



Illuminating the Path to Thyroid Disorder Management Using Artificial Intelligence: A Narrative Review

Farnaz Atighi ¹, Parsa Yazdanpanahi ^{1,*}, Alireza Keshtkar ², Alireza Karimi ¹, Arzhang Naseri ³, Mohammad Hossein Dabbaghmanesh ^{3,**}

¹ Student Research Committee, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran

² Research Center for Noncommunicable Diseases, Jahrom University of Medical Sciences, Jahrom, Iran

³ Endocrinology and Metabolism Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

*Corresponding Author: Student Research Committee, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran. Email: yazdanpanahiparsa@gmail.com

**Corresponding Author: Endocrinology and Metabolism Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. Email: dabbaghm@sums.ac.ir

Received: 17 July, 2024; Revised: 20 October, 2024; Accepted: 17 November, 2024

Abstract

Background: Dysregulation of thyroid function, manifesting as hyperthyroidism or hypothyroidism, can profoundly impact an individual's overall health. In this context, artificial intelligence (AI) applications have the potential to revolutionize diagnostic approaches, treatment strategies, and patient monitoring.

Objectives: This study comprehensively reviews the latest literature on AI applications in thyroid functional and autoimmune disorders.

Methods: An online search was conducted on databases using search queries crafted with MeSH terms related to AI and thyroid disorders. After screening, studies aligned with our research focus were selected for this narrative review.

Results: Multiple studies have explored the use of AI technologies, including machine learning (ML) and deep learning (DL), to enhance laboratory workflows for thyroid function tests (TFT) and improve the accuracy of TFT interpretation by incorporating clinical data. In imaging, DL-based models have demonstrated the potential to assist less experienced radiologists in interpreting scintigraphy and ultrasound images. Artificial intelligence has also provided valuable insights into identifying diagnostic genes for thyroid-related autoimmune disorders and understanding the effects of environmental factors, such as chemicals, on thyroid gland function. Some ML models have been developed to predict the risk of hypothyroidism following radioiodine therapy (RAI). Furthermore, AI has shown promise in personalized levothyroxine dose adjustments, predicting treatment responses, and accurately diagnosing complications such as thyroid-associated ophthalmopathy (TAO). Finally, ML-based models forecasting the risk of suicide attempts in patients with major depressive disorder (MDD) and predicting pregnancy outcomes, such as gestational diabetes mellitus (GDM) and preterm delivery, based on TFT results, appear beneficial in addressing these significant health issues.

Conclusions: The current state of AI in diagnosing and treating thyroid function disorders is promising, with applications primarily focused on improving diagnostic accuracy, consistency, and personalized treatment approaches. However, challenges remain that prevent these models from fully substituting professionals. Addressing these challenges is crucial to ensure AI effectively contributes to the management of patients with thyroid diseases.

Keywords: Thyroid Functional Disease, Hypothyroidism, Hyperthyroidism, Thyroiditis, Artificial Intelligence, Machine Learning, Deep Learning, Thyroid Function Test, Thyroid-Associated Ophthalmopathy, Personalized Medicine

1. Context

The thyroid gland is the maestro of the endocrine system, orchestrating the body's metabolism through its hormonal "batons." Located at the front of the neck, this butterfly-shaped gland primarily produces

triiodothyronine (T3) and thyroxine (T4), two hormones that impact all tissues. Additionally, it secretes calcitonin, a hormone responsible for maintaining calcium homeostasis. The thyroid's function is regulated by the hypothalamic-pituitary-thyroid axis, where the hypothalamus releases thyrotropin-releasing hormone

(TRH), and the pituitary gland secretes thyroid-stimulating hormone (TSH) to stimulate thyroid activity (1). Disruptions in the regulation, production, or secretion of these hormones can lead to adverse physiological effects. The most common thyroid function disorders are hyperthyroidism and hypothyroidism. Hyperthyroidism is caused by the overproduction of thyroid hormones, while hypothyroidism results from insufficient hormone production. Numerous factors contribute to these conditions, affecting a significant proportion of the global population (2, 3).

Integrating artificial intelligence (AI) into various medical fields holds the potential to revolutionize the diagnosis, monitoring, and management of thyroid function disorders (Figure 1). The initial attempts at computer-assisted diagnosis (CAD) in this domain date back to 1963, when probability matrices were employed to differentiate between the diverse etiologies of thyroid dysfunction using clinical and laboratory data (4). With the continuous evolution of AI technologies, there is an increasing capacity to improve healthcare resource efficiency, support clinical decision-making, provide personalized recommendations and treatments, and enhance understanding of complex thyroid pathologies by leveraging vast datasets (5). In this review, we aim to present a comprehensive overview of various AI applications in addressing thyroid function disorders, rather than focusing on a detailed examination of a single area.

2. Evidence Acquisition

We crafted search queries by combining a diverse array of MeSH terms: “artificial intelligence,” “machine learning,” “deep learning,” “thyroid gland,” “thyroid hormones,” “thyroid diseases,” “hyperthyroidism,” “hypothyroidism,” “thyroiditis,” “Graves’ disease,” “diagnosis,” and “thyroid function test.” These queries were used to conduct an extensive search in PubMed and Google Scholar online databases to collect full-text available human original studies, systematic reviews, and meta-analyses on the applications of AI in thyroid function disorders, published in English from January 2014 to May 2024. Various types of studies were included in this review to deepen the understanding of multiple aspects of the topic and study methodologies. Studies published in the last decade were included because the application of AI in the field of thyroid diseases has recently gained significant attention, and we aimed to gather the most up-to-date literature in this field. Our search covered titles and abstracts to ensure a broad exploration of relevant literature. After screening the

titles and abstracts, a number of studies aligned with our research focus were selected to provide this narrative review.

3. Results

Physicians may utilize an approach involving patient history, physical examination, laboratory testing, imaging studies, and biopsies to diagnose thyroid disorders. Diagnosing thyroid function diseases is challenging due to the broad spectrum of non-specific clinical presentations, such as weight changes, exhaustion, memory and cognitive issues, muscular pain, etc. Furthermore, it is essential to note that some patients, for instance elderly people, may not always display evident clinical symptoms (6, 7).

Thyroid function tests (TFT) are frequently utilized to diagnose thyroid function disorders, comprising blood tests for TSH, total and free T3 and T4, and T3 resin uptake (T3RU). In certain instances, anti-thyroid antibodies like thyroid peroxidase antibodies (TPO-Ab) and TSH receptor antibodies (TR-Ab) may also be considered. Thyroid-stimulating hormone level assessment through immunoassay is the most prevalent initial screening method for thyroid dysfunction (8). However, it is susceptible to various confounding factors and may not always precisely reflect the actual thyroid gland function status (9). Artificial intelligence can offer new opportunities to enhance accuracy, efficiency, and patient care quality within thyroid laboratory testing.

3.1. Thyroid Function Test and Routine Clinical Laboratory Data

First, AI can improve TFT immunoassay designs and mitigate the impact of confounding variables through in-silico modeling (10). Utilizing machine learning (ML) technologies is promising for transforming molecular identification methods for diagnosing thyroid function disorders. For instance, the deployment of a least squares support vector machine (LS-SVM) trained by biomolecular features extracted from serum TSH infrared spectroscopy data demonstrated a resilient capability to accurately assess a broad spectrum of blood TSH levels (11). Furthermore, by analyzing medical records and clinical datasets, ML presents opportunities to customize reference ranges and diagnostic thresholds for TFTs according to physiological variations in hormone levels. A study has indicated that ML has the potential to enhance the automated verification process of TFT results (12).

Second, AI can aid in diagnosing thyroid function disorders by integrating overall health conditions, clinical evaluations, and the interpretation of TFT

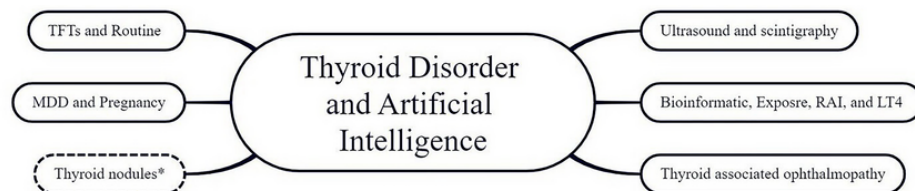


Figure 1. Key artificial intelligence (AI)-powered applications in thyroid disorders. The "*" symbol indicates that the application of AI in the management of thyroid nodules is not explored in the current study, although it is addressed in a separate survey conducted by the authors (5).

results. With advancements in natural language processing (NLP) technologies, a study assessed the proficiency of ChatGPT and Google Bard in interpreting TFT results. The findings revealed that the precision and reliability of these chatbots fall short of the standards required for clinical application, confirming that they are not yet suitable substitutes for professional medical consultation (13).

Numerous studies have investigated ML and deep learning (DL)-based predictive and interpretative models for classifying thyroid dysfunctions. These models have been trained on various exclusive or publicly available thyroid disease datasets, primarily those from the University of California Irvine (UCI) repository (14-20). Generally, ML models have demonstrated superior performance compared to DL models, primarily due to the limitations in the size and features of datasets used for training DL models (15). Conducting a comparative assessment of the performance of different ML models is challenging due to the significant heterogeneity in model frameworks and training/testing datasets. However, based on our review, decision trees, random forests, and artificial neural networks have shown higher effectiveness in interpreting and categorizing TFT. Table 1 provides the performance metrics of the best-performing models from the included studies.

Most studies have concentrated on evaluating the performance of ML models in binary classification (e.g., illness versus normal) of thyroid function status. Multi-class categorization remains more challenging due to the limited data available for minority classes. To address this issue, researchers have employed two strategies to manage imbalanced datasets: Oversampling, which involves augmenting minority class instances by generating synthetic samples or duplicating existing ones (17), and under-sampling, which reduces the majority class instances to match the minority class size (15).

A promising future direction involves fostering collaborative efforts across multiple institutions to pool datasets, creating comprehensive and diverse datasets that ensure better model generalization and performance.

On the other hand, it is important to acknowledge that while extensive datasets provide a wealth of training data, they can potentially hinder model performance when dealing with low-quality or complex data. In such cases, data pre-processing techniques, such as data cleaning, normalization, and feature engineering, are commonly employed to enhance model efficacy (14). Mir and Mittal demonstrated that the accuracy of thyroid disease classifiers improved after excluding laboratory features from the training dataset (18). Specifically, a support vector machine (SVM) classifier achieved an accuracy of 99.08% in predicting thyroid disorders within testing cohorts using only patient history and physical evaluations, eliminating the need for laboratory-based tests. By contrast, a bagging classifier trained on the original dataset, including laboratory results, achieved an accuracy of 98.56%. Although feature selection through DL requires greater computational resources and is costlier, it holds significant promise for enhancing future model performance.

Research has also explored identifying the most valuable predictors and detectors of thyroid dysfunction, extending beyond traditional TFTs to include routine clinical laboratory data and supplementary diagnostic methods (21, 22, 24, 25). For example, an ML model was developed to analyze the effects of various micronutrients and vitamins on TSH hormone regulation. The study revealed a direct correlation between TSH concentration and vitamin B9 levels, along with a slightly negative association with magnesium levels (23). Future research should prioritize identifying potential molecules that contribute to the onset of thyroid dysfunction, enabling timely diagnosis

Table 1. Thyroid Function Tests and Routine Clinical Laboratory Data and Artificial Intelligence-Powered Diagnosis of Thyroid Function Status

First Author, Year, Reference	Aim	Technique	Dataset	Sample Size	Performance Metrics
Cheng et al., 2022 (16)	Diagnosis of thyroid dysfunction from thyroid datasets (TFT)	DT, LR, XGB, and SVM	21 features, including TFT	11565 NI and 1170 elevated TSH from 1 center	XGB: AUC 0.87, and ACC 0.86
Islam et al., 2022 (17)	Same	CatB, Extra-Trees, ANN, LGBM, SVC, KNN, RF, XGB, DT, and GaussianNB	UCI sick-euthyroid dataset (25 features)	2870 sick and 292 NI	ANN: F1-score 0.957, precision 0.957, recall 0.959, and ACC 0.9587
Abbad Ur Rehman et al. 2021 (19)	Same	KNN, NB, SVM, LR, and DT	DHQ Teaching Hospital thyroid dataset (10 features)	170 NI, 66 hyper, and 73 hypo	NB: ACC 100%, recall 100%, and F1-score 100%
Rasitha, 2016 (14)	Same	LDA	UCI hypothyroid dataset (29 features)	3481 NI, 194 compensated hypo, 95 primary hypo, and 2 secondary hypo	Precision 0.996, recall 0.996, ACC 99.62, and ROC 0.996
Yadav and Pal, 2020 (20)	Same	DT, RF, CART, and bagging ensemble model	UCI thyroid disease dataset (30 features)	3710 patients (2 classes)	Bagging ensemble model: ACC 100%
Chaganti et al., 2022 (15)	Same	RF, LR, SVM, ADA, GBM and CNN, LSTM, and CNN-LSTM	UCI thyroid disease dataset (30 features)	233 primary hypo, 359 compensated hypo, 346 increasing binding proteins, 456 concurrent non-thyroidal illness, and 400 NI	RF: F1-score 0.99, precision 0.99, recall 0.99, and ACC 0.99
Mir and Mittal, 2020 (18)	Same	Boosting, bagging, NB, SVM, and J48	21 features, including TFT	489 NI, 488 hyper, and 487 hypo from 1 center	SVM: ACC 99.08, precision 0.991, recall 0.991, and ROC 0.994
Yoshimura Noh et al., 2024 (21)	Diagnosis of thyroid dysfunctions from routines	ANN (Prediction one) and LR	JHEP (11 features), JND (9 features), and Ito (32 features) datasets including routine lab	20653 GD, 3435 painless thyroiditis, 4266 NI, and 18937 HT from 1 center	Thyrotoxicosis: AUC 0.977 Hypothyroidism: AUC 0.877
Hu et al., 2022 (22)	Same	GBDT, SVM, LR, and ANN	EMRs (23 features, including routine lab)	176727 patients from 4 centers	Hyperthyroidism: AUC 93.8% Hypothyroidism: AUC 90.9%
Ghali et al., 2020 (23)	Prediction of TSH by macroelements and vitamins	ANFIS, ANN, and MLR	7 vitamins and macronutrients	Blood sample of 1 patient	ANFIS: R2 0.914
Shin et al., 2023 (24)	Detection of hyperthyroidism	LGB	662 pairs of TFT and HR	175 patients (2 classes) from 1 center	Sensitivity 86.14%, specificity 98.28%, NPV 95.32, and PPV 94.57%
Choi et al., 2022 (25)	Same	DL	174331 ECGs for training, 48648 for external validation	146672 patients from 2 centers	AUC 0.926 for internal validation and AUC 0.883 for external validation

Abbreviations: DT, decision tree; LR, logistic regression; XGB, X gradient boosting; SVM, support vector machine; CatB, CatBoost; ANN, artificial neural network; GaussianNB, Gaussian naive bayes; LGBM, light gradient-boosting machine; SVC, support vector classifier; KNN, K-nearest neighbors; RF, random forest; LDA, linear discriminant analysis; CART, classification and regression tree; BP-AdaBoost, back propagation-adaptive boosting; CNN, convolutional neural network; ANFIS, adaptive neuro-fuzzy inference system; MLR, multiple linear regression; ADA, AdaBoost; GBM, gradient boosting machine; LSTM, long short-term memory; DL, deep learning; NI, normal; AUC, area under curve; ACC, accuracy; TFT, thyroid function tests.

and preventive measures, such as dietary iodine supplementation.

In another study, a DL algorithm was developed using diverse electrocardiograms (ECG) to detect subclinical and overt hyperthyroidism, achieving an area under the curve (AUC) of 0.926 for internal validation and 0.883 for external validation. However, the inherent complexity and lack of transparency in the decision-making processes of DL models—commonly referred to as the "black box" problem—may compromise the generalizability and reliability of their results (25). To enhance reliability and clinical applicability, future research should focus on developing explainable AI systems. These systems can provide greater insight into the decision-making processes of DL-based models, ensuring transparency. Additionally, validation using real-world patient cohorts and comparative studies with

expert physicians are essential to establish trust and efficacy.

3.2. Ultrasound and Scintigraphy

In the diagnosis of thyroid function disorders, thyroid scintigraphy plays a pivotal role in providing valuable insights into the secretory activity of the thyroid gland. However, accurate interpretation of thyroid scintigraphy heavily relies on the expertise of nuclear medicine specialists and is notably time-consuming (26). Given the promising performance of DL in medical imaging analysis, several studies have integrated DL networks to diagnose and differentiate various thyroid functional states using scintigraphy (27-30) (Table 2).

Kappa coefficient metrics for DL convolutional neural networks (DCNNs) have demonstrated a commendably

Table 2. Ultrasound and Nuclear Medicine & Artificial Intelligence-Powered Diagnosis of Thyroid Function Status

First Author, Year, Reference	Aim	Technique	Dataset	Sample Size	Performance Metrics
Acharya et al., 2014 (31)	Diagnosis of thyroid dysfunctions	SVM, DT, fuzzy classifier, and KNN	7 features from US images	232 NI and 294 HT from 1 center	Fuzzy classifier: ACC 84.6%
Zhang et al., 2022 (32)	Same	DL	Features from US images, US videos, and 6 features of TFT	37424 HT and 69089 NI from 1 hospital	By US videos + TFT: AUC 0.949 and ACC 0.892
Zhao et al., 2022 (33)	Same	DL	Features from US images	20666 HT and 18613 non-HT from 2 centers	Acc 0.892, AUC 0.940, and F1-score 0.892
Vasile et al., 2021 (34)	Same	DL	Features from US images	767 autoimmune, 672 micro-nodular, 720 nodular, and 638 NI from 4 centers	Acc 98.78, and AUC 0.98
Qiao et al., 2021 (28)	Same	DL	1430 thyroid scintigraphies	175 NL, 834 GD, and 421 subacute thyroiditis from 1 center	Subacute thyroiditis: F1-score 84.98 %, precision 77.99%, recall 93.33%, and ACC 89.00% GD: F1-score 88.62 %, precision 93.36%, recall 84.33%, and ACC 92.78%
Yang et al., 2021 (30)	Same	DL	3389 thyroid scintigraphies from 3 centers	4 classes of scintigraphy pattern	Overall ACC 92.73%,
Zhao et al., 2023 (29)	Same	DL	3194 thyroid SPECT	742 NI, 808 GD, 826 subacute thyroiditis, and 818 tumors from 3 centers	Subacute thyroiditis: F1-score 0.958, recall 93.9, precision 97.6, and AUC 0.992 GD: F1-score 0.981, recall 100.0, precision 96.3, and AUC 0.999
Kikuchi et al., 2023 (27)	Same	LGBM	7013 F-18 FDG PET/CT scans	182 hypo, 265 hyper, and 6566 NI from 1 center	Hypothyroidism: AUC 0.77 Hyperthyroidism: AUC 0.78

Abbreviations: SVM, support vector machine; DT, decision tree; DL, deep learning; KNN, K-nearest neighbors; US, ultrasound; HT, hypothyroidism; TFT, thyroid function tests; SPECT, single-photon emission computed tomography scan; GD, Graves' disease; LGBM, light gradient boosting machine; F-18 FDG PET/CT, fluorine-18 fluorodeoxyglucose positron emission tomography; NL, normal; AUC, area under curve; ACC, accuracy.

high level of agreement (> 0.715) between DCNN-generated outputs and definitive diagnoses, highlighting the precision of these models. For example, the Residual Neural Network (ResNet) performed exceptionally well among various CNN frameworks in a study conducted by Zhao et al. This model achieved an F1-score of 0.981 and an AUC of 0.993 for diagnosing Graves' disease based on SPECT images following 5-fold cross-validation (29).

DCNN-based models have also outperformed junior nuclear medicine residents, although findings have shown variability when compared to senior residents (28, 29). Integrating DL-assisted diagnostic systems holds promise for aiding less experienced physicians, who may sometimes overlook subtle features during scintigraphy assessments (28). Despite this potential, practical implementation of DL assistance in clinical settings may face challenges due to the "black box" problem, a lack of external multicentric validation, and limited training datasets, particularly since scintigraphy is an infrequent clinical procedure.

To address these issues, Zhao et al. utilized gradient-weighted class activation mapping (Grad-CAM) to generate attentional heat maps, highlighting regions of the original input image that the model focused on. This approach improved the interpretability of their proposed framework (29). Future research should prioritize incorporating such interpretative

methodologies, like Grad-CAM, to enhance the transparency and understanding of DL networks, thereby increasing their trustworthiness and clinical applicability.

3.3. Thyroid Ultrasound Computer-Aided Diagnosis Systems

Thyroid ultrasound computer-aided diagnosis (CAD) systems have emerged as powerful tools for the accurate and efficient diagnosis of thyroid diseases, particularly thyroid nodules. Computer-aided diagnosis systems can also assist in detecting diffuse inflammatory changes in the gland. For instance, ThyroScan, an ML-based CAD system, extracts significant grayscale features from thyroid ultrasound images to detect Hashimoto's thyroiditis with relatively high accuracy (84.6%) (31). By incorporating ultrasound videos rather than static images and generating synthetic, high-dimensional, detailed instances alongside clinical and serological data, DL-based models have demonstrated promising performance in diagnosing thyroid dysfunctions (32, 33).

Deep convolutional neural networks (DCNNs) have shown superior performance by leveraging transfer learning, which involves reusing a model trained for one task in another context. It is important to note that the framework of CNN layers significantly influences model performance. For example, adding a dropout

Table 3. Artificial Intelligence Prediction of Major Depressive Disorder and Pregnancy Outcomes by Thyroid Function Status

First Author, Year, Reference	Aim	Technique	Dataset	Sample Size	Performance Metrics
Stroek et al., 2023 (41)	Congenital hypothyroidism screening	RF	Features from the Deutch National database	458 CH-T, 82 CH-C, 2332 false-positive referrals, and 1670 NI	ACC 0.77
Sun et al., 2021 (42)	Prediction of pregnancy outcome	LR, RF, XGB, and DL	Obstetrics and pre- and post-conception serum TSH features	3428 delivery from 1 center	XGB: PRETERM BIRTH: AUC 0.812 LOW APGAR SCORE: AUC 0.987, RF: INDUCTION: AUC 0.650
Araya et al., 2021 (43)	Prediction of GDM	PCA	29 thyroidal and non-thyroidal features in 1-2 trimesters	39 pregnancies from 1 center	NA
Mennickent et al., 2023 (44)	Same	LR, L-SVM, PLS-DA, CART, and XGB	75 thyroidal and non-thyroidal features in 1-2 trimesters	12 GDM and 54 NGT from 3 centers	PLS-DA: AUC 0.940
Zhou et al., 2022 (45)	Prediction of preterm delivery	GAM	Features from routine prenatal examination	3176 preterm birth, 2127 Spontaneous preterm birth, and 1049 Iatrogenic preterm birth from 1 center	NA
Zhang et al., 2021 (46)	Prediction of postpartum depression	RF, DT, XGB, LR, and MLP	EMRs (32 features)	14187 non-PPD and 1010 PPD from 1 center for training, 50459 non-PPD and 3513 PPD from another center for validation	LR: AUC 0.937 training and 0.886 validation
Yuan et al., 2023 (47)	Prediction of abortion	LR and XGB	48 features	340 abortions and 677 IVF-treated delivery from 1 center	XGB: AUC 0.759 and F1-score 0.566
Yang et al., 2023 (40)	Prediction of suicide attempt	LASSO	Hamilton depression and anxiety symptoms and biological features	1372 non-attempts and 208 attempts from 1 center	AUC 0.72
Li et al., 2021 (38)	Same	GBDT	Hamilton depression and anxiety symptoms and biological features	1372 non-attempts, 235 recent attempts, and 111 late attempts from 1 center	RECENT: ACC 87% LATE : ACC 88%
Qiao et al., 2022 (39)	Prediction of MDD prognosis	SVM	Hamilton depression and anxiety symptoms and TFT features	2086 MDD from 1 center	ROC-AUC 0.86

Abbreviations: GDM, gestational diabetes mellitus; NGT, normal glucose tolerance; LR, logistic regression; XGB, X gradient boosting; SVM, support vector machine; RF, Random forest; PCA, principal component analysis; L-SVM, linear support vector machine; MLP, multilayer perception; LASSO, least absolute shrinkage and selection operator; CART, classification and regression tree; GAM, generalized additive model; DT, decision tree; GBDT, gradient-boosting decision tree; DL, deep learning; PLS-DA, partial least-squares discriminant analysis; EMR, electronic medical record; NL, normal; AUC, area under curve; ACC, accuracy; MDD, major depressive disorder.

layer effectively reduces the overfitting problem, improving model reliability (34). Detailed data from studies utilizing AI to identify thyroid dysfunctions via ultrasonography are presented in Table 2.

3.4. Pregnancy and Major Depressive Disorders

To this point, this review has focused on AI applications for diagnosing thyroid function disorders in the general population, without accounting for concurrent conditions or illnesses. Since the symptoms of thyroid dysfunction often develop gradually, individuals may remain asymptomatic for an extended period. Early diagnosis and intervention are therefore critical to preventing adverse outcomes, particularly in scenarios such as pregnancy and congenital hypothyroidism, where thyroid hormone levels are essential for fetal and neonatal brain development and pregnancy outcomes (35).

Moreover, untreated thyroid dysfunction can exacerbate coexisting conditions, such as cardiovascular, neurological, or psychological disorders

(36, 37). For example, major depressive disorder (MDD) represents a significant public health concern due to its impact on quality of life and its association with a higher risk of suicide. Thyroid function status has shown a strong correlation with the development of MDD and response to treatment. Developing ML-based models to predict the risk of suicidal attempts and MDD improvement based on thyroid function data could provide valuable insights into this issue (38-40).

Consequently, the application of AI to address thyroid dysfunctions within specific subgroups has garnered increasing attention in the current literature. Detailed findings are summarized in Table 3.

3.5. Bioinformatics, Exposure, Radioiodine Therapy, and Levothyroxine Dose Adjustment

Numerous factors contribute to the development of thyroid dysfunction conditions, including thyroid gland inflammation, autoimmune processes (e.g., Graves' disease and Hashimoto's thyroiditis), iodine intake, medications, thyroid surgery, radiation therapy, congenital developmental abnormalities (e.g., thyroid

agenesis or dysplasia), pituitary disorders, and pituitary or thyroid tumors (6, 7).

Autoimmune thyroid disorders result from a complex interplay between genetic predisposition, dietary influences, coexisting disorders, and environmental triggers (48, 49). By leveraging ML techniques in bioinformatics—an interdisciplinary field combining computer science, mathematics, and statistics for biological data analysis—new insights can be gleaned from large-scale biological datasets such as genomic sequences, protein structures, gene expression profiles, and regulatory mechanisms. Several studies have explored the discovery of diagnostic genes associated with thyroid and related autoimmune disorders (50, 51) (Table 4). For example, Shen et al. identified 75 genes linked to hyperthyroidism using a random walk ML model for gene analysis, achieving an AUC of 0.90. This model also encoded gene interaction networks based on existing gene regulation data to clarify the pathological mechanisms of gene interactions (52).

In recent years, exposure to endocrine-disrupting chemicals (EDCs) has become an increasing public health concern. These chemicals, found in various environmental sources such as air, water, and food, disrupt the normal functioning of the endocrine system by mimicking or blocking the action of natural hormones in the body (53). Some studies have explored the development of predictive ML models to screen TSH receptor agonists and predict their inhibitory or non-inhibitory activity and mechanisms (54, 55) (Table 4).

Radioiodine therapy (RAI) is a widely used treatment for hyperthyroidism, particularly in Graves' disease, as well as for removing residual thyroid tissue following thyroid cancer surgery (56). Hypothyroidism is a common early consequence of RAI, requiring lifelong thyroid hormone replacement therapy. To address this, researchers have developed ML models to identify risk factors associated with early hypothyroidism after RAI and to screen patients at higher risk. Findings suggest that male gender, older age, lower radioactive iodine uptake, smaller thyroid volume, and certain blood biochemical markers may contribute to an increased risk of hypothyroidism (57, 58). Additionally, ML models have been designed to predict the therapeutic dose of RAI. For instance, random forest (RF) models created for this purpose have demonstrated strong performance (58).

Moreover, patients who have undergone thyroidectomy require exogenous hormone replacement with levothyroxine (LT4). Adjusting the LT4 dose for these patients necessitates careful

consideration of individual variability in thyroid hormone requirements, as well as factors influencing the absorption and metabolism of LT4, to achieve euthyroidism. Regression and classification models, such as fuzzy logistic systems and decision trees, have shown promise in providing personalized approaches to LT4 dose adjustment (59-61). Future research should aim to develop models with even higher performance for this purpose. Table 4 also summarizes investigations conducted on the performance of AI systems in predicting post-RAI treatment hypothyroidism and LT4 dose adjustments.

3.6. Thyroid-Associated Ophthalmopathy

Thyroid-associated ophthalmopathy (TAO), also known as Graves' ophthalmopathy, is a potentially sight-threatening ocular condition characterized by upper lid retraction, lid lag, and eye protrusion. Numerous studies have evaluated the performance of AI in extracting eye parametric data from orbital imaging modalities, particularly computed tomography (CT) scans and magnetic resonance imaging (MRI), to detect this disorder and assess its severity (62-64) (Table 5). These investigations have primarily employed deep convolutional neural networks (DCNNs) with various architectures, such as Visual Geometry Group Network (VGG), Residual Neural Network (ResNet), and GoogleNet Inception (GoogLeNet). The performance of these models has been comparable to that of experienced ophthalmologists. For instance, Karlin et al. introduced an ensemble DL model that demonstrated greater accuracy than ophthalmologists in diagnosing TAO from external orbital images (65).

Researchers have also developed predictive models for treatment response and post-decompression surgery outcomes in TAO patients (66, 67). For example, using dual facial photographs from 109 patients who had undergone orbital decompression surgery, a generative adversarial network (GAN) achieved an accuracy of 90.9% and an AUC of 0.957. This model demonstrated one of the highest performances among predictive models for surgical outcomes in TAO patients (66). In another study, Zhai et al. highlighted the potential utility of T2 mapping and T2 Iterative Decomposition of Water and Fat with Echo Asymmetry and Least-Squares Estimation (IDEAL) as predictors for response to glucocorticoid therapy in patients with active and moderate-to-severe TAO (67).

The limitations of the mentioned models primarily arise from the fact that they are mostly trained on abnormal medical imaging, as healthy individuals typically do not undergo orbital imaging. This reliance

Table 4. Artificial Intelligence Applications in Bioinformatics, Exposure, Radioiodine Therapy, and Levothyroxine Dose Adjustment

First Author, Year, Reference	Aim	Technique	Dataset	Sample Size	Performance Metrics
Atas, 2023 (50)	Prediction of autoimmune concomitant dx with HT	SVM, RF, LR, KNN, MLP, and a ML hybrid model	OMIM, PUBMED, Entrez Gene on NCBI, NCBI dbSNP, and SWISS Prot database	162 genes	ACC 0.815, precision 0.731, recall 1.0, and F1-score 0.800
Li et al., 2024 (51)	HT diagnosis based on genes	LASSO	HRA001684, GSE29315 and GSE163203 datasets	2000 highly variant genes	NA
Shen et al., 2021 (52)	Prediction of genes of hyperthyroidism	RW-RVM, RF, ANN, and NB	DisGeNET	269 genes	AUC 0.90
Liu et al., 2023 (54)	Screening Tshr agonists	RF, MLP, SVM, and GAT	Updated TSHR agonist dataset from PubChem fingerprints	7 molecular representations	RF: AUC 0.984, and ACC 0.941
Xu et al., 2022 (55)	Detection of TSHR inhibitory chemicals	RF, XGB, and LR	Tox21 database	5952 compounds from a cAMP analysis	RF: ACC 0.85, recall 0.89, and AUC 0.92
Gao et al., 2021 (58)	Prediction of 131i therapeutic dose	BPNN, RBFNN, SVM, BP-AdaBoost, and RF	EMRs (17 features)	353 patients from several centers	RF: ACC 100%
Duan et al., 2022 (57)	Prediction of hypothyroidism after RAIT	ML	EMR (138 clinical and lab test features)	471 GD patients from 1 center	AUC 0.74 and F1-score 0.74
Chen et al., 2019 (59)	Levothyroxine dosage post-thyroidectomy	DT	LT4 doses and TSH levels	320 patients from 1 center	Correctly predicted dose adjustment 75%, confidence interval = 65% - 82%
Barrio et al., 2023 (60)	Same	ANN, RF, OLS, and LR	Demographic, clinical, and laboratory data	951 patients from 1 center	Met postopTSH goal 45.3%
Hemmati et al., 2023 (61)	Same	Fuzzy logic	NA	Thyrosim application to simulate thyroid hormone courses of a virtual thyroidectomized patient	NA

Abbreviations: HT, Hashimoto's thyroiditis; GD, Graves' disease; TSHR, thyroid stimulating hormone receptor; RAIT, radioactive iodine therapy; LT4, levothyroxine; LR, logistic regression; XGB, extreme gradient boosting; SVM, support vector machine; ANN, artificial neural network; KNN, K-nearest neighbors; RF, random forest; OLS, ordinary least squares; BP-AdaBoost, back propagation-adaptive boosting; RBFNN, radial basis functions neural Network; MLP, multilayer perceptron; LASSO, least absolute shrinkage and selection Operator; RW-RVM, random walk-relevance vector machine; GAT, graph attention network; BPNN, back propagation neural network; EMR, electronic medical record; NL, normal; AUC, area under curve; ACC, accuracy.

Table 5. Artificial Intelligence Applications in Thyroid-Associated Ophthalmopathy

First author, Reference, Number	Aim	Technique	Dataset	Sample Size	Performance Metrics
Lee et al., 2022 (62)	TAO diagnosis	DL	Orbital CT scans	99 mild GO, 94 mod-to-severe, and 95 NI from 1 center	AUC 0.895 - 0.979
Lin et al., 2024 (63)	Same	DL	Orbital CT scans	459 mild, 355 severe, and 373 NI from 1 center	ACC 89.5% and AUC 0.96 - 0.99
Lin et al., 2021 (64)	Same	DL	Orbital MRI	50 active phase and 110 inactive phase from 1 center	ACC 0.863, precision 0.680, and F1-score 0.712
Karlin et al., 2023 (65)	Same	DL	External orbital photographs	2288 images from 1 clinical dataset	ACC 89.2%, recall 93.4%, precision 79.7%, and F1 score 86.0%
Yoo et al., 2020 (66)	Prediction of post-orbital decompression surgery appearance	Generative adversarial network (GAN)	Facial photographs	109 pairs of matched pre- and postoperative facial images from amGoogle image search	ACC 90.9% and AUC 0.957
Zhai et al., 2021 (67)	Prediction of the therapeutic efficacy of IV glucocorticoids	Binary LR	Orbital MRI and clinical characteristics	35 responsive and 28 unresponsive orbit	AUC 0.844

Abbreviations: TAO: thyroid associated ophthalmopathy; LR: logistic regression; DL: deep learning; CT: computed tomography scan; MRI: magnetic resonance imaging; NI: normal; AUC: area under curve; ACC: accuracy.

on abnormal cases can lead to sample bias, as the absence of normal control data may skew the results, thereby limiting the generalizability of the models. To overcome these challenges, future research should focus on training models using larger and more diverse datasets that include normal controls, as well as other

differential diagnoses for TAO. These models should integrate both clinical and imaging data to identify correlations between imaging results and clinical presentations. Such an approach could expedite the diagnostic process and help prevent potential

permanent eye damage or vision loss through earlier intervention.

4. Discussion and Conclusions

Applying AI to the diagnosis and treatment of thyroid disorders holds immense potential to revolutionize patient care. The AI-powered models can interpret TFT results and provide comprehensive diagnostic assessments based on clinical and laboratory data. Most studies have utilized the University of California, Irvine (UCI) ML repository to train their interpretive models. The UCI repository is a publicly available, prominent resource for researchers and practitioners, offering a diverse collection of datasets that span various domains, including thyroid disease datasets that incorporate various TFT components and clinical variables.

Research has also explored AI applications for detecting thyroid dysfunction, ranging from routine clinical laboratory data to other supplementary diagnostic methods. The training datasets in these studies mainly consist of laboratory results, alongside demographic and clinical variables acquired from electronic medical records of single or multiple centers. Feature selection engineering allows for identifying the most valuable variables for training improved models.

Deep learning-based models primarily hold the potential to enhance the diagnostic accuracy of medical imaging for thyroid dysfunctions, particularly by analyzing thyroid scintigraphies and ultrasound images. DL image processing has also performed excellently in diagnosing and evaluating the severity of TAO from orbital CT scans. However, since imaging modalities for assessing thyroid function disorders are relatively uncommon, the lack of large, diverse, and balanced datasets remains a significant challenge in training such models. Researchers should collaborate across institutions to pool datasets and ensure better generalization.

Artificial intelligence can also aid in predicting pregnancy complications, improving newborn screening programs, forecasting treatment responses, assessing the likelihood of suicidal attempts in patients with MDD, and optimizing levothyroxine dose adjustment. To train AI models for these purposes, it is essential to include not only features of concurrent conditions, but also TFT results, routine laboratory data, and demographic, biological, psychosocial, and clinical variables to create trustworthy models.

While AI technology holds significant potential to improve the management of thyroid disorders, challenges such as data quality, model interpretability,

and the need for extensive validation have made it a long way from replacing specialists. By addressing these challenges, future research can develop models that provide a high level of reliability and advance the field of endocrinology.

Acknowledgements

The authors would like to formally express their deep appreciation to NSh, a Professor of TEFL at Shiraz University of Medical Sciences, for her invaluable assistance in the grammatical correction of this article.

Footnotes

Authors' Contribution: Study conception and design: A. K. and M. H. D.; data acquisition: A. N. and A. K.; preparation and critical revision of the manuscript for important intellectual content: P. Y. and F. A. All authors reviewed the results and approved the final version of the manuscript.

Conflict of Interests Statement: M. H. D. is the editor-in-chief (EIC) of this journal.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication.

Funding/Support: This study was not supported by financial sources, such as investment and commercial agencies or not-for-profit sectors.

References

- Campbell I. Thyroid and parathyroid hormones and calcium homeostasis. *Anaesth Intensive Care Med.* 2011;**12**(10):465-8. <https://doi.org/10.1016/j.mpaic.2011.07.003>.
- Garmendia Madariaga A, Santos Palacios S, Guillen-Grima F, Galofre JC. The incidence and prevalence of thyroid dysfunction in Europe: a meta-analysis. *J Clin Endocrinol Metab.* 2014;**99**(3):923-31. [PubMed ID: 24423323]. <https://doi.org/10.1210/jc.2013-2409>.
- Karimi F, Kalantarhormozi MR, Dabbaghmanesh MH, Ranjbar Omrani G. Thyroid disorders and the prevalence of antithyroid antibodies in Shiraz population. *Arch Iran Med.* 2014;**17**(5):347-51. [PubMed ID: 24784864].
- Overall JE, Williams JM. Conditional probability program for diagnosis of thyroid function. *JAMA.* 1963;**183**:307-13. [PubMed ID: 13940794]. <https://doi.org/10.1001/jama.1963.03700050039009>.
- Yazdanpanahi P, Atighi F, Keshtkar A, Hamidi R, Rezaeimanesh M, Karimi A, et al. The Current Progress of Artificial Intelligence in Approach to Thyroid Nodules: A Narrative Review. *Shiraz E-Med J.* 2024;**25**(11). <https://doi.org/10.5812/semj-148493>.
- Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med.* 2000;**160**(4):526-34. [PubMed ID: 10695693]. <https://doi.org/10.1001/archinte.160.4.526>.

7. Goichot B, Caron P, Landron F, Bouee S. Clinical presentation of hyperthyroidism in a large representative sample of outpatients in France: relationships with age, aetiology and hormonal parameters. *Clin Endocrinol (Oxf)*. 2016;**84**(3):445-51. [PubMed ID: 25959282]. <https://doi.org/10.1111/cen.12816>.
8. Ross DS. *Laboratory assessment of thyroid function*. 2024. Available from: <https://www.uptodate.com/contents/laboratory-assessment-of-thyroid-function>.
9. Ling C, Sun Q, Khang J, Lastarria M, Strong J, Stolze B, et al. Does TSH reliably detect hypothyroid patients? *AnnThyroid Res*. 2018;**4**(1):122.
10. Gruson D, Dabla P, Stankovic S, Homsak E, Gouget B, Bernardini S, et al. Artificial intelligence and thyroid disease management: considerations for thyroid function tests. *Biochem Med (Zagreb)*. 2022;**32**(2):20601. [PubMed ID: 35799984]. [PubMed Central ID: PMC9195598]. <https://doi.org/10.11613/BM.2022.020601>.
11. Mello C, Marangoni A, Poppi R, Noda I. Fast determination of thyroid stimulating hormone in human blood serum without chemical preprocessing by using infrared spectroscopy and least squares support vector machines. *Anal Chim Acta*. 2011;**696**(1-2):47-52. [PubMed ID: 21621031]. <https://doi.org/10.1016/j.aca.2011.04.015>.
12. Dip G, Colombo V, Lejona S, Lombardo A. B-125 Artificial Intelligence Application to Improve the Algorithm for Autoverification of Results of Biochemical Markers that Assess Thyroid Function. *Clin Chem*. 2023;**69**(Supplement_1):hvad097.459. <https://doi.org/10.1093/clinchem/hvad097.459>.
13. Stevenson E, Walsh C, Hibberd L. Can artificial intelligence replace biochemists? A study comparing interpretation of thyroid function test results by ChatGPT and Google Bard to practising biochemists. *Ann Clin Biochem*. 2024;**61**(2):143-9. [PubMed ID: 37699796]. <https://doi.org/10.1177/00045632231203473>.
14. Rasitha G. Predicting Thyroid Disease using Linear Discriminant Analysis (LDA) Data Mining Technique. *Commun Appl Electronics*. 2016;**4**(1):4-6. <https://doi.org/10.5120/cae2016651990>.
15. Chaganti R, Rustam F, De La Torre Diez I, Mazon JLV, Rodriguez CL, Ashraf I. Thyroid Disease Prediction Using Selective Features and Machine Learning Techniques. *Cancers (Basel)*. 2022;**14**(16). [PubMed ID: 36010907]. [PubMed Central ID: PMC9405591]. <https://doi.org/10.3390/cancers14163914>.
16. Cheng X, Li S, Deng L, Luo W, Wang D, Cheng J, et al. Predicting Elevated TSH Levels in the Physical Examination Population With a Machine Learning Model. *Front Endocrinol (Lausanne)*. 2022;**13**:839829. [PubMed ID: 35282438]. [PubMed Central ID: PMC8907627]. <https://doi.org/10.3389/fendo.2022.839829>.
17. Islam SS, Haque MS, Miah MSU, Sarwar TB, Nugraha R. Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study. *PeerJ Comput Sci*. 2022;**8**:e898. [PubMed ID: 35494828]. [PubMed Central ID: PMC9044232]. <https://doi.org/10.7717/peerj-cs.898>.
18. Mir YI, Mittal S. Thyroid Disease Prediction Using Hybrid Machine Learning Techniques: An Effective Framework. *Int J Sci Technol Res*. 2020;**9**(2):2868-74.
19. Abbad Ur Rehman H, Lin C, Mushtaq Z, Su S. Performance Analysis of Machine Learning Algorithms for Thyroid Disease. *Arabian J Sci Engin*. 2021;**46**(10):9437-49. <https://doi.org/10.1007/s13369-020-05206-x>.
20. Yadav DC, Pal S. Prediction of thyroid disease using decision tree ensemble method. *Hum-Intelligent Syst Integration*. 2020;**2**(1-4):89-95. <https://doi.org/10.1007/s42454-020-00006-y>.
21. Yoshimura Noh J, Yoshihara A, Hiruma S, Ichikawa M, Hirose R, Koshibu M, et al. An Evaluation of the Efficacy of Machine Learning in Predicting Thyrotoxicosis and Hypothyroidism: A Comparative Assessment of Biochemical Test Parameters Used in Different Health Checkups. *Intern Med*. 2024;**63**(16):2259-68. [PubMed ID: 38171877]. [PubMed Central ID: PMC11414366]. <https://doi.org/10.2169/internalmedicine.2825-23>.
22. Hu M, Asami C, Iwakura H, Nakajima Y, Sema R, Kikuchi T, et al. Development and preliminary validation of a machine learning system for thyroid dysfunction diagnosis based on routine laboratory tests. *Commun Med (Lond)*. 2022;**2**:9. [PubMed ID: 35603277]. [PubMed Central ID: PMC9053267]. <https://doi.org/10.1038/s43856-022-00071-1>.
23. Ghali UM, Usman AG, Degm MAA, Alsharksi AN, Naibi AM, Abba SI. Applications of artificial intelligence-based models and multi-linear regression for the prediction of thyroid stimulating hormone level in the human body. *Int J Adv Sci Technol*. 2020;**29**(4):3690-9.
24. Shin K, Kim J, Park J, Oh TJ, Kong SH, Ahn CH, et al. A machine learning-assisted system to predict thyrotoxicosis using patients' heart rate monitoring data: a retrospective cohort study. *Sci Rep*. 2023;**13**(1):21096. [PubMed ID: 38036639]. [PubMed Central ID: PMC10689821]. <https://doi.org/10.1038/s41598-023-48199-x>.
25. Choi B, Jang JH, Son M, Lee MS, Jo YY, Jeon JY, et al. Electrocardiographic biomarker based on machine learning for detecting overt hyperthyroidism. *Eur Heart J Digit Health*. 2022;**3**(2):255-64. [PubMed ID: 36713007]. [PubMed Central ID: PMC9707932]. <https://doi.org/10.1093/ehjdh/ztac013>.
26. Almomammed HI, Mansour S, Alhulwah AH, Mayhoub FH, Arafah AM. Scintigraphy has the potential to replace thyroid stimulating hormone and ultrasonography in hyperthyroidism diagnosis. *Saudi J Biol Sci*. 2020;**27**(7):1722-5. [PubMed ID: 32565688]. [PubMed Central ID: PMC7296473]. <https://doi.org/10.1016/j.sjbs.2020.05.015>.
27. Kikuchi T, Hanaoka S, Nakao T, Nomura Y, Yoshikawa T, Alam A, et al. Significance of FDG-PET standardized uptake values in predicting thyroid disease. *Eur Thyroid J*. 2023;**12**(1). [PubMed ID: 36562641]. [PubMed Central ID: PMC9986380]. <https://doi.org/10.1530/ETJ-22-0165>.
28. Qiao T, Liu S, Cui Z, Yu X, Cai H, Zhang H, et al. Deep learning for intelligent diagnosis in thyroid scintigraphy. *J Int Med Res*. 2021;**49**(1):300060520982842. [PubMed ID: 33445994]. [PubMed Central ID: PMC7812409]. <https://doi.org/10.1177/0300060520982842>.
29. Zhao H, Zheng C, Zhang H, Rao M, Li Y, Fang D, et al. Diagnosis of thyroid disease using deep convolutional neural network models applied to thyroid scintigraphy images: a multicenter study. *Front Endocrinol (Lausanne)*. 2023;**14**:1224191. [PubMed ID: 37635985]. [PubMed Central ID: PMC10453808]. <https://doi.org/10.3389/fendo.2023.1224191>.
30. Yang P, Pi Y, He T, Sun J, Wei J, Xiang Y, et al. Automatic differentiation of thyroid scintigram by deep convolutional neural network: a dual center study. *BMC Med Imaging*. 2021;**21**(1):179. [PubMed ID: 34823482]. [PubMed Central ID: PMC8620916]. <https://doi.org/10.1186/s12880-021-00710-4>.
31. Acharya UR, Sree SV, Krishnan MM, Molinari F, Zieleznik W, Bardales RH, et al. Computer-aided diagnostic system for detection of Hashimoto thyroiditis on ultrasound images from a Polish population. *J Ultrasound Med*. 2014;**33**(2):245-53. [PubMed ID: 24449727]. <https://doi.org/10.7863/ultra.33.2.245>.
32. Zhang Q, Zhang S, Pan Y, Sun L, Li J, Qiao Y, et al. Deep learning to diagnose Hashimoto's thyroiditis from sonographic images. *Nat Commun*. 2022;**13**(1):3759. [PubMed ID: 35768466]. [PubMed Central ID: PMC9243092]. <https://doi.org/10.1038/s41467-022-31449-3>.
33. Zhao W, Kang Q, Qian F, Li K, Zhu J, Ma B. Convolutional Neural Network-Based Computer-Assisted Diagnosis of Hashimoto's Thyroiditis on Ultrasound. *J Clin Endocrinol Metab*. 2022;**107**(4):953-63. [PubMed ID: 34907442]. [PubMed Central ID: PMC8947219]. <https://doi.org/10.1210/clinem/dgab870>.
34. Vasile CM, Udristoiu AL, Ghenea AE, Popescu M, Gheonea C, Niculescu CE, et al. Intelligent Diagnosis of Thyroid Ultrasound Imaging Using an Ensemble of Deep Learning Methods. *Medicina (Kaunas)*.

- 2021;57(4). [PubMed ID: 33921597]. [PubMed Central ID: PMC8073676]. <https://doi.org/10.3390/medicina57040395>.
35. Saki F, Dabbaghmanesh MH, Ghaemi SZ, Forouhari S, Ranjbar Omrani G, Bakhshayeshkaram M. Thyroid function in pregnancy and its influences on maternal and fetal outcomes. *Int J Endocrinol Metab*. 2014;12(4). e19378. [PubMed ID: 25745488]. [PubMed Central ID: PMC4338651]. <https://doi.org/10.5812/ijem.19378>.
 36. Aweimer A, El-Battrawy I, Akin I, Borggreffe M, Mugge A, Patsalis PC, et al. Abnormal thyroid function is common in takotsubo syndrome and depends on two distinct mechanisms: results of a multicentre observational study. *J Intern Med*. 2021;289(5):675-87. [PubMed ID: 33179374]. <https://doi.org/10.1111/joim.13189>.
 37. Jia Y, Chang CH, Hughes TF, Jacobsen E, Wang S, Berman SB, et al. Predictors of Dementia in the Oldest Old: A Novel Machine Learning Approach. *Alzheimer Dis Assoc Disord*. 2020;34(4):325-32. [PubMed ID: 32701513]. [PubMed Central ID: PMC7677183]. <https://doi.org/10.1097/WAD.0000000000000400>.
 38. Li XY, Tabarak S, Su XR, Qin Z, Chai Y, Zhang S, et al. Identifying clinical risk factors correlate with suicide attempts in patients with first episode major depressive disorder. *J Affect Disord*. 2021;295:264-70. [PubMed ID: 34482058]. <https://doi.org/10.1016/j.jad.2021.08.028>.
 39. Qiao D, Liu H, Zhang X, Lei L, Sun N, Yang C, et al. Exploring the potential of thyroid hormones to predict clinical improvements in depressive patients: A machine learning analysis of the real-world based study. *J Affect Disord*. 2022;299:159-65. [PubMed ID: 34856305]. <https://doi.org/10.1016/j.jad.2021.11.055>.
 40. Yang W, Wang X, Kang C, Yang L, Liu D, Zhao N, et al. Establishment of a risk prediction model for suicide attempts in first-episode and drug naive patients with major depressive disorder. *Asian J Psychiatr*. 2023;88:103732. [PubMed ID: 37586124]. <https://doi.org/10.1016/j.ajp.2023.103732>.
 41. Stroek K, Visser A, van der Ploeg CPB, Zwaveling-Soonawala N, Heijboer AC, Bosch AM, et al. Machine learning to improve false-positive results in the Dutch newborn screening for congenital hypothyroidism. *Clin Biochem*. 2023;116:7-10. [PubMed ID: 36878346]. <https://doi.org/10.1016/j.clinbiochem.2023.03.001>.
 42. Sun Y, Zheng W, Zhang L, Zhao H, Li X, Zhang C, et al. Quantifying the Impacts of Pre- and Post-Conception TSH Levels on Birth Outcomes: An Examination of Different Machine Learning Models. *Front Endocrinol (Lausanne)*. 2021;12:755364. [PubMed ID: 34777251]. [PubMed Central ID: PMC8586450]. <https://doi.org/10.3389/fendo.2021.755364>.
 43. Araya J, Rodriguez A, Lagos-SanMartin K, Mennickent D, Gutierrez-Vega S, Ortega-Contreras B, et al. Maternal thyroid profile in first and second trimester of pregnancy is correlated with gestational diabetes mellitus through machine learning. *Placenta*. 2021;103:82-5. [PubMed ID: 33099203]. <https://doi.org/10.1016/j.placenta.2020.10.015>.
 44. Mennickent D, Ortega-Contreras B, Gutierrez-Vega S, Castro E, Rodriguez A, Araya J, et al. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia. *PLoS One*. 2023;18(1). e0280513. [PubMed ID: 36638142]. [PubMed Central ID: PMC9838876]. <https://doi.org/10.1371/journal.pone.0280513>.
 45. Zhou Y, Liu Y, Zhang Y, Zhang Y, Wu W, Fan J. Identifying Non-Linear Association Between Maternal Free Thyroxine and Risk of Preterm Delivery by a Machine Learning Model. *Front Endocrinol (Lausanne)*. 2022;13:817595. [PubMed ID: 35282469]. [PubMed Central ID: PMC8907667]. <https://doi.org/10.3389/fendo.2022.817595>.
 46. Zhang Y, Wang S, Hermann A, Joly R, Pathak J. Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women. *J Affect Disord*. 2021;279:1-8. [PubMed ID: 33035748]. [PubMed Central ID: PMC7738412]. <https://doi.org/10.1016/j.jad.2020.09.113>.
 47. Yuan G, Lv B, Du X, Zhang H, Zhao M, Liu Y, et al. Prediction model for missed abortion of patients treated with IVF-ET based on XGBoost: a retrospective study. *Peer J*. 2023;11. e14762. [PubMed ID: 36743954]. [PubMed Central ID: PMC9893909]. <https://doi.org/10.7717/peerj.14762>.
 48. Ashouri E, Dabbaghmanesh MH, Ranjbar Omrani G. Presence of more activating KIR genes is associated with Hashimoto's thyroiditis. *Endocrine*. 2014;46(3):519-25. [PubMed ID: 24174177]. <https://doi.org/10.1007/s12020-013-0080-2>.
 49. Kalantar K, Ghandehari F, Malek-Hosseini S, Golmoghaddam H, Rostamzadeh D, Dabbaghmanesh MH, et al. Association of FCRL3 rs7528684 polymorphism with risk of Hashimoto's thyroiditis in Iranian patients. *Meta Gene*. 2020;24. <https://doi.org/10.1016/j.mgene.2020.100663>.
 50. Atas PK. A novel hybrid model to predict concomitant diseases for Hashimoto's thyroiditis. *BMC Bioinformatics*. 2023;24(1):319. [PubMed ID: 37620755]. [PubMed Central ID: PMC10464155]. <https://doi.org/10.1186/s12859-023-05443-5>.
 51. Li Y, Lu X, Cao W, Liu N, Jin X, Li Y, et al. Exploring the diagnostic value of endothelial cell and angiogenesis-related genes in Hashimoto's thyroiditis based on transcriptomics and single cell RNA sequencing. *Arch Biochem Biophys*. 2024;757:110013. [PubMed ID: 38670301]. <https://doi.org/10.1016/j.abb.2024.110013>.
 52. Shen F, Cai W, Gan X, Feng J, Chen Z, Guo M, et al. Prediction of Genetic Factors of Hyperthyroidism Based on Gene Interaction Network. *Front Cell Dev Biol*. 2021;9:700355. [PubMed ID: 34409035]. [PubMed Central ID: PMC8365469]. <https://doi.org/10.3389/fcell.2021.700355>.
 53. Diamanti-Kandarakis E, Bourguignon JP, Giudice LC, Hauser R, Prins GS, Soto AM, et al. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. *Endocr Rev*. 2009;30(4):293-342. [PubMed ID: 19502515]. [PubMed Central ID: PMC2726844]. <https://doi.org/10.1210/er.2009-0002>.
 54. Liu W, Wang Z, Chen J, Tang W, Wang H. Machine Learning Model for Screening Thyroid Stimulating Hormone Receptor Agonists Based on Updated Datasets and Improved Applicability Domain Metrics. *Chem Res Toxicol*. 2023;36(6):947-58. [PubMed ID: 37209109]. <https://doi.org/10.1021/acs.chemrestox.3c00074>.
 55. Xu X, Wang C, Gui B, Yuan X, Li C, Zhao Y, et al. Application of machine learning to predict the inhibitory activity of organic chemicals on thyroid stimulating hormone receptor. *Environ Res*. 2022;212(Pt A):113175. [PubMed ID: 35351457]. <https://doi.org/10.1016/j.envres.2022.113175>.
 56. Premawardhana LD, Lazarus JH. Management of thyroid disorders. *Postgrad Med J*. 2006;82(971):552-8. [PubMed ID: 16954449]. [PubMed Central ID: PMC2585732]. <https://doi.org/10.1136/pgmj.2006.047290>.
 57. Duan L, Zhang HY, Lv M, Zhang H, Chen Y, Wang T, et al. Machine learning identifies baseline clinical features that predict early hypothyroidism in patients with Graves' disease after radioiodine therapy. *Endocr Connect*. 2022;11(5). [PubMed ID: 35521803]. [PubMed Central ID: PMC9175589]. <https://doi.org/10.1530/EC-22-0119>.
 58. Gao H, Yuan W, Gao Y, Wang Y, Yao J, Gao L, et al. Prediction of 131I Therapeutic Dose and Prognosis in Hyperthyroidism Patients Using Mechanical Learning Model. *Res Square*. 2021;Preprint. <https://doi.org/10.21203/rs.3.rs-938232/v1>.
 59. Chen SS, Zaborek NA, Doubleday AR, Schaefer SC, Long KL, Pitt SC, et al. Optimizing Levothyroxine Dose Adjustment After Thyroidectomy With a Decision Tree. *J Surg Res*. 2019;244:102-6. [PubMed ID: 31279993]. [PubMed Central ID: PMC7041547]. <https://doi.org/10.1016/j.jss.2019.06.025>.
 60. Barrio M, Raeburn CD, McIntyre R, Albuja-Cruz M, Haugen BR, Pozdveyev N. Computer-Assisted Levothyroxine Dose Selection for the Treatment of Postoperative Hypothyroidism. *Thyroid*. 2023;33(5):547-55. [PubMed ID: 37084246]. <https://doi.org/10.1089/thy.2023.0033>.

61. Hemmati M, Tabesh H, Peyghambari A, Bazrafshan H. In silico dose adjustment of levothyroxine after total thyroidectomy using fuzzy logic methodology: A proof-of-concept study. *Heliyon*. 2023;**9**(1):e12797. [PubMed ID: 36685426]. [PubMed Central ID: PMC9852674]. <https://doi.org/10.1016/j.heliyon.2023.e12797>.
62. Lee J, Seo W, Park J, Lim WS, Oh JY, Moon NJ, et al. Neural network-based method for diagnosis and severity assessment of Graves' orbitopathy using orbital computed tomography. *Sci Rep*. 2022;**12**(1):12071. [PubMed ID: 35840769]. [PubMed Central ID: PMC9287334]. <https://doi.org/10.1038/s41598-022-16217-z>.
63. Lin LY, Zhou P, Shi M, Lu JE, Jeon S, Kim D, et al. A Deep Learning Model for Screening Computed Tomography Imaging for Thyroid Eye Disease and Compressive Optic Neuropathy. *Ophthalmol Sci*. 2024;**4**(1):100412. [PubMed ID: 38046559]. [PubMed Central ID: PMC10692956]. <https://doi.org/10.1016/j.xops.2023.100412>.
64. Lin C, Song X, Li L, Li Y, Jiang M, Sun R, et al. Detection of active and inactive phases of thyroid-associated ophthalmopathy using deep convolutional neural network. *BMC Ophthalmol*. 2021;**21**(1):39. [PubMed ID: 33446163]. [PubMed Central ID: PMC7807896]. <https://doi.org/10.1186/s12886-020-01783-5>.
65. Karlin J, Gai L, LaPierre N, Danesh K, Farajzadeh J, Palileo B, et al. Ensemble neural network model for detecting thyroid eye disease using external photographs. *Br J Ophthalmol*. 2023;**107**(11):1722-9. [PubMed ID: 36126104]. <https://doi.org/10.1136/bjo-2022-321833>.
66. Yoo TK, Choi JY, Kim HK. A generative adversarial network approach to predicting postoperative appearance after orbital decompression surgery for thyroid eye disease. *Comput Biol Med*. 2020;**118**:103628. [PubMed ID: 32174327]. <https://doi.org/10.1016/j.compbiomed.2020.103628>.
67. Zhai L, Luo B, Wu H, Wang Q, Yuan G, Liu P, et al. Prediction of treatment response to intravenous glucocorticoid in patients with thyroid-associated ophthalmopathy using T2 mapping and T2 IDEAL. *Eur J Radiol*. 2021;**142**:109839. [PubMed ID: 34252869]. <https://doi.org/10.1016/j.ejrad.2021.109839>.