The Journal of Clinical Endocrinology & Metabolism, 2022, **107**, 3222–3230
https://doi.org/10.1210/clinem/dgac544
Advance access publication 19 September 2022
Clinical Research Article

Machine Learning–Based Prediction of Elevated PTH Levels Among the US General Population

Hajime Kato,[1,2] Yoshitomo Hoshino,[1,2] Naoko Hidaka,[1,2] Nobuaki Ito,[1,2] Noriko Makita,[1,2] Masaomi Nangaku,[1] and Kosuke Inoue[3]

[1]Division of Nephrology and Endocrinology, The University of Tokyo Hospital, Tokyo 113-8655, Japan
[2]Osteoporosis Center, The University of Tokyo Hospital, Tokyo 113-8655, Japan
[3]Department of Social Epidemiology, Graduate School of Medicine, Kyoto University, Kyoto 604-8146, Japan

Correspondence: Kosuke Inoue, MD, PhD, Department of Social Epidemiology, Graduate School of Medicine, Kyoto University, Fl 2, Science Frontier Laboratory, Yoshida-konoe-cho, Sakyo-ku, Kyoto, Kyoto 604-8146, Japan. Email: inoue.kosuke.2j@kyoto-u.ac.jp.

Abstract

Context: Although elevated parathyroid hormone (PTH) levels are associated with higher mortality risks, the evidence is limited as to when PTH is expected to be elevated and thus should be measured among the general population.

Objective: This work aimed to build a machine learning–based prediction model of elevated PTH levels based on demographic, lifestyle, and biochemical data among US adults.

Methods: This population-based study included adults aged 20 years or older with a measurement of serum intact PTH from the National Health and Nutrition Examination Survey (NHANES) 2003 to 2006. We used the NHANES 2003 to 2004 cohort (n = 4096) to train 6 machine-learning prediction models (logistic regression with and without splines, lasso regression, random forest, gradient-boosting machines [GBMs], and SuperLearner). Then, we used the NHANES 2005 to 2006 cohort (n = 4112) to evaluate the model performance including area under the receiver operating characteristic curve (AUC).

Results: Of 8208 US adults, 753 (9.2%) showed PTH greater than 74 pg/mL. Across 6 algorithms, the highest AUC was observed among random forest (AUC [95% CI] = 0.79 [0.76-0.81]), GBM (AUC [95% CI] = 0.78 [0.75-0.81]), and SuperLearner (AUC [95% CI] = 0.79 [0.76-0.81]). The AUC improved from 0.69 to 0.77 when we added cubic splines for the estimated glomerular filtration rate (eGFR) in the logistic regression models. Logistic regression models with splines showed the best calibration performance (calibration slope [95% CI] = 0.96 [0.86-1.06]), while other algorithms were less calibrated. Among all covariates included, eGFR was the most important predictor of the random forest model and GBM.

Conclusion: In this nationally representative data in the United States, we developed a prediction model that potentially helps us to make accurate and early detection of elevated PTH in general clinical practice. Future studies are warranted to assess whether this prediction tool for elevated PTH would improve adverse health outcomes.

Key Words: parathyroid hormone, hyperparathyroidism, machine learning, prediction model, NHANES

Abbreviations: γGTP, γ-glutamyl transferase; 25(OH)D, 25-hydroxyvitamin D; AUC, area under the receiver operating characteristic curve; BMI, body mass index; eGFR, estimated glomerular filtration rate; GBM, gradient-boosting machine; LDH, lactate dehydrogenase; NHANES, National Health and Nutrition Examination Survey; PTH, parathyroid hormone.

Parathyroid hormone (PTH) is the main regulator of calcium homeostasis. Elevation of PTH levels was observed particularly among patients with primary hyperparathyroidism and secondary hyperparathyroidism due to vitamin D deficiency and chronic kidney diseases with inappropriately treated calcium and phosphate level (1, 2). While the causal relationship between PTH and long-term health outcomes has not been fully elucidated, some previous studies reported the association between elevated PTH and all-cause or cardiovascular mortality even among the general population (3–8). Given these findings and other common manifestations of primary hyperparathyroidism (eg, osteoporosis, vertebral fractures, hypercalciuria, nephrolithiasis) (2), the early and judicious detection of elevated PTH is important in general clinical settings. However, a previous study showed the fact that PTH was evaluated in only 55% of the patients who were likely to have primary hyperparathyroidism (9). Particularly, primary normocalcemic hyperparathyroidism, a mild phenotype of primary hyperparathyroidism, would be more difficult to be suspected in general clinical practice. Furthermore, as vitamin D deficiency—a common condition with a prevalence rate of 5.9% among US adults—sometimes lead to osteoporosis due to secondary hyperparathyroidism (10, 11), prediction of elevated PTH is also useful for the diagnosis of such secondary hyperparathyroidism due to vitamin D deficiency. In this context, it is imperative to develop a high-performance prediction model for elevated PTH from standard biochemical information to assist the effective screening and avoid the underdetection of this endocrine disorder among the general population.

Over the last decade, in line with the availability of big data and rapidly increasing computational capacity, machine-

Received: 1 June 2022. **Editorial Decision:** 12 September 2022. **Corrected and Typeset:** 29 September 2022

learning algorithms have attracted substantial attention because of their high performance to predict health outcomes (12, 13). Although a few previous studies in the United States applied and built prediction models for primary hyperparathyroidism (14, 15), they did not include comprehensive data (eg, socioeconomic status and standard biochemical data including liver function and electrolytes) in their prediction model. Moreover, other diseases that might be underdiagnosed in the general population but impair health status, including primary normocalcemic hyperparathyroidism and secondary hyperparathyroidism due to vitamin D deficiency, were not included in their model, limiting their utility to answer the clinically important question: Who should be targeted for PTH evaluation in general clinical settings?

To address this knowledge gap, using a nationally representative database of US adults along with several machine-learning algorithms, we developed several machine learning–based models to predict elevated PTH levels from commonly available demographic information and biomarkers. We then compared the performance of each model, and assessed variables that largely contribute to the prediction of elevated PTH levels.

Materials and Methods

Study Design and Setting

The US National Health and Nutrition Examination Survey (NHANES) is a large-scale, multistage, nationally representative survey of the civilian, noninstitutionalized population in the United States conducted by the National Center for Health Statistics. Data from US adults aged 20 years or older who participated in the 2 cycles (2003 to 2004 and 2005 to 2006) of NHANES were used in this study. Structured interview data and physical examination results, including laboratory data of blood and urine samples, are collected continuously and released in 2-year cycles. The detailed design and participants of the NHANES cohort are described elsewhere (16). All participants gave their written informed consent and the approval to participate in NHANES study protocols as per the research ethics review board of the National Center for Health Statistics (17). This study followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement (18) (Supplemental Checklist) (19).

Study Samples

There was a total of 8948 participants aged 20 years or older at enrollment for whom serum PTH level was available. We excluded individuals who lacked data on education status (n = 12), marital status (n = 2), body mass index (BMI) (n = 157), and poverty-income ratio (n = 432). We additionally excluded people with missing data on serum calcium levels (n = 29), serum phosphate levels (n = 2), glycated hemoglobin A_{1c} (n = 28), aspartate aminotransferase (n = 65), total protein (n = 10), and lactate dehydrogenase (LDH) (n = 2). The final analytical cohort contained 8208 participants.

Predictors and Outcome

The variables for the prediction models were selected from the NHANES data. Because the present study focused on building the prediction model of elevated PTH from information obtained in general practice, we included demographic and lifestyle data (age, sex, race/ethnicity, poverty-income ratio, education status, marital status, smoking status, comorbidities, prescription medications, and BMI) and standard biochemical data as predictors. Information on demographic and lifestyle characteristics was collected at the survey enrollment. Participants who smoked at least 100 cigarettes during their lifetime were categorized as smokers, with former smokers defined as individuals who smoked at least 100 cigarettes and not currently smoking. As comorbidities, we selected self-reported information on the physician diagnoses of diabetes mellitus, cardiovascular disease, and cancer. The use of antihypertensives and statins was also self-reported. Measured weights and heights were used to calculate BMI.

Standard biochemical profile including albumin, total protein, cholesterol, triglycerides, total bilirubin, alanine aminotransferase, aspartate transaminase, alkaline phosphatase, γ-glutamyl transferase (γGTP), LDH, uric acid, blood urea nitrogen, sodium, potassium, chloride, calcium, phosphate, and creatinine levels were measured by Beckman Synchron LX20. We adjusted calcium levels for hypoalbumia as previously reported (20). The estimated glomerular filtration rate (eGFR) was calculated by the Chronic Kidney Disease Epidemiology Collaboration equation (21). Glycated hemoglobin A_{1c} was measured by high-performance liquid chromatography. Serum 25-hydroxyvitamin D (25(OH)D) levels were measured using the DiaSorin radioimmunoassay (RIA) kit. 25(OH)D values were converted to liquid chromatography–tandem mass spectrometry equivalent values (ng/mL) as previously reported (8, 22).

Serum intact PTH was assayed by an electrochemiluminescence immunoassay on the Elecys 1010 autoanalyzer (Roche Diagnostics). We defined elevated PTH levels as PTH greater than 74 pg/mL (23).

Statistical Analysis

Based on the TRIPOD statement (18), we split the data into the training data set and test data set using different periods (ie, training data set, NHANES 2003-2004 [n = 4096]; test data set, NHANES 2005-2006 [n = 4112]). Using the training set, we developed the conventional prediction model and 4 machine-learning models to predict the probability of elevated PTH levels. First, as the conventional model, we fit a logistic regression model including demographic, lifestyle, and biochemical parameters. These parameters were age, sex (male or female), race/ethnicity (non-Hispanic White, non-Hispanic Black, Mexican American, or others), poverty-income ratio, education status (< 9th grade, 9th-11th grade, high school, or general education degree, or > high school), marital status (married or not), smoking status (never, former, or current), prior history of diabetes mellitus, prior history of cardiovascular disease, prior history of cancer, antihypertensive prescription, statin prescription, and standard biochemical markers. Given the possible nonlinear relationship of elevated PTH with eGFR, we also built the model adding cubic splines for eGFR.

Using the predictors listed earlier, the following 4 machine-learning prediction models were also constructed: logistic regression with lasso regularization (lasso regression) (24), random forest (25), gradient-boosting machines (GBM) (26), and SuperLearner (27). Briefly, lasso regression enhances standard regression models by enabling us to select important predictors (feature selection). Both random forest and GBM are an ensemble of decision trees. The random forest combines

Table 1. Demographic characteristics according to serum parathyroid hormone levels in the National Health and Nutrition Examination Survey 2003 to 2006

	PTH > 74 pg/mL	PTH ≤ 74 pg/mL
Total No.	753	7455
PTH, pg/mL		
Mean (SD)	104.67 (43.90)	39.63 (14.28)
Median (IQR)	91 (80-108)	38 (29-49)
Age, y	58.98 (19.41)	48.01 (18.63)
Male sex, n (%)	357 (47.4)	3591 (48.2)
Race/ethnicity, n (%)		
Non-Hispanic White	304 (40.4)	3952 (53.0)
Non-Hispanic Black	250 (33.2)	1457 (19.5)
Mexican American	144 (19.1)	1517 (20.3)
Others	55 (7.3)	529 (7.1)
Poverty-income ratio	2.37 (1.49)	2.66 (1.61)
Education status, n (%)		
< 9th grade	140 (18.6)	924 (12.4)
9th-11th grade	119 (15.8)	1098 (14.7)
High school or GED	177 (23.5)	1811 (24.3)
> High school	317 (42.1)	3622 (48.6)
Marital status, n (%)		
Married	357 (47.4)	4166 (55.9)
Not married	396 (52.6)	3289 (44.1)
Smoking, n (%)		
Never	432 (57.4)	3781 (50.7)
Former	227 (30.1)	1941 (26.0)
Current	94 (12.5)	1733 (23.2)
Diabetes, n (%)	107 (14.2)	714 (9.6)
Cardiovascular disease, n (%)	205 (27.2)	743 (10.0)
Cancer, n (%)	85 (11.3)	612 (8.2)
Antihypertensive prescription, n (%)	377 (50.1)	1930 (25.9)
Statin prescription, n (%)	169 (22.4)	941 (12.6)
BMI	30.69 (8.52)	28.41 (6.21)
eGFR, mL/min/1.73 m^2	75.2 (30.18)	98.64 (24.52)
Albumin, g/dL	4.10 (0.38)	4.17 (0.40)
Total protein, g/dL	7.18 (0.51)	7.14 (0.51)
HbA_{1c}, %	5.69 (0.94)	5.57 (1.00)
Cholesterol, mg/dL	197.93 (48.48)	203.16 (43.53)
Triglycerides, mg/dL	144.50 (99.62)	146.84 (118.35)
Total bilirubin, mg/dL	0.77 (0.70)	0.73 (0.29)
Alanine aminotransferase, U/L	23.31 (15.05)	25.64 (28.24)
Aspartate aminotransferase, U/L	25.46 (11.11)	25.75 (24.53)
Alkaline phosphatase, U/L	61.02 (30.30)	56.69 (24.58)
γ-Glutamyl transferase, U/L	61.24 (117.4)	58.44 (103.61)
Lactate dehydrogenase, U/L	140.19 (31.27)	128.20 (31.36)
Uric acid, mg/dL	6.10 (1.71)	5.27 (1.38)
Blood urea nitrogen, mg/dL	16.75 (11.33)	12.30 (5.14)
Serum sodium, mmol/L	139.58 (2.54)	138.89 (2.26)
Serum potassium, mmol/L	4.02 (0.40)	3.97 (0.33)
Serum chloride, mmol/L	104.04 (3.36)	103.60 (2.70)
Serum albumin-adjusted calcium, mg/dL	9.53 (0.47)	9.59 (0.33)
Serum phosphate, mg/dL	3.69 (0.62)	3.82 (0.54)
25-Hydroxyvitamin D, ng/mL	19.60 (7.80)	25.24 (8.80)

Abbreviations: BMI, body mass index; eGFR, estimated glomerular filtration rate; HbA_{1c}, glycated hemoglobin A_{1c}; GED, General Educational Development; IQR, interquartile range; PTH, parathyroid hormone.

outputs provided by the decision trees randomly generated from predictors, while GBM is an additive model of decision trees estimated by gradient descent. SuperLearner is an ensemble of machine-learning algorithms that creates an optimal weight for the initial set of candidate models (ie, logistic regression, lasso regression, random forest, and GBM). The parameters of each algorithm were tuned using 10-fold cross-validation.

In the test set, the prediction performance of each model is evaluated by computing the area under the receiver operating characteristic curve and prospective prediction results (ie, sensitivity, specificity, positive predictive value, and negative predictive value). To calculate these values under the class imbalance in the outcome, we chose the threshold of prediction results using the Youden index (28). Calibration was evaluated using calibration intercepts and slope for each model (29). Given the possibility that 2 major mineral and bone metabolism biomarkers—serum calcium and phosphate—may not be frequently measured in general clinical practice, we also built the prediction model without information on serum albumin-adjusted calcium and phosphate.

We conducted the following 4 additional analyses. First, we evaluated the diagnostic performance of the prediction model to distinguish patients with biochemical profiles close to primary hyperparathyroidism (ie, elevated PTH with high or high-normal serum calcium level). The cutoff for high or high-normal calcium levels was greater than 9.6 mg/dL, as previously reported (8). Second, although we did not include 25(OH)D in our primary prediction model as it is not commonly included in standard biochemical examination, we rebuilt the prediction model of elevated PTH including 25(OH)D as an additional predictor. Third, because PTH is less likely to be measured among individuals without advanced chronic kidney disease, we built the prediction model of elevated PTH among those with eGFR greater than or equal to 60 mL/min/1.73 m^2 (n = 7492). Last, we rebuilt the prediction model of elevated PTH including urine albumin to creatinine ratio as an additional predictor among individuals with available urine data (n = 5445). All analyses were performed using the R version 4.2.1.

Results

The baseline characteristics of the study participants are shown in Table 1. Compared to adults with PTH levels less than or equal to 74 pg/mL, those with elevated PTH levels were older, non-Hispanic Black, individuals with lower income levels, less educated, and unmarried. They were also more likely to have comorbidities such as diabetes, cardiovascular disease, and cancer, and take antihypertensive and statin prescriptions. When comparing the biochemical data among these groups, participants with elevated PTH levels had lower

Table 2. Predictive ability of the logistic regression model, tree-based algorithms, and SuperLearner for elevated parathyroid hormone levels

	AUC	Sensitivity	Specificity	PPV	NPV
Model without calcium and phosphate					
Logistic regression	0.69 (0.66-0.72)	0.69	0.61	0.14	0.95
Logistic regression + spline models	0.77 (0.74-0.80)	0.69	0.72	0.19	0.96
Lasso regression	0.73 (0.70-0.76)	0.67	0.69	0.17	0.96
Random forest	0.79 (0.76-0.81)	0.62	0.80	0.22	0.96
Gradient boosting	0.78 (0.75-0.81)	0.67	0.77	0.21	0.96
SuperLearner	0.79 (0.76-0.81)	0.72	0.70	0.18	0.96
Model without calcium and phosphate					
Logistic regression	0.67 (0.64-0.71)	0.55	0.73	0.16	0.95
Logistic regression + spline models	0.76 (0.73-0.79)	0.65	0.75	0.19	0.96
Lasso regression	0.72 (0.69-0.75)	0.53	0.80	0.20	0.95
Random forest	0.76 (0.74-0.79)	0.67	0.69	0.17	0.96
Gradient boosting	0.78 (0.75-0.80)	0.67	0.72	0.18	0.96
SuperLearner	0.75 (0.72-0.78)	0.64	0.75	0.19	0.96

The prediction model included age, sex, race/ethnicity, poverty-income ratio, education status, marital status, smoking status, prior history of diabetes mellitus, prior history of cardiovascular disease, prior history of cancer, antihypertensive prescription, statin prescription, and standard biochemical markers. PPVs were generally low for all algorithms because of the small number of outcomes overall.
Abbreviations: AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

levels of eGFR, serum phosphate, and cholesterol and higher levels of alkaline phosphatase, γGTP, LDH, uric acid, blood urea nitrogen, and 25(OH)D than others with low or normal PTH levels. Decreasing trends in PTH levels according to increased eGFR (< 30, 30 to < 60, 60 to < 90, ≥ 90) are shown in Supplementary Table S1 (19).

Prediction of Hyperparathyroidism With Predictors

Overall, 753 participants (9.2% of 8208 participants) had elevated levels of PTH. Across 6 algorithms, the highest predictive performance was shown in random forest (AUC [95% CI] = 0.79 [0.76-0.81], sensitivity = 0.62, specificity = 0.80), GBM (AUC [95% CI] = 0.78 [0.75-0.81], sensitivity = 0.67, specificity = 0.77), and SuperLearner (AUC [95% CI] = 0.79 [0.76-0.81], sensitivity = 0.72, specificity = 0.70) (Table 2 and Fig. 1). While the logistic regression showed the lowest predictive performance (AUC [95% CI] = 0.69 [0.66-0.72], sensitivity = 0.69, specificity = 0.61), the performance was substantially improved when we added the cubic splines for eGFR in the logistic regression model (AUC [95% CI] = 0.77 [0.74-0.80], sensitivity = 0.69, specificity = 0.72). These patterns were consistently observed when we excluded serum calcium and phosphate levels from the model while AUC was slightly lower than that in the model with serum calcium and phosphate levels (eg, logistic regression with splines, AUC [95% CI] = 0.76 [0.73-0.79]; random forest, AUC [95% CI] = 0.76 [0.74-0.79]; GBM, AUC [95% CI] = 0.78 [0.75-0.80]; SuperLearner, AUC [95% CI] = 0.75 [0.72-0.78]) (see Table 2 and Fig. 2). Among the 6x algorithms, logistic regression models with splines showed the best calibration performance (calibration slope [95% CI] = 0.96 [0.86-1.06]; Fig. 3).

Fig. 4 shows the importance of variables in the random forest model and GBM. Renal function (eGFR) was the most important predictor in both models. In the random forest model, age, calcium, and other laboratory data including blood urea nitrogen level, alkaline phosphatase, and uric acid levels were frequently used to build the algorithm. Likewise, serum calcium, phosphate, sodium, uric acid levels, and BMI were frequently used to build the GBM algorithm.

Additional Analyses

We also found high predictive performance for elevated PTH levels with serum calcium levels greater than 9.6 mg/dL (eg, logistic regression with splines, AUC [95% CI] = 0.76 [0.71-0.80]) (Table 3). When we added 25(OH)D as a predictor, we found the improvement of predictive performance (eg, AUC [95% CI] = 0.82 [0.80-0.84] for logistic regression with splines, GBM, and SuperLearner) (Supplementary Table S2) (19). AUC was around 0.73 when we restricted individuals to those with eGFR greater than or equal to 60 mL/min/1.73 m^2 (Supplementary Table S3) (19) and when we additionally included urine albumin to creatinine ratio in the prediction model among individuals with urine data (Supplementary Table S4) (19). The predictive performance

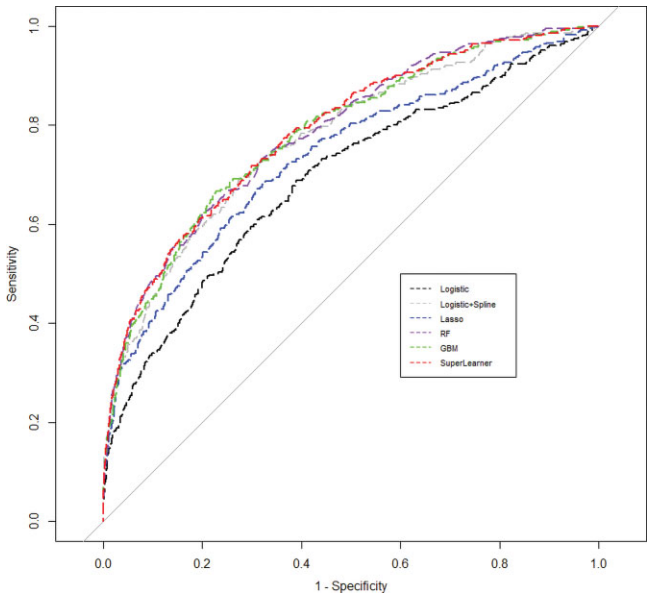

Figure 1. Receiver operating characteristic curve of the logistic regression model, tree-based algorithms, and SuperLearner to predict elevated parathyroid hormone levels. GBM, gradient-boosting machines; Lasso, logistic regression with lasso regularization, Logistic, logistic regression model; RF, random forest.

was not different between logistic regression with and without splines in these analyses.

Discussion

In this analysis of the total of 8025 participants from a national population-based survey data, we applied several machine-learning approaches (ie, lasso regression, random forest, GBM, and SuperLearner) to differentiate participants with elevated PTH from those with normal or low PTH levels. Among these algorithms, the random forest model, GBM, and SuperLearner achieved the highest predictive performance in discrimination using demographical data and clinical data, such as serum electrolytes including calcium and phosphate, liver function, and lipid profile. However, these machine-learning algorithms showed poor calibration performance. Instead, logistic regression models with spline models achieved high performance of both discrimination and calibration. The predictive performance for elevated PTH levels remained high even without calcium and phosphate data—2 major mineral and bone metabolism biomarkers that are closely related to PTH.

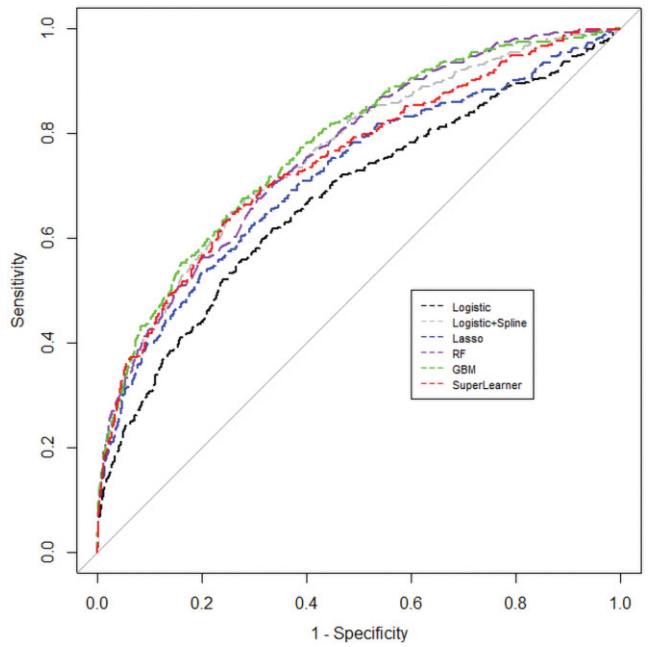

Figure 2. Receiver operating characteristic curve of the logistic regression model, tree-based algorithms, and SuperLearner to predict elevated parathyroid hormone levels without information on serum calcium and phosphorus levels. GBM, gradient-boosted machines; Lasso, logistic regression with lasso regularization, Logistic, logistic regression model; RF, random forest.

To the best of our knowledge, this is the first study that has applied several models including machine-learning approaches to predict elevated PTH levels among the US general population. Our results suggest that we can predict elevated PTH from general information even without serum calcium, phosphate, and vitamin D level, which are not routinely evaluated in the general clinical setting. A few previous studies reported that machine learning–based prediction models effectively differentiate patients with primary hyperparathyroidism and those without primary hyperparathyroidism. For instance, Somnay et al (14) developed a Bayesian network model using age, sex, and serum calcium, phosphate, PTH, vitamin D, and creatinine levels and reported C statistics of 0.989 among 6777 patients with surgically treated primary hyperparathyroidism and 5053 controls. However, the utility of this prediction model might be limited to specific circumstances, because their cohort was based on 3 high-volume endocrine surgery programs. Moreover, because almost half of their patients with primary hyperparathyroidism showed elevated calcium and PTH levels, the authors' findings cannot be extended to patients with primary normocalcemic hyperparathyroidism, which is not a rare condition (0.18% to 8.9% of the general population) and potentially underdiagnosed among the general population (30, 31). Greer et al (15) also reported the prediction model of primary hyperparathyroidism with an accuracy of 0.86. Their findings based on hospital electronic health record data would also not be extended to the general population or general clinical settings given the unique feature of the university hospital specializing in endocrine disorders.

In both the random forest model and GBM, we observed that GFR was the most important variable to predict elevated PTH among the general population. Impaired renal function is one of the key factors leading to elevated PTH levels (32). In our cohort, of 753 individuals with hyperparathyroidism, 102 individuals (14%) exhibited eGFR less than 40; the condition under which more than 50% of the patients present with hyperparathyroidism (32). Because not a few individuals would have impaired renal function associated with secondary hyperparathyroidism, eGFR is critical information to consider whether the individuals have elevated PTH. The importance of eGFR to predict elevated PTH was also supported by the fact that the difference in predictive ability between logistic regression models with and without splines of eGFR diminished when we built the model among people without impaired renal function or when we additionally included urine albumin to creatinine ratio in the model.

In addition, our prediction model showed that age and other biochemical predictors (blood urea nitrogen, and uric acid)

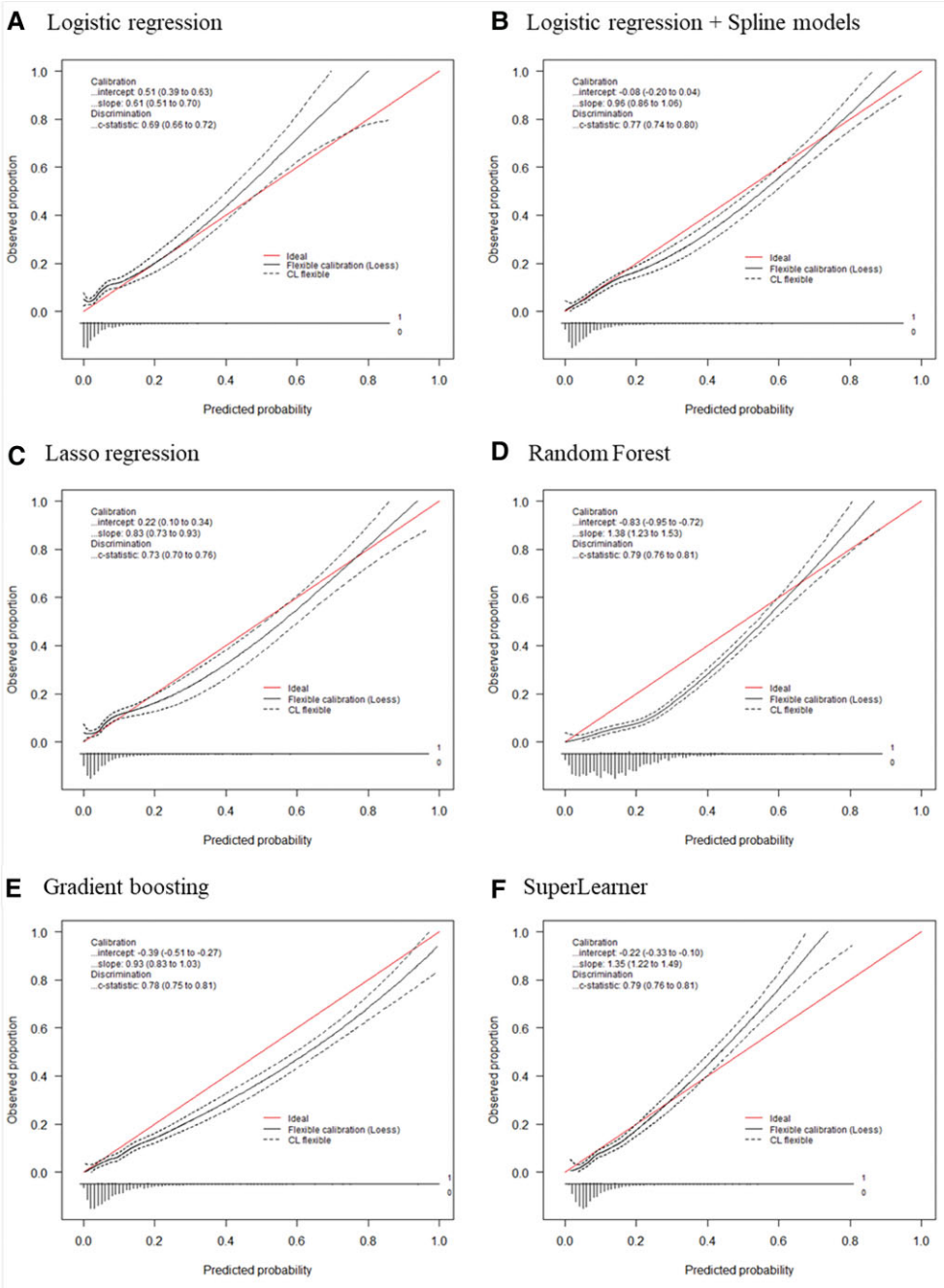

Figure 3. Calibration plots of the logistic regression model, tree-based algorithms, and SuperLearner for elevated parathyroid hormone levels.

were other important variables to predict elevated PTH as well as major mineral and bone metabolism biomarkers (ie, serum calcium and phosphate levels). PTH levels are known to increase with age (33–35). According to previous studies, the age-related increase in PTH levels might be induced by a fall in renal function and an age-related decrease in calcium absorption possibly due to low vitamin D levels (36, 37). Furthermore, a previous study using the NHANES 2003 to 2006 observed that hyperuricemia suppressed 1-αhydroxylase leading to higher PTH (38). While there was no report about the association between blood urea nitrogen and PTH levels, blood urea nitrogen might contribute to predicting elevated PTH levels by indirectly reflecting renal function or dehydration resulting from nephrogenic diabetes insipidus due to hypercalcemia induced by primary hyperparathyroidism.

Our study has several limitations. First, PTH was measured at one time during the survey and thus may not reflect patients' chronic status. There is a circadian rhythm and seasonal variation in PTH levels (39, 40), and a single measurement of PTH levels might be affected by circadian rhythm. However, because seasonal variation in PTH levels might be inversely associated with the seasonal variation of vitamin D levels (40), the prediction model indirectly reflected the vitamin D status by other predictors and thus partly compensated for the seasonal variation in PTH levels. Second, we included only participants with PTH measured, and thus cannot rule out the possibility of selection bias. Third, because the purpose of our study

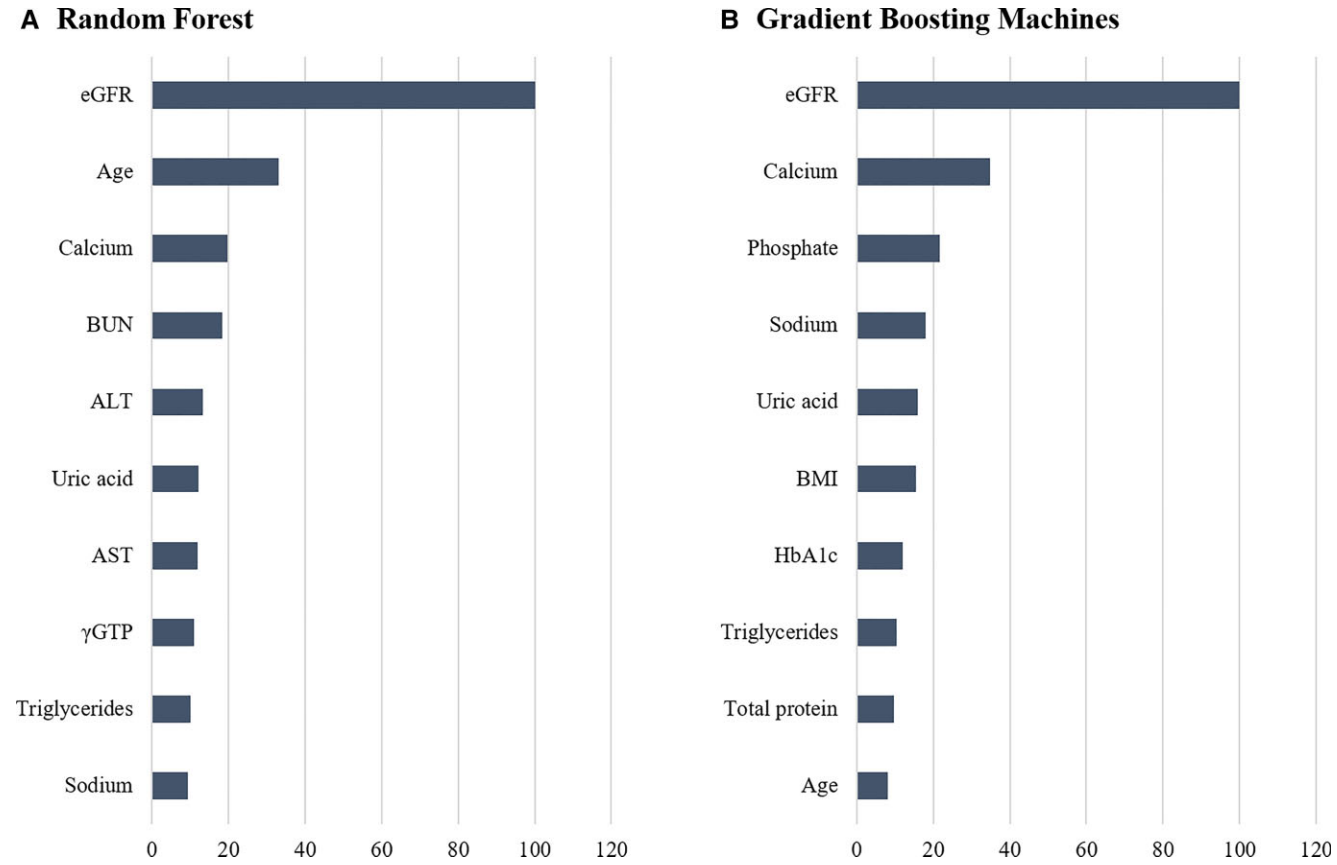

Figure 4. Variable importance of each predictor in the random forest and gradient-boosting machine algorithms. The variable importance is a measure scaled to have a maximum value of 100. A, random forest; B, gradient-boosting machines.

was to build the prediction model of elevated PTH from demographic and clinical data, we did not employ the NHANES survey weight, which is generally recommended to use to produce nationally representative descriptive statistics. Thus, our findings may suffer from sampling bias and have limited generalizability. Fourth, while we split the NHANES data into training and test data, we did not evaluate the validity of our prediction models in external data. Last, we did not have detailed information to differentiate the pathology of hyperparathyroidism (eg, primary hyperparathyroidism and secondary hyperparathyroidism due to vitamin D deficiency or chronic kidney disease). However, the diagnosis of the etiology of hyperparathyroidism was based on a multimodal approach including biochemical data, ultrasonography, and scintigraphy, and this is outside of the scope of our study.

In conclusion, among US adults, we found that the application of flexible models including machine-learning approaches has the potential to improve the discriminative ability for elevated PTH levels from generally available demographic, lifestyle, and biochemical data. Among all algorithms, logistic regression with splines showed better calibration performance than other machine-learning algorithms. These prediction models, if well discriminated and calibrated, would improve the medical management of hyperparathyroidism (including primary hyperparathyroidism and secondary hyperparathyroidism due to vitamin D deficiency) by helping clinicians to evaluate PTH levels leading to the early diagnosis and management of this endocrine disorder. Future investigations are needed to validate our findings and assess whether using prediction models of elevated PTH in clinical practice reduces long-term adverse health outcomes among the general population.

Table 3. Predictive ability of the logistic regression model, tree-based algorithms, and SuperLearner for elevated parathyroid hormone levels with serum calcium levels greater than 9.6 mg/dL

	AUC	Sensitivity	Specificity	PPV	NPV
Logistic regression	0.73 (0.68-0.78)	0.61	0.71	0.11	0.97
Logistic regression + spline models	0.76 (0.71-0.80)	0.61	0.71	0.11	0.97
Lasso regression	0.72 (0.67-0.77)	0.61	0.74	0.06	0.99
Random forest	0.73 (0.67-0.78)	0.68	0.66	0.11	0.97
Gradient boosting	0.71 (0.66-0.77)	0.67	0.70	0.06	0.99
SuperLearner	0.74 (0.68-0.79)	0.66	0.67	0.11	0.97

The prediction model included age, sex, race/ethnicity, poverty-income ratio, education status, marital status, smoking status, prior history of diabetes mellitus, prior history of cardiovascular disease, prior history of cancer, antihypertensive prescription, statin prescription, and standard biochemical markers without serum albumin-adjusted calcium and phosphate levels. PPVs were generally low for all algorithms because of the small number of outcomes overall.
Abbreviations: AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

Funding

KI was supported by the Japan Society for the Promotion of Science (21K20900 and 22K17392), the Japan Endocrine Society, Meiji Yasuda Life Foundation of Health and Welfare, and the Program for the Development of Next-generation Leading Scientists with Global Insight (L-INSIGHT) sponsored by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the article for publication.

Disclosures

The authors declare no potential conflict of interest related to the subject matter of the paper.

Data Availability

Original data generated and analyzed during this study are included in this published article or the data repositories listed in "References."

References

1. Fraser WD. Hyperparathyroidism. *Lancet*. 2009;374(9684):145-158.
2. Walker MD, Silverberg SJ. Primary hyperparathyroidism. *Nat Rev Endocrinol*. 2018;14(2):115-125.
3. Anderson JL, Vanwoerkom RC, Horne BD, *et al*. Parathyroid hormone, vitamin D, renal dysfunction, and cardiovascular disease: dependent or independent risk factors? *Am Heart J*. 2011;162(2):331-339.e2.
4. Kritchevsky SB, Tooze JA, Neiberg RH, *et al*. Health ABC Study. 25-Hydroxyvitamin D, parathyroid hormone, and mortality in black and white older adults: the Health ABC study. *J Clin Endocrinol Metab*. 2012;97(11):4156-4165.
5. Domiciano DS, Machado LG, Lopes JB, *et al*. Bone mineral density and parathyroid hormone as independent risk factors for mortality in community-dwelling older adults: a population-based prospective cohort study in Brazil. The São Paulo Ageing & Health (SPAH) Study. *J Bone Miner Res*. 2016;31(6):1146-1157.
6. Saleh FN, Schirmer H, Sundsfjord J, Jorde R. Parathyroid hormone and left ventricular hypertrophy. *Eur Heart J*. 2003;24(22):2054-2060.
7. Rashid G, Bernheim J, Green J, Benchetrit S. Parathyroid hormone stimulates endothelial expression of atherosclerotic parameters through protein kinase pathways. *Am J Physiol Renal Physiol*. 2007;292(4):1215-1218.
8. Kato H, Ito N, Makita N, Nangaku M, Leung AM, Inoue K. Association of serum parathyroid hormone levels with all-cause and cause-specific mortality among U.S. adults. *Endocr Pract*. 2022;28(1):70-76.
9. Press DM, Siperstein AE, Berber E, *et al*. The prevalence of undiagnosed and unrecognized primary hyperparathyroidism: a population-based analysis from the electronic medical record. *Surgery*. 2013;154(6):1232-1238.
10. Schleicher RL, Sternberg MR, Looker AC, *et al*. National estimates of serum total 25-hydroxyvitamin D and metabolite concentrations measured by liquid chromatography-tandem mass spectrometry in the US population during 2007-2010. *J Nutr*. 2016;146(5):1051-1061.
11. Cashman KD. Vitamin D deficiency: defining, prevalence, causes, and strategies of addressing. *Calcif Tissue Int*. 2020;106(1):14-29.
12. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375(13):1216-1219.
13. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-2509.
14. Somnay YR, Craven M, McCoy KL, *et al*. Improving diagnostic recognition of primary hyperparathyroidism with machine learning. *Surgery*. 2017;161(4):1113-1121.
15. Greer ML, Davis K, Stack BC. Machine learning can identify patients at risk of hyperparathyroidism without known calcium and intact parathyroid hormone. *Head Neck*. 2022;44(4):817-822.
16. NHANES—National Health and Nutrition Examination Survey Homepage. Accessed June 1, 2022. https://www.cdc.gov/nchs/nhanes/index.htm.
17. NHANES—NCHS Research Ethics Review Board Approval. Accessed June 1, 2022. https://www.cdc.gov/nchs/nhanes/irba98.htm.
18. Moons KG, Altman DG, Reitsma JB, *et al*. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-73.
19. Kato H, Hoshino Y, Hidaka N, *et al*. Supplementary data for "Machine Learning-Based Prediction of Elevated PTH Levels Among the US General Population." Accessed September 5, 2022. https://datadryad.org/stash/share/_jggEk3vlCh96sjT8fyuGqLI8MeViSlaYb3fAbpoBFg.
20. Payne RB, Little AJ, Williams RB, Milner JR. Interpretation of serum calcium in patients with abnormal serum proteins. *Br Med J*. 1973;4(5893):643-646.
21. Levey AS, Stevens LA, Schmid CH, *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604-612.
22. NHANES—Analytical Note for 25-Hydroxyvitamin D Data Analysis. https://wwwn.cdc.gov/nchs/nhanes/vitamind/analyticalnote.aspx
23. Center for Health Statistics N. Laboratory Procedure Manual Parathyroid Hormone ECL/Origen-Electrochemiluminescent Important Information for Users.
24. CRAN—Package glmnet. https://cran.r-project.org/web/packages/glmnet/index.html
25. CRAN—Package. https://cran.r-project.org/web/packages/ranger/index.html
26. CRAN—Package xgboost. https://cran.r-project.org/web/packages/xgboost/index.html
27. van der Laan MJ, Polley EC, Hubbard AE. Super Learner. *Stat Appl Genet Mol Biol*. 2007;6:Article25.
28. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. *Epidemiology*. 2005;16(1):73-81.
29. Van Calster B, McLernon DJ, van Smeden M, *et al*. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230.
30. Cusano NE, Cipriani C, Bilezikian JP. Management of normocalcemic primary hyperparathyroidism. *Best Pract Res Clin Endocrinol Metab*. 2018;32(6):837-845.
31. Schini M, Jacques RM, Oakes E, Peel NFA, Walsh JS, Eastell R. Normocalcemic hyperparathyroidism: study of its prevalence and natural history. *J Clin Endocrinol Metab*. 2020;105(4):E1171-E1186.
32. Isakova T, Wahl P, Vargas GS, *et al*. Fibroblast growth factor 23 is elevated before parathyroid hormone and phosphate in chronic kidney disease. *Kidney Int*. 2011;79(12):1370-1378.
33. Aloia JF, Feuerman M, Yeh JK. Reference range for serum parathyroid hormone. *Endocr Pract*. 2006;12(2):137-144.
34. Björkman M, Sorva A, Tilvis R. Responses of parathyroid hormone to vitamin D supplementation: a systematic review of clinical trials. *Arch Gerontol Geriatr*. 2009;48(2):160-166.
35. Paik JM, Farwell WR, Taylor EN. Demographic, dietary, and serum factors and parathyroid hormone in the National Health and Nutrition Examination Survey. *Osteoporos Int*. 2012;23(6):1727-1736.
36. Marcus R, Madvig P, Young G. Age-related changes in parathyroid hormone and parathyroid hormone action in normal humans. *J Clin Endocrinol Metab*. 1984;58(2):223-230.

37. Eastell R, Yergey AL, Vieira NE, Cedel SL, Kumar R, Riggs BL. Interrelationship among vitamin D metabolism, true calcium absorption, parathyroid function, and age in women: evidence of an age-related intestinal resistance to 1,25-dihydroxyvitamin D action. *J Bone Miner Res*. 1991;6(2):125-132.
38. Chen W, Roncal-Jimenez C, Lanaspa M, *et al*. Uric acid suppresses 1 alpha hydroxylase in vitro and in vivo. *Metabolism*. 2014;63(1):150-160.
39. Calvo MS, Eastell R, Offord KP, Bergstralh EJ, Burritt MF. Circadian variation in ionized calcium and intact parathyroid hormone: evidence for sex differences in calcium homeostasis. *J Clin Endocrinol Metab*. 1991;72(1):69-76.
40. Pasco JA, Henry MJ, Kotowicz MA, *et al*. Seasonal periodicity of serum vitamin D and parathyroid hormone, bone resorption, and fractures: the Geelong Osteoporosis Study. *J Bone Miner Res*. 2004;19(5):752-758.