

Sensitivity and specificity of clinical diagnosis of Graves' disease and non-Graves' disorders compared to serum TSH receptor antibodies

Received: 25/11/2020

Accepted: 30/03/2021

Mohammed Qader Meena^{1*}

Abstract

Background and objective: Sensitivity and specificity of clinical diagnosis of different hyperthyroid conditions were studied and compared with serum thyrotropin receptor antibodies test, as a standard test for diagnosing Graves' disease. This study aimed to assess the degree of sensitivity and specificity of clinical diagnosis of hyperthyroid conditions compared to the thyrotropin receptor antibodies test, which was selected as a standard test for diagnosis of Graves' disease in the study.

Methods: We studied patients presenting to endocrine outpatient clinics and other outpatient settings in Erbil city, presenting with clinical features of thyrotoxicosis. Serum or plasma thyroid-stimulating hormone, free T4, and thyrotropin receptor antibodies measurements were done for all the patients. Thyroid peroxidase antibodies and thyroid ultrasonography were done for most patients. Radioactive thyroid uptake was available for a limited number of patients. Sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, and accuracy index of clinical diagnosis of the thyrotoxicosis cases were calculated.

Results: A total of 150 patients were included in our study, female 102, males 48, mean age 38 (range 16-78) years, pretest diagnosis of (Graves' disease N=58, and non-Graves' disease N=92 and posttest diagnosis of (Graves' disease n= 59 vs. non-Graves' disease=91). Sensitivity of 74.6%, specificity (84.60%), positive predictive value (75.9%), negative predictive value (83.7%), positive likelihood ratio (4.847), negative likelihood ratio (0.30), and accuracy index (80.7%) of the clinical diagnosis were estimated.

Conclusion: The study showed that clinical diagnosis of Graves' disease and non-Graves' disease disorders, using thorough clinical examination and primary essential investigations, has high sensitivity and specificity index. However, the serum thyrotropin receptor antibodies test remains the standard method in the differential diagnosis of Graves' disease. Other studies are needed to study ultrasound and scintigraphy compared to the thyrotropin receptor antibodies test.

Keywords: Graves' disease; TSH receptor antibodies; Thyrotoxicosis; Toxic multinodular goiter; Thyroiditis.

Introduction

Thyrotoxicosis is a clinical state resulting from inappropriately high thyroid hormone levels.¹ Hyperthyroidism has a prevalence of 1% to 2% in women and 0.1% to 0.2% in men. The most common causes of an overactive thyroid are Graves' disease and toxic multinodular goiter.² Graves' disease is caused by the development of unique human autoantibodies to the thyroid-

stimulating hormone (thyrotropin, TSH) receptor; these autoantibodies act as TSH receptor agonists.²

Other common conditions causing hyperthyroidism include the passive release of thyroid hormones from damaged thyroid follicles; inflammation of the thyroid gland (thyroiditis), which may be autoimmune, post-viral, or drug-induced; and extrathyroidal sources of thyroid

¹ Department of Medicine, College of Medicine, Hawler Medical University, Erbil, Iraq.

* Correspondence: mmeneh@gmail.com

hormone, most often iatrogenic or self-administered.²

Measurement of thyrotropin receptor antibodies (TRAb) plays a crucial role in the differential diagnosis of hyperthyroidism, which has important therapeutic and prognostic implications.^{3,4} The autoimmune production of TRAb is central to the pathogenesis of Graves' disease.⁵

European clinicians and investigators advocate using TRAb as the primary test in the initial workup of hyperthyroidism.¹ The American Thyroid Association (ATA) and the American Association of Clinical Endocrinologists, in their joint guidelines, indicate a thyroid scan as the primary differential diagnostic test.⁶

Graves' disease is diagnosed based on typical clinical features of hyperthyroidism such as weight loss, fatigue, heat intolerance, tremor, palpitations, and diffuse thyroid enlargement, plus specific clinical features of Graves' disease, including orbitopathy, thyroid dermopathy (pretibial myxedema), and thyroid acropachy.⁵

Serum analyses typically show suppressed TSH and elevated thyroid hormones, tetraiodothyronine (T4; thyroxine), and triiodothyronine (T3).⁶

Additional diagnostic tests can include imaging, commonly, ultrasound and radioisotope uptake study, and thyroid autoantibodies, which can help to distinguish Graves' disease from other causes of thyrotoxicosis.⁷

The 3rd generation Thyrotropin Binding Inhibitor Immunoglobulin (TBII) assays have been found to have a sensitivity of over 97.2% and a specificity of over 98.3%.⁷ Furthermore, the first immunoassay method declared to measure serum TRAb concentration has recently been successfully developed in an automated commercial platform with a sensitivity of 100% and specificity of 99%.⁷ The sensitivity and specificity of clinical diagnosis of different hyperthyroid conditions were compared with serum

TRAb test, as a standard test for diagnosing Graves' disease. This study aimed to assess the degree of sensitivity and specificity of clinical diagnosis of hyperthyroid conditions compared to the TRAb test, which was selected as a standard test for diagnosis of Graves' disease in the study. The aim of this study was to demonstrate the accuracy of clinical diagnosis compared with the TRAb test in the diagnostic approach of Grave's disease, using clinical features and other basic laboratory and imaging studies.

Methods

A cross-sectional study was conducted on 150 patients presenting with clinical features of thyrotoxicosis, in the Endocrinology clinic, Erbil Teaching Hospital, and other outpatient settings, in Erbil City, from December 2018 to February 2020. After clinical evaluation and basic investigation as thyroid function test (TFT), ultrasound scan (US) of thyroid and thyroid peroxidase antibodies (TPO Ab) were sent for most of the patients, and TSH receptor antibodies (TRAb) for all the patients. The decision was made on clinical diagnosis, as Graves hyperthyroid or non-Graves hyperthyroid disorders. TRAb were sent for the two groups. Posttest diagnosis were evaluated.

Elecsys Anti-TSH Receptor Antibody was used for estimating plasma or serum levels of TRAb. Elecsys Anti-TSH Receptor antibody is a fully automated test for the detection of autoantibodies to the TSH receptor.

Elecsys Thyrotropin (TSH) Assay was used to estimate serum levels of TSH.

COBAS Elecsys anti-TPO were used for assessing TPO Antibodies. Free T4 (FT4) was assessed using Roche Cobas Laboratories Elecsys FT4 II Immunoassay for the in vitro quantitative determination of free thyroxine in human serum and plasma.

After complete clinical assessment, including thyroid status and neck examination for the thyroid gland,

the clinical status of thyrotoxicosis was confirmed by serum TSH level and Free T4 levels.

Ultrasound examination of the thyroid gland was performed for most of the patients included in this study to inspect for evidence of diffuse thyroid enlargement, as in the case of Graves' disease and thyroiditis, or for multinodular thyroid enlargement as in toxic multinodular goiter, or toxic solitary thyroid adenomas. Doppler ultrasound of thyroid gland was not obtained for suspected Graves' disease patients because of technical unavailability. We could not include thyroid uptake scan for our patients in the study because of the unavailability of the test.

Few patients had 99mTechnitium pertechnetate scans available after primary diagnosis of Graves' disease, which was competent with posttest results. The radioactive Iodine Uptake (RAIU) test was excluded in the study.

Pretest clinical evaluation was performed for all the patients, and the patients were categorized as provisional diagnosis of Graves' disease and non-Graves' disease, and posttest serum TRAb were estimated and diagnosis of Graves' disease or non-Graves' disease were applied according to TRAb positivity and the serum titer.

Clinical diagnosis was given for two groups: The first group was provisionally diagnosed as Graves' disease according to clinical features and some preliminary tests, and the second group were subgrouped as non-Graves diseases disorders, including Toxic multinodular Goiter, toxic solitary adenoma, different types of thyroiditis, including subacute painful thyroiditis, postpartum thyroiditis, and other types of thyroiditis.

The study was approved by the ethics committee, College of Medicine, Hawler Medical University. Informed verbal consent was taken from all the patients included in the study.

The data were analyzed using the following equations

$Se = TP / (TP + FN)$; $Sp = TN / (TN + FP)$;
 $PPV = TP / (TP + FP)$; $NPV = TN / (TN + FN)$;
 $PLR = Se / (1 - Sp)$; $NLR = (1 - Se) / Sp$;
 $AI = (TP + TN) / (TP + TN + FP + FN)$.

Statistical Analysis

The statistical package for the social sciences (SPSS 24) and Microsoft Excel were used for data analysis.

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood (PLR), negative likelihood ratio (NLR), and accuracy index (AI) of clinical diagnosis against serum TRAb test were calculated.

Chi-square and Fisher's exact test were used to show the relationship between groups. A *P* value of ≤ 0.05 was regarded as significant.

The data were assumed not to be normally distributed or have violations for assumptions of normality. Mean rank of TRAb, using Mann Whitney *U* test, between Graves' disease and non-Graves' disease posttest, and different variables were calculated, and significance (*P* value) were estimated,

McNemar's test was used for assessing the significance of difference in pretest and posttest results for the two groups.

Results

One hundred fifty cases were studied, with a pre-TRAb test clinical diagnosis of Graves' disease or non-Graves' disease hyperthyroidism. A total of 102 (68%) were females, and 48 (32%) were males. The age of the patients was between 16-78 years, with a mean age and standard deviation of 38.11 ± 12.781 years.

In pretest evaluation, 58 patients (38.66%) were categorized as Graves' disease, while after performing TRAb test 59 (39.33%) were found to have Graves' disease. In pretest, 92 (61.33%) were diagnosed as non-graves disorders, and posttest found to be 91 (60.66%) were found to be non-graves disorders.

Graves positive and Graves-Negative were estimated in the total number of patients, and among the male and female groups,

in pretest and posttest groups (Table 1). Running McNemar's test, the significance between the groups was calculated with *P* values (Table 1). A McNemar's Chi-square test was run to determine if there was a difference in the Graves' disease pre- and post TRAb test. It was found that the *P*-value associated with this test statistic was *P* value = 0.8527, = 0.705, = 0.669,

in all cases, male group, and female group, respectively (Table 1).

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated: sensitivity (74.60%), specificity (84.6), PPV (75.90%), and NPV (83.70%), as shown in Table 2.

Table 1 Graves' disease and non-Graves' disease cases with significant relations and *P* values

Clinical Diagnosis	Positive (Graves)	Negative (Non-Graves)		<i>P</i> value
Graves' disease	44	14	58	0.852
Non-Graves' disease	15	77	92	
Total	59	91	150	
Males				
Graves' disease	17	4		0.705
Non-Graves' disease	3	24		
Total	20	28	48	
Females				
Graves' disease	27	10		0.669
Non-Graves' disease	12	53		
Total	39	63	102	

Table 2: Sensitivity, specificity, positive predictive value, negative predictive values of all cases

All (N = 150)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Total (N=150)	74.6%	84.6%	75.9%	83.7%
Females (N=102)	69.2%	84.1%	73.0%	81.5%
Males (N=48)	85.0%	85.7%	81.0%	88.9%

NPV: Negative Predictive Value, NPV: Negative Predictive Value

True positive, false positive, true negative, false negative, positive likelihood ratio (PLR) (4.847), negative likelihood ratio (NLR) (0.30), and accuracy index (AI) (80.7%) were calculated (Table 3). The Mean Rank of serum TRAb levels in relation to different variables was estimated

with a significant *P* value of ≤ 0.05 regarded as significant (Table 4). The TRAb positive group has Mean Rank (104.15), compared with TRAb negative group of (56.920), which was significantly different, with *P* < 0.001.

Table 3 True positive, false positive, true negative, false negative, PLR, NLR, AI of all studied

	All (N =150) (%)	Female (N=102)	Male (N=48)
True Positive	44 (29.33%)	28.00	18.00
False Positive	14 (9.3%)	10.00	4.00
True Negative	77(51.33)	53.00	24.00
False negative	15 (10%)	11.00	2.00
PLR	4.847		
NLR	0.300		
AI (%)	0.807		

PLR: Positive Likelihood Ratio, NLR: Negative Likelihood Ratio, AI: Accuracy Index.

Table 4 Mean Rank of TRAb in correlation with different variables with *P* value

Characteristics		N (Total = 150)	TRAb mean rank	<i>P</i> value
Age	16-50	95	79.92	0.092
	51-78	55	67.87	
Gender	Female	102	77.35	0.434
	Male	48	71.56	
Posttest	Graves	59	104.15	<0.001
	non-graves	91	56.92	
Clinical Goiter	Goiter	101	74.14	0.382
	No Goiter	49	79.82	

Discussion

Serum TSH receptor antibody measurement has the highest sensitivity and specificity of any single blood test used to evaluate suspected hyperthyroidism and should be used as an initial screening test.⁵

The diagnosis of hyperthyroidism is based on characteristic clinical features and biochemical abnormalities. If pathognomonic features such as ophthalmopathy or dermopathy are absent and a diffuse goiter is not detected, radionuclide scanning can confirm the diagnosis.⁸

We studied a population of cases presenting with clinical features of thyrotoxicosis, assessing the clinical diagnosis against serum TRAb. In this study, we found high sensitivity and specificity for clinical diagnosis, depending on specific clinical features of Graves' disease and other preliminary tests such as ultrasound of thyroid and thyroid scintigraphy (in a few cases).

TSH Receptor Antibodies are specific biomarkers for Graves' disease.^{9,10}

A meta-analysis of 21 studies showed that the overall pooled sensitivity and specificity of the serum TRAb concentration measured with second- and third-generation binding assays were 97 and 98%, respectively.⁷

TSH-R-Ab are specific biomarkers for Graves' disease. The measurement of TRAb is a sensitive and specific tool for rapid and accurate diagnosis and differential diagnosis of Graves' hyperthyroidism.¹¹

Another retrospective study examined the sensitivity and specificity of thyrotropin receptor antibody immunoassays test in diagnosing Graves' disease and non-Graves' disease hyperthyroidism, using thyroid scintigraphy and ultrasound of thyroid gland compared with TRAb testing. The study concluded that thyroid scintigraphy remains the most accurate method to differentiate causes of thyrotoxicosis. However, TRAb assays can be alternatively adopted in this setting,

limiting the use of thyroid scintigraphy [Technetium Pertechnetate Thyroid Uptake (TcTU evaluation)] to TRAb negative patients. Thyroid ultrasound is less accurate than both TRAb/TSI and thyroid scintigraphy, but the 'thyroid inferno' pattern provides a high PPV for Graves' disease.¹²

Many studies examined the sensitivity and specificity of TRAb in the diagnosis of Graves' disease and differentiating it from non-Graves' disease disorders.^{5,13,14}

True Graves' disease hyperthyroidism cannot occur without TRAb. Graves' disease is almost unique among autoimmune diseases in that the most important clinical manifestation of the disease, hyperthyroidism, is entirely dependent on, and completely recapitulated by, the interaction of an autoantibody with its autoantigen. Hence, testing for the TSHR antibody should be particularly useful in the diagnosis of Graves' disease hyperthyroidism.¹⁵

Another study that compared the performance characteristics of different TBII and bioassays for TRAbs reported 100% specificity for all the assays. The likelihood of a TRAb-positive individual having Graves' disease was >1000 to >3000 fold greater (depending on the assay type) than a TRAb-negative person.¹⁶

Despite this simple concept, TRAb is not always used in the United States as a first-line test in the differential diagnosis of hyperthyroidism. For example, in their joint guidelines, the American Thyroid Association (ATA) and the American Association of Clinical Endocrinologists indicate a thyroid scan as the primary differential diagnostic test.⁶

In establishing the diagnosis of Graves' disease and differentiating it from other causes of thyrotoxicosis, some argue that TRAb assays are not necessary to diagnose Graves' disease and for its differential diagnosis from other causes of thyrotoxicosis.¹⁶

The British Thyroid Association

recommends testing for TRAb in special situations only.¹⁷ In contrast, many European clinicians and investigators advocate the use of TRAb as the primary test in the initial workup of hyperthyroidism.¹

Although some experts doubt its value in subjects with typical features of Graves' disease, we believe that TRAb assays should be done in all patients to positively establish a diagnosis and to help in differentiating between the various causes of thyrotoxicosis.¹⁸

In our study, we found high sensitivity and specificity of clinical diagnosis in the differentiation of Graves' disease and non-Graves' disease diseases, assuming the high specificity and specificity of TRAb test as a standard test for diagnosis of Graves' disease.

In another study by Bell L, Hunter AL reported that the clinical diagnosis of Graves' disease or non-Graves' disease hyperthyroidism has remarkably poorer sensitivity and specificity than the TRAb test. They concluded that the clinical diagnosis of Graves' disease or non-Graves' disease hyperthyroidism may not reliably identify patients with or without Graves' disease, respectively, and that testing for TRAb is of value in the initial clinical assessment of all patients presenting with hyperthyroidism.⁵ It was a retrospective study with a larger number of Graves' disease cases reported than our study, which was a cross-sectional study. The number of Graves' disease cases was less than non-Graves' disease cases.

In our study, sensitivity was 74.60% and specificity 84.60%, indicating high sensitivity and specificity of clinical diagnosis of Graves' disease or non-Graves' disease hyperthyroidism, supported by other tests. Apart from thorough clinical examination, the likely cause in this case is possibly due to increased use of some preliminary diagnostic tools such as ultrasound scan, TPO Abs, and even, in a few cases, the availability of radio uptake iodine scans.

However, TRAb measurements using modern 2nd-3rd generation receptor assays are increasingly more freely available, quickly done, and cheap (certainly in high volume laboratories).³

They offer a greater advantage over TPOAb and thyroid scintigraphy in terms of higher sensitivity and specificity, logistical considerations, and cost savings.¹⁸ Furthermore, newer automated 3rd-generation assays provide excellent sensitivity and specificity with high PPV and NPV in subjects with biochemical hyperthyroidism.¹⁹

For that TRAb test has superiority over the other tests, for being easily available in most laboratory centers, relatively cheap, technically more accessible, compared with radioactive iodine or Technetium 99 scintigraphy, which is not accessible in many medical centers, as in our locality.

The overall pretest Graves' disease and non-Graves' disease cases were the same, but true positive, true negative, false positive, and false negative rates were different posttest. The true positive cases were 44, false negatives were 15, true negatives were 77, and false positives were 14.

In this study, the mean TRAb in the Graves group was significantly higher than in the non-Graves group ($P < 0.001$). Our work did not study the differential pathologies of non-Graves' disease, such as toxic nodules, toxic multinodular goiter, and different types of thyroiditis.

Conclusion

We found that clinical diagnosis supported with primary investigations has high sensitivity and specificity for the diagnosis of Graves' disease. However, other preliminary tests such as radioactive thyroid uptake scanning, specific ultrasound scans, or Doppler study of the thyroid gland are to be available. While the gold standard for thyroid-stimulating immunoglobulins is the bioassay serum thyroid-stimulating immunoglobulin [TSI], the TRAb test remains the more easily

available in most laboratory centers, less expensive, with high sensitivity and specificity in the differential diagnosis of Graves' disease. We conclude to use the TRAb test to diagnose Graves' disease despite the sensitivity and specificity of clinical diagnosis for the reasons mentioned above.

We suggest studying a larger number of patients with suspected Graves' disease, comparing TRAb test with RAIU testing, and Doppler study of the thyroid gland.

Funding

Not applicable.

Competing interests

The author declares that he has no competing interests.

References

1. Kahaly GJ, Bartalena L, Hegedüs L. The American Thyroid Association/American Association of Clinical Endocrinologists guidelines for hyperthyroidism and other causes of thyrotoxicosis: A European perspective. *Thyroid*. 2011;21(6):585–91. <https://doi.org/10.1089/thy.2011.2106.ed3>.
2. Anthony H, Wilmar MW. Hyperthyroid Disorders. In: Shlomo M, Richard JA, Allison BG, editors. *Williams Textbook of Endocrinology*. 14th ed. Philadelphia: Elsevier; 2020. P. 452–3.
3. Matthews DC, Syed AA. The role of TSH receptor antibodies in the management of Graves' disease. *Eur J Intern Med*. 2011;22(3):213–6. <https://doi.org/10.1016/j.ejim.2011.02.006>.
4. Vasileiou M, Gilbert J, Fishburn S, Boelaert K; Guideline Committee. Thyroid disease assessment and management: summary of NICE guidance. *BMJ*. 2020;368:m41. <https://doi.org/10.1136/bmj.m41>.
5. Bell L, Hunter AL, Kyriacou A, Mukherjee A, Syed AA. Clinical diagnosis of Graves' or non-Graves' hyperthyroidism compared to TSH receptor antibody test. *Endocrine Connections*. 2018;7(4):504–10. <https://doi.org/10.1530/EC-18-0082>.
6. Bahn Chair RS, Burch HB, Cooper DS, Garber JR, Greenlee MC, Klein I, et al. Hyperthyroidism and other causes of thyrotoxicosis: management guidelines of the American Thyroid Association and American Association of Clinical Endocrinologists. *Thyroid*. 2011;21(6):593–646. <https://doi.org/10.1089/thy.2010.0417>.
7. Tozzoli R, Bagnasco M, Giavarina D, Bizzaro N. TSH receptor autoantibody immunoassay in patients with Graves' disease: improvement of diagnostic accuracy over different generations of methods. Systematic review and meta-analysis. *Autoimmun Rev*. 2012;12(2):107–13. <https://doi.org/10.1016/j.autrev.2012.07.003>.
8. Smith TJ, Hegedüs L. Graves' Disease. *N Engl J Med*. 2016;375(16):1552–65. <https://doi.org/10.1056/NEJMra1510030>.
9. Bartalena L. Diagnosis and management of Graves' disease: a global overview. *Nat Rev Endocrinol*. 2013;9(12):724–34. <https://doi.org/10.1038/nrendo.2013.193>.
10. Bartalena L, Burch HB, Burman KD, Kahaly GJ. A 2013 European survey of clinical practice patterns in the management of Graves' disease. *Clin Endocrinol*. 2016;84(1):115–20. <https://doi.org/10.1111/cen.12688>.
11. Kahaly GJ, Bartalena L, Hegedüs L, Leenhardt L, Poppe K, Pearce SH. European Thyroid Association Guideline for the Management of Graves' Hyperthyroidism. *Eur Thyroid J*. 2018;7(4):167–86. <https://doi.org/10.1159/000490384>.
12. Scappaticcio L, Trimboli P, Keller F, Imperiali M, Piccardo A, Giovanella L. Diagnostic testing for Graves' or non-Graves' hyperthyroidism: A comparison of two thyrotropin receptor antibody immunoassays with thyroid scintigraphy and ultrasonography. *Clin Endocrinol (Oxf)*. 2020;92(2):169–78. <https://doi.org/10.1111/cen.14130>.
13. Meena MQ. Diagnosing Graves' Disease and non-Graves hyperthyroidism using TSH receptor antibody test versus non-TSH receptor antibody test methods of diagnosis. *Open Journal of Endocrine and Metabolic Diseases*. 2020;10(2):7–17. <https://doi.org/10.4236/ojemd.2020.102002>.
14. Diana T, Wüster C, Kanitz M, Kahaly GJ. Highly variable sensitivity of five binding and two bio-assays for TSH-receptor antibodies. *Journal of Endocrinological Investigation*. 2016;39(10):1159–65. <https://doi.org/10.1007/s40618-016-0478-9>.
15. Barbesino G, Tomer Y. Clinical utility of TSH receptor antibodies. *J Clin Endocrinol Metab*. 2013;98(6):2247–55. <https://doi.org/10.1210/jc.2012-4309>.
16. Kamath C, Adlan MA, Premawardhana LD. The role of thyrotrophin receptor antibody assays in graves' disease. *J Thyroid Res*. 2012;2012:525936. <https://doi.org/10.1155/2012/525936>.
17. Beastall GH, Beckett GJ, Franklyn JA, Fraser WD, Hickey J, John R, et al. British Thyroid Association. U K Guidelines for the use of thyroid function tests. (Accessed October 2, 2020, available at <http://www.british-thyroid-association.org>).
18. Syme NR, Toft AD, Stoddart M, Beckett GJ. Clinical performance of the Roche Cobase411 automated assay system for thyrotropin-receptor antibodies for the diagnosis of Graves' disease. *Ann Clin Biochem*. 2011;48(5):471–3. <https://doi.org/10.1258/acb.2011.011013>.

19. Rugge JB, Bougatsos C, Chou R. Screening and treatment of thyroid dysfunction: an evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2015;162(1):35–45. <https://doi.org/10.7326/M14-1456>.