dL	2.9 (1.9-4.3)	2.8 (2.0-4.0)	2.9 (1.9-4.4)	.91
Total T3 at GD, pg/mL	238 (183-356)	218 (180-651)	240 (183-348)	.44
TRAb at GD, IU/L	7.51 (3.97-19.00)	5.19 (3.55-32.00)	7.74 (4.02-18.00)	.39
Overt hyperthyroidism	323 (83.0)	39 (92.8)	284 (81.8)	.07
Probability of AF, %	1.5 (0.5-4.3)	8.3 (2.0-31.0)	1.3 (0.4-3.4)	<.001
Probability of female, %	81.7 (27.5-98.0)	65.3 (14.4-89.2)	83.5 (30.8-98.3)	.11

[a]AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; GD, Graves disease; TIA, transient ischemic attack; TRAb, thyrotropin receptor antibodies.
[b]Values are presented as median (IQR) or number (percentage).

flow, or newly irregular menstrual cycles at or after the diagnosis of GD.

Atrial fibrillation diagnosis was identified by International Classification of Diseases codes in the EMR in patients with GD and validated through ECG data. Graves disease–related AF was defined as AF diagnosed up to 30 days before or any time after the diagnosis of GD. We excluded AF in the setting of surgery (single episode of AF in the postoperative period), valve disease (greater than or equal to moderate aortic stenosis, mitral stenosis, or primary mitral regurgitation), cardiomyopathy, or other structural heart disease. We did not differentiate between atrial fibrillation and atrial flutter.

Heart failure (HF) diagnosis was identified by International Classification of Diseases codes in the EMR and confirmed using the modified Framingham criteria.[12] Graves disease–related HFrEF was defined as patients who presented with a new episode of congestive HF associated with left ventricular EF of less than 50% that was not explained by ischemia, valve disorder, or other cardiomyopathy at or after GD diagnosis.

Study Endpoints

The primary endpoints of the study were the development of GD-related AF and HFrEF. A secondary endpoint was menstrual changes related to GD.

Statistical Analysis

Continuous variables are expressed as median and interquartile range (IQR). Categorical variables are expressed as frequency and percentage. Differences between groups were tested using the Pearson χ^2 test or two-sample Student *t* test, as appropriate. Analysis for risk factors associated with the development of GD-related AF and HFrEF was performed using the proportional hazards method. For multivariate analysis, all variables with a *P* less than .15 at univariate analysis were entered into the model, and a backward stepwise selection method was

Mayo Clin Proc. ■ April 2022;97(4):730-737 ■ https://doi.org/10.1016/j.mayocp.2021.08.029
www.mayoclinicproceedings.org

used to generate the final model. The number of variables was constrained according to the number of events to avoid overfitting. The concordance statistic (C-statistic) was used to estimate the AUC in timed analyses. Comparison between non-nested models (models including AI-enhanced ECG versus other models) was evaluated through concordance comparison; first, the variance of the difference between C-statistics was calculated. The test statistic (normal z-score) was then computed and the two-sided *P* value calculated. The matched-pairs *t* test was used to compare probability of being a female obtained from AI-enabled ECG; patients acted as their own controls. A subset analysis was performed on females younger than 45 years of age at GD diagnosis to study the association between female probability during GD and menstrual changes. Statistical analysis was performed using the JMP Pro software version 14.1.0 (SAS Institute Inc, Cary, NC) and BlueSky Statistics software v. 7.10 (BlueSky Statistics LLC, Chicago, IL). Statistical significance was accepted for *P* less than .05.

TABLE 2. Risk Factors Associated With Graves Disease-Related AF[a]

Univariate analysis for AF	Hazard ratio	95% CI	*P*
Age at GD, years	1.07 per year	1.04-1.09	<.001
Male	1.91	1.01-3.62	.047
BMI, kg/m^2	1.01 per unit	0.97-1.04	.73
Hypertension	4.41	2.22-8.75	<.001
Diabetes mellitus	1.80	0.86-3.75	.12
Hyperlipidemia	2.75	1.49-5.08	.001
History of CAD	5.04	2.41-10.51	<.001
History of stroke/TIA	2.82	0.87-9.14	.08
COPD	2.00	0.79-5.08	.15
Chronic kidney disease	2.94	1.36-6.33	.006
Anemia	1.62	0.80-3.28	.18
Ever smoker	0.82	0.45-1.51	.53
Free T4 at GD, ng/dL	1.03 per unit	0.90-1.16	.64
Total T3 at GD, pg/mL	1.02 per 10 units	0.99-1.06	.16
TRAb at GD, IU/L	1.02 per unit	0.99-1.05	.17
Overt hyperthyroidism	2.62	0.81-8.49	.11
Probability of AF, %	1.04 per 1%	1.03-1.05	<.001
Probability of female, %	0.99 per 1%	0.99-1.00	.08

[a]AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; GD, Graves disease; TIA, transient ischemic attack; TRAb, thyrotropin receptor antibodies.

RESULTS

Risk of GD-Related AF

A total of 430 patients met the inclusion criteria for risk factor analysis of GD-related AF. Baseline characteristics, comorbidities, and thyroid hormone levels are shown in Table 1. Graves disease–related AF was diagnosed in 43 (10.0%) patients with a median duration of 11 (IQR, 0 to 863) days after GD diagnosis. Electrocardiograms used were obtained 27 (IQR, 4 to 690) days before the diagnosis of AF.

Univariate risk factors are shown in Table 2. The C-statistic for AI-ECG–derived probability of AF was 0.79 (95% CI, 0.72 to 0.86). At multivariate analysis, risk factors for AF were AI-ECG–derived probability of AF (hazard ratio [HR], 1.5; 95% CI, 1.2 to 1.6 per 10%; *P*<.001), older age (HR, 1.05; 95% CI, 1.03 to 1.07 per year; *P*<.001), and overt hyperthyroidism (HR, 3.9; 95% CI, 1.2 to 12.7; *P*=.03). The model had a C-statistic of 0.85 (95% CI, 0.80-0.90). On the other hand, when analyzing conventional risk factors alone, older age (HR, 1.07; 95% CI, 1.05 to 1.09; *P*<.001), male sex (HR, 2.1; 95% CI, 1.1 to 4.1; *P*=.02), and overt hyperthyroidism (HR 4.2; 95% CI, 1.3 to 13.8; *P*=0.02) were independent risk factors with a C-statistic of 0.79 (95% CI, 0.73 to 0.85). The model performed significantly lower than the one with the AI-enabled ECG data (*P*=.001).

Upon dividing AI-enabled ECG-derived AF probability into four quartiles based on the duration between ECG and AF diagnosis (<4 days, 4 to 27 days, 27 to 690 days, and >690 days), AF probability increased with closer duration between ECG and AF diagnosis. Median AI-derived AF probability was 24.4 (IQR, 4.8 to 50.9), 19.0 (IQR, 3.0 to 37.8), 7.6 (IQR, 1.7 to 43.2), and 3.6 (IQR, 0.6 to 10.4) for the four quartiles, respectively (Supplemental Table 1, available online at http://www.mayoclinicproceedings.org).

Risk of GD-Related HFrEF

A total of 495 patients met the inclusion criteria for risk factor analysis of GD-related HFrEF. Baseline characteristics,

TABLE 3. Baseline Characteristics of Patients With Graves Disease With and Without HFrEF[a,b]

	All GD patients N=495	GD patients with HFrEF n= 12	GD patients without HFrEF n=483	*P*
Age at GD, years	53 (38-66)	60 (45-71)	53 (37-66)	.26
Male	121 (24.4)	3 (25.0)	118 (24.4)	.96
BMI, kg/m^2	28.3 (24.3-33.8)	34.3 (26.2-37.0)	28.2 (24.2-33.6)	.26
Hypertension	212 (42.8)	7 (58.3)	205 (42.4)	.27
Diabetes mellitus	62 (12.5)	1 (8.3)	61 (12.6)	.66
Hyperlipidemia	179 (36.1)	3 (25.0)	176 (36.4)	.42
History of CAD	29 (5.9)	1 (8.3)	28 (6.0)	.71
History of stroke/TIA	15 (3.0)	1 (8.3)	14 (2.9)	.28
COPD	33 (6.7)	1 (8.3)	32 (6.6)	.81
Chronic kidney disease	34 (6.9)	3 (25.0)	31 (6.4)	.01
Anemia	71 (14.3)	4 (33.3)	67 (13.9)	.06
Ever smoker	230 (46.5)	3 (25.0)	227 (47.0)	.13
Free T4 at GD, ng/dL	2.9 (1.9-4.4)	4.1 (2.1-9.3)	2.9 (1.9-4.3)	.16
TRAb at GD, IU/L	7.86 (4.05-18.00)	33.00 (4.56-40.00)	7.56 (4.02-17.17)	.06
Overt hyperthyroidism	372 (83.4)	11 (91.6)	361 (83.2)	.44
Presence of AF (previous or new)	117 (23.6)	9 (75.0)	108 (22.4)	<.001
Probability of low EF, %	1.04 (0.5-2.8)	4.6 (2.6-48.7)	1.02 (0.5-2.6)	<.001
Probability of female, %	76.6 (20.8-97.2)	61.2 (19.1-87.8)	77.5 (20.8-97.4)	.45

[a]AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; GD, Graves disease; TIA, transient ischemic attack; TRAb, thyrotropin receptor antibodies.
[b]Values are presented as median (IQR) or number (percentage).

comorbidities, and thyroid hormone levels are shown in Table 3. Heart failure with reduced EF was diagnosed in 12 (2.4%) patients with a median duration of 69 (IQR, 0 to 919) days after GD diagnosis. Median EF was 39% (IQR 34% to 47%). Electrocardiograms used were obtained 69 (IQR, 5 to 856) days before HFrEF diagnosis.

Univariate risk factors are shown in Table 4. The C-statistic for AI-ECG–derived probability of low EF was 0.85 (95% CI, 0.79 to 0.91). Risk factors at multivariate analysis had an AI-ECG probability of low EF (HR, 1.4; 95% CI, 1.1 to 1.6 per 10%; *P*=.001) and presence of AF (HR, 8.3; 95% CI, 2.2 to 30.9; *P*=.002). The model had a C-statistic of 0.89 (95% CI, 0.84 to 0.94). When analyzing conventional risk factors alone, thyrotropin receptor antibody (TRAb) (HR, 1.07; 95% CI, 1.02 to 1.13; *P*=.004) was the only risk factor in the final model with a C-statistic of 0.66 (95% CI, 0.37 to 0.95), which performed lower than the model using AI-ECG–derived data, although it did not achieve statistical significance (*P*=.14).

Artificial Intelligence ECG-Derived Female Probability

A total of 215 patients (mean age at GD of 59±16 years, 154 [71.6%] female) with ECGs both before and during GD were identified. Artificial intelligence ECG-derived female probability decreased with GD diagnosis in women (mean 84.8% vs 76.8%, *P*<.001). On the other hand, 199 patients (mean age at GD of 58±16 years, 137 [68.8%] female) with GD had ECGs during and after GD. Artificial intelligence ECG-derived female probability increased after attaining euthyroidism in women (mean 75.8% during GD vs 82.8% after GD, *P*=.005). Neither similar nor opposite changes were noted in men (Figure).

Mayo Clin Proc. ■ April 2022;97(4):730-737 ■ https://doi.org/10.1016/j.mayocp.2021.08.029
www.mayoclinicproceedings.org

To analyze the association of AI-ECG–derived female probability with menstrual changes, 72 females younger than 45 years of age with ECGs during GD were identified. Of these, 30 (41.7%) developed menstrual changes at time of GD diagnosis and had a significantly lower AI-ECG–derived female probability from ECGs obtained during GD when compared with females with no menstrual changes (median 77.3% (IQR 57.9% to 94.4%) versus median 97.7% (IQR 92.4% to 99.5%), $P<.001$).

TABLE 4. Risk Factors Associated With Graves Disease-Related HFrEF[a]

Univariate analysis for HFrEF	Hazard ratio	95% CI	*P*
Age at GD, years	1.02 per year	0.99-1.06	.23
Male	1.07	0.29-3.95	.92
BMI, kg/m^2	1.04 per unit	0.99-1.09	.08
Hypertension	1.87	0.59-5.89	.29
Diabetes mellitus	0.61	0.08-4.74	.64
Hyperlipidemia	0.56	0.15-2.08	.39
History of CAD	1.47	0.19-11.42	.71
History of stroke/TIA	3.23	0.42-25.15	.26
COPD	1.36	0.18-10.57	.77
Chronic kidney disease	4.30	1.61-15.91	.03
Anemia	2.88	0.87-9.59	.08
Ever smoker	0.38	0.10-1.42	.15
Free T4, ng/dL	1.23 per unit	1.02-1.46	.02
TRAb, IU/L	1.07 per unit	1.02-1.13	.004
Overt hyperthyroidism	2.10	0.27-16.30	.48
Presence of AF (previous or new)	10.02	2.71-37.02	.001
Probability of low EF, %	1.04 per 1%	1.02-1.05	<.001
Probability of female, %	0.99	0.98-1.01	.46

[a]AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; GD, Graves disease; TIA, transient ischemic attack; TRAb, thyrotropin receptor antibodies.

DISCUSSION

To the best of our knowledge, this is the first study to investigate the application of ECG data derived from machine learning platforms in identifying patients at highest risk of AF and HFrEF secondary to GD. Our major finding was that AI-ECG–derived probabilities of AF and low EF were strongly associated with developing GD-related AF and HFrEF, respectively, at follow-up. Furthermore, we found that the AI-enabled ECG signature of female sex in women was altered by GD and was associated with menstrual changes.

Graves Disease–Related AF and HFrEF

Excess thyroid hormone levels have direct toxic effects on the heart altering energy production, intracellular metabolism, and contractile function with subsequent myocardial damage.[1,13] This can manifest as left ventricular hypertrophy, diastolic dysfunction, cardiac chamber enlargement, HF, and arrhythmias — most commonly AF.[1,13] Conventional risk factors associated with higher occurrence of AF in our study were similar to those reported previously and included older age, male sex, hypertension, hyperlipidemia, and history of coronary artery disease, chronic kidney disease, and overt hyperthyroidism.[3,14] Similarly, conventional risk factors for HFrEF associated with GD included chronic kidney disease, higher free T4 levels, higher TRAb levels, and the presence of AF, consistent with previous reports.[4] The AI-enabled ECG platforms were able to identify patients at risk of AF and HFrEF with C-statistics of 0.79 and 0.85, respectively. Furthermore, incorporation of data from the AI-enabled ECG platforms significantly enhanced the ability of the clinical models to identify patients at risk of GD-related AF with a C-statistic of 0.85. As for HFrEF, AI-enabled ECG did improve the C-statistic from 0.66 to 0.89; however, this did not reach statistical significance, likely due to the underpowered nature of the analysis. One possible explanation is that the structural changes caused by the excess thyroid hormone can cause ECG patterns not recognized by the human eye that machine learning can detect even before clinical AF or HF results. These AI-enabled ECGs have been proven effective in identifying patients from the general population who might have silent AF during sinus rhythm ECGs as well as in identifying patients with, or at increased risk of, left ventricular systolic dysfunction.[7,8] The low-EF model was further validated prospectively and in special populations of patients presenting to the emergency department,

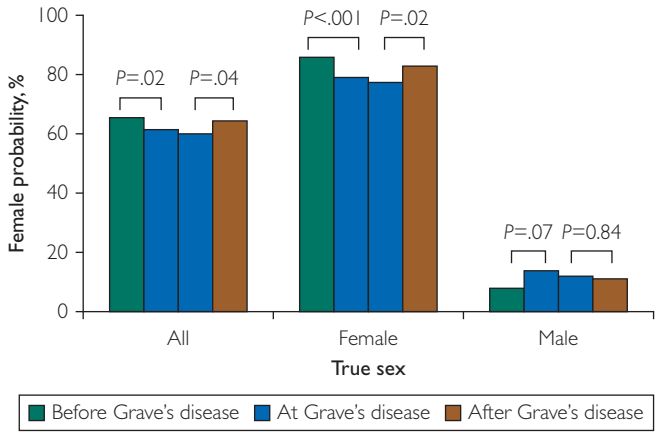

FIGURE. Changes in artificial intelligence (AI) electrocardiogram (ECG)–derived female probability with the diagnosis of Graves disease (GD) and upon restoration of euthyroidism. Artificial intelligence ECG-derived female probability decreases upon GD diagnosis in women and increases again after restoration of euthyroidism. Similar or opposite changes are not observed in men, presumably due to low power of the analysis. The matched pairs *t* test was used for analysis and so only patients having ECGs in both time intervals were included in each of the two analyses; hence, probability at GD differed slightly for the two analyses.

admitted to the cardiac intensive care unit, and those with coronavirus disease 2019 infection.[15-18] Given the negative impact of GD-related AF and HFrEF on outcomes,[2,5] it is important to identify patients at increased risk and follow them closely. Moreover, it may be pertinent to treat hyperthyroidism in these patients more promptly given that longer duration of thyrotoxicosis increases risk of cardiovascular morbidity.[6] Our study further showed that upon dividing duration between ECG and AF diagnosis into four quartiles, the closer ECGs to AF diagnosis had higher AF probability. Therefore, one might argue that serial ECGs could be helpful in patients with GD; however, the frequency of monitoring is yet to be determined in future studies.

AI-Enabled ECG Female Probability in GD

Hyperthyroidism increases the production rate of androstenedione and testosterone, which also leads to increased estrone and estradiol levels through peripheral conversion.[10] Our findings showed a decreased probability of female derived from AI-enabled ECG in women upon the diagnosis of GD that increased again with the restoration of euthyroidism. Previous studies reported possible effects of sex on the ECG that are possibly caused by hormonal differences.[19,20] Furthermore, AI-augmented ECG has been proven effective in identifying the signature of female sex with an AUC reaching 0.97.[11] However, both the current and the previous studies did not assess for hormonal levels given their retrospective nature. Therefore, it remains a hypothesis that changes in the AI-ECG–derived female probability are reflective of changes in the hormonal milieu induced by the excess thyroid hormones. At the same time, our finding that women who had menstrual changes as part of their presentation with GD were identified by AI-ECG as having significantly lower ECG-derived female probability may further support the hypothesis. Analysis of changes in AI-ECG–derived female probability with the diagnosis of GD in men was likely underpowered and larger studies incorporating hormonal levels are needed to confirm these findings and to assess the association between AI-ECG–derived female probability and the actual hormonal levels.

Study Limitations

Our study is best understood in the setting of its limitations. First, given the retrospective nature of the study, patients included were a subset of GD patients who had ECGs around the diagnosis of GD likely in the setting of another clinical indication, which could limit the generalizability of the study to GD patients who did not have ECGs. Second, given the rarity of the disease, our study included only a small number of GD-related HFrEF cases, which could have affected the power of the analysis. Finally, we did not have data on hormonal levels that we could correlate with changes in AI-ECG–derived female probability upon the diagnosis and resolution of hyperthyroidism. However, the derived female probability was associated with the clinical manifestation of menstrual changes that could potentially be reflective of hormonal levels. Future, larger-scale, prospective studies are needed to confirm the

association of AI-ECG–derived female probability with sexual hormones levels.

CONCLUSION

Artificial intelligence–enabled ECG platforms performed well in identifying patients at highest risk of GD-related AF and HFrEF. These patients may benefit from closer follow-up and more prompt treatment of hyperthyroidism. Moreover, the AI-ECG–derived female probability was reflective of menstrual changes in GD. Future studies to confirm its role in screening for infertility in women with GD are needed.

SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at http://www.mayoclinicproceedings.org. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Abbreviations and Acronyms: **AF,** atrial fibrillation; **AI,** artificial intelligence; **AUC,** area under curve; **ECG,** electrocardiogram; **EF,** ejection fraction; **EMR,** electronic medical record; **GD,** Graves disease; **HFrEF,** heart failure with reduced ejection fraction; **HR,** hazard ratio; **IQR,** interquartile range; **TRAb,** thyrotropin receptor antibody

Potential Competing Interests: The authors report no potential competing interests.

Correspondence: Address to Grace Lin, MD, Division of Cardiovascular Diseases, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905 (Lin.Grace@mayo.edu; Twitter: @Grace_Lin_MD).

ORCID

Jwan A. Naser: https://orcid.org/0000-0002-6210-7669; Marius N. Stan: https://orcid.org/0000-0002-1408-8125; Patricia A. Pellikka: https://orcid.org/0000-0001-6800-3521; Paul A. Friedman: https://orcid.org/0000-0001-5052-2948

REFERENCES

1. Osuna PM, Udovcic M, Sharma MD. Hyperthyroidism and the heart. *Methodist Debakey Cardiovasc J*. 2017;13(2):60-63.
2. Naser JA, Pislaru S, Stan MN, Lin G. Incidence, risk factors, natural history, and outcomes of atrial fibrillation in patients with Graves' disease. *Circulation*. 2020;142(suppl 3):A14539.
3. Frost L, Vestergaard P, Mosekilde L. Hyperthyroidism and risk of atrial fibrillation or flutter — a population-based study. *Arch Intern Med*. 2004;164(15):1675-1678.
4. Siu CW, Yeung CY, Lau CP, Kung AWC, Tse HF. Incidence, clinical characteristics and outcome of congestive heart failure as the initial presentation in patients with primary hyperthyroidism. *Heart*. 2007;93(4):483-487.
5. Naser JA, Pislaru S, Stan MN, Lin G. Incidence, risk factors, and outcomes of heart failure in patients with Graves' disease. *Circulation*. 2020;142 (abstr).
6. Lillevang-Johansen M, Abrahamsen B, Jorgensen HL, Brix TH, Hegedus L. Duration of hyperthyroidism and lack of sufficient treatment are associated with increased cardiovascular risk. *Thyroid*. 2019;29(3):332-340.
7. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25(1):70-74.
8. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet*. 2019;394(10201):861-867.
9. Unuane D, Tournaye H, Velkeniers B, Poppe K. Endocrine disorders & female infertility. *Best Pract Res Clin Endocrinol Metab*. 2011;25(6):861-873.
10. Southren AL, Olivo J, Gordon GG, Vittek J, Brener J, Rafii F. The conversion of androgens to estrogens in hyperthyroidism. *J Clin Endocrinol Metab*. 1974;38(2):207-214.
11. Attia ZI, Friedman PA, Noseworthy PA, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ Arrhythm Electrophysiol*. 2019;12(9):e007284.
12. Senni M, Tribouilloy CM, Rodeheffer RJ, et al. Congestive heart failure in the community — a study of all incident cases in Olmsted County, Minnesota, in 1991. *Circulation*. 1998;98(21):2282-2289.
13. Vargas-Uricoechea H, Bonelo-Perdomo A, Sierra-Torres CH. Effects of thyroid hormones on the heart. *Clin Investig Arterioscler*. 2014;26(6):296-309.
14. Selmer C, Olesen JB, Hansen ML, et al. The spectrum of thyroid disease and risk of new onset atrial fibrillation: a large population cohort study. *BMJ*. 2012;345:e7895.
15. Attia ZI, Kapa S, Noseworthy PA, Lopez-Jimenez F, Friedman PA. Artificial intelligence ECG to detect left ventricular dysfunction in COVID-19: a case series. *Mayo Clin Proc*. 2020;95(11):2464-2466.
16. Jentzer JC, Kashou AH, Attia ZI, et al. Left ventricular systolic dysfunction identification using artificial intelligence-augmented electrocardiogram in cardiac intensive care unit patients. *Int J Cardiol*. 2021;326:114-123.
17. Adedinsewo D, Carter RE, Attia Z, et al. Artificial intelligence-enabled ECG algorithm to identify patients with left ventricular systolic dysfunction presenting to the emergency department with dyspnea. *Circ Arrhythm Electrophysiol*. 2020;13(8):e008437.
18. Attia ZI, Kapa S, Yao X, et al. Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction. *J Cardiovasc Electrophysiol*. 2019;30(5):668-674.
19. van Noord C, Dorr M, Sturkenboom MCJM, et al. The association of serum testosterone levels and ventricular repolarization. *Eur J Epidemiol*. 2010;25(1):21-28.
20. Malik M, Hnatkova K, Kowalski D, Keirns JJ, van Gelderen EM. QT/RR curvatures in healthy subjects: sex differences and covariates. *Am J Physiol Heart Circ Physiol*. 2013;305(12):H1798-H1806.