

Clinical scoring scales in thyroidology: A compendium

Sanjay Kalra^{1,2}, Sachin K. Khandelwal², Aakshit Goyal³

Department of Endocrinology, ¹Bharti Hospital and ²BRIDE, Karnal, Haryana, ³Government Medical College, Patiala, Punjab, India

ABSTRACT

This compendium brings together traditional as well as contemporary scoring and grading systems used for the screening and diagnosis of various thyroid diseases, dysfunctions, and complications. The article discusses scores used to help diagnose hypo- and hyperthyroidism, to grade and manage goiter and ophthalmopathy, and to assess the risk of thyroid malignancy.

Key words: Goiter, grading system, malignancy, ophthalmopathy, scoring system, Thyroid dysfunction, Thyroid screening

INTRODUCTION

Clinical scores have traditionally been used in thyroidology to help in the diagnosis of thyroid dysfunction. The time honored Wayne's score and Billewicz score are well known to older generations of endocrinologists, for their utility in the clinical diagnosis of hyperthyroidism and hypothyroidism, respectively.

Sadly, however, these scores do not find place in current textbooks of surgery or medicine, having been discarded as being too old fashioned. Current students of medicine therefore do not have easy access to these, and to newer scoring systems in thyroidology.

Newer clinical scores and grading systems have also been developed by workers, highlighting the numerous diagnostic developments and advances in the field. Results from imaging studies such as ultrasound are also used in grading thyroid malignancy.

This compendium seeks to bring together, and comment upon, the various scoring and grading systems used in clinical endocrinology. It is hoped that this collation will be of use to endocrinologists, physicians, and all workers in the field of thyroidology.

GOITER

The diagnosis of goiter is a clinical one arrived at by inspection and palpation. The presence of goiter does not necessarily indicate thyroid dysfunction.

The World Health Organization has classified goiter in a simple, objective manner, into three grades.^[1] This classification can be used to grade, compare, and monitor thyroid enlargement [Table 1], with minimal inter- and intraobserver bias.^[2]

HYPERTHYROIDISM

Grave's disease is used to describe the diffuse hyperplasia of the thyroid gland. The signs and symptoms of Grave's disease can be classified into three groups.^[3]

The first groups of symptoms occur because of hyperfunctioning of the thyroid gland and increased catecholamine sensitivity. The second group related to the enlargement of the thyroid, i.e., goitre and its accompanying signs such as thyroid bruit. The third constellation of symptoms and signs is the eye signs, which are a long list

Access this article online

Quick Response Code:



Website:
www.ijem.in

DOI:
10.4103/2230-8210.83332

Corresponding Author: Dr. Sanjay Kalra, BRIDE, Kunjpura Road, Karnal – 132 001, Haryana, India. E-mail: brideknl@gmail.com

Table 1: WHO classification of Goiter

1. Grade 0 – no goiter presence is found (the thyroid impalpable and invisible)
2. Grade 1 – neck thickening is present in result of enlarged thyroid, palpable, however, not visible in normal position of the neck; the thickened mass moves upwards during swallowing. Grade 1 includes also nodular goiter if thyroid enlargement remains invisible
3. Grade 2 – neck swelling, visible when neck is in normal position, corresponding to enlarged thyroid- found in palpation

of eponymous conditions, dreaded by medical students. In addition, there are miscellaneous clinical features, affecting the skin, muscles and other systems.^[3] These symptoms and signs have been used to prepare a simple clinical score for diagnosis and screening.

Wayne’s Index is more than half a century old,^[4] but has shown remarkable success in helping diagnose hyperthyroidism [Table 2].

Nine symptoms and 10 signs are listed, each with differential weightage in scoring. The signs are scored both positively and negatively, i.e., absence of some signs (but not all) gives negative marking. Similarly, two symptoms, i.e., decreased appetite and preference for heat, have negative scores.

The score ranges from + 45 to –25. A score greater than 19 implies toxic hyperthyroidism, while a score less than 11 implies euthyroidism, A score between 11 and 19 is equivocal. Though arrived at by trial and error, it has shown a diagnostic accuracy of 85%.

Wayne’s Index was earlier used to help to diagnose hyperthyroidism and limit the number of investigations required.^[4] At present, its value lies in its emphasis on clinical thyroidology, and its utility in explaining the clinical features of Grave’s disease to thyroid practitioner.

OPHTHALMOPATHY

The clinical activity score (CAS), for Grave’s ophthalmopathy, published in 1997,^[5] has become a widely accepted tool to help decide the management of the condition. Grave’s ophthalmopathy is a biphasic disease, with an initial phase of active inflammation, followed by a “burnt out” phase with stable proptosis and impaired eye muscle mobility.

The CAS, which is based on four classical signs of inflammation (pain, redness, swelling and impaired function), consists of 10 equally weighted items [Table 3]. The total CAS may range from 0 to 10. The higher the CAS, the greater is the response to immunosuppression.

Table 2: Wayne’s Index- Showing the scoring of signs and symptoms for the diagnosis of hypothyroidism

Symptoms of recent onset and/or increased severity	Score	Signs	Present Absent	
			Present	Absent
Dyspnea on effort	+1	Palpable thyroid	+3	-3
Palpitations	+2	Bruit over thyroid	+2	-2
Tiredness	+2	Exophthalmoses	+2	-
Preference for heat	-5	Lid retraction	+2	-
Preference for cold	+5	Lid lag	+1	-
Excessive sweating	+3	Hyperkinesis	+4	-2
Nervousness	+2	Hands hot	+2	-2
Appetite: increased	+3	Hands moist	+1	-1
Appetite: decreased	-3	Casual pulse rate: >80/min	-	-3
Weight increased	-3	>90/min	+3	-
Weight decreased	+3	Atrial fibrillation	+4	-

Table 3: Point system used for the clinical activity score

Pain	1	Painful, oppressive feeling on or behind the globe during the last 2 weeks
	2	Pain on attempted up, side or down gaze during the last 4 weeks
Redness	3	Redness of the eyelids
	4	Diffuse redness of the conjunctiva covering at least one quadrant
Swelling	5	Swelling of eyelids
	6	Chemosis
	7	Swollen caruncle
	8	Increase of proptosis ≥ 2 mm during a period of 1-3 months
Impaired function	9	Decrease of eye movements in any direction $\geq 5^\circ$ during a period of 1-3 months
	10	Decrease of visual acuity of ≥ 1 line on the Snellen chart (using a pin hole) during a period of 1-3 months

A CAS ≥ 4 implies an active inflammatory stage of Grave’s ophthalmopathy. Using this CAS cut-off, a specificity of 86%, sensitivity of 55%, positive predictive value of 80%, and a negative predictive value of 64% have been reported in predicting therapeutic outcome. CAS does not differ in duration of disease.^[5]

CAS is able to predict therapeutic outcome based on classical signs and symptoms of inflammation. The sign of “heat” is not used in CAS as it is difficult to detect subtle rises of orbital temperature without special instruments.

A study on CAS^[5] found that intraobserver variance was minimal for pain- related symptoms. It was higher for redness of the conjunctiva, which should be diffuse, and cover at least one quadrant to be defined as inflammatory. The CAS is purely clinical, and helps select appropriate therapy for patients with Grave’s ophthalmopathy.

The NOSPECS classification has been used for over a decade to grade the changes seen in thyroid ophthalmopathy

in an objective manner. The NOSPECS scores the disease based on soft tissue involvement, corneal involvement, and sight loss [Table 4].^[6]

It is an objective method of assessing disease progression, rather than inflammation. It can be used to grade and monitor patients with minimal interobserver variability.

The NOSPECS classification is also a simple mnemonic which helps emphasize the clinical features of thyroid disease, while making it simple for medical students to remember them.

In passing, one may mention another mnemonic related to thyroid disease. The TEARS mnemonic is useful for remembering the initial management of the condition [Table 5].^[7]

HYPOTHYROIDISM

The Billewicz score utilizes 8 symptoms and 6 signs to assess the thyroid status, and diagnose hypothyroidism. In a detailed study on 256 euthyroid controls, and hypothyroid patients, 13 symptoms and 8 signs were studied. Initially 14 symptoms and signs were selected and weighted differentially, based on the frequency of their occurrence in hypothyroid patients [Table 6].^[8]

The score may range from + 67 to – 47, with the highest weightage being given to a sluggish ankle jerk and slow movements. Billewicz *et al.* do not forget to prescribe an effective manner in which to elicit these signs and symptoms, in order to minimize interobserver variability.^[8]

Diminished sweating is assessed in a warm room or a centrally heated hall (Billewicz practiced in Aberdeen, Scotland). Dry skin is defined as dryness of skin noted spontaneously, or requiring treatment. Cold intolerance implies a preference for a warm room, extra clothing, or bed clothing. The weight increase is scored as present if the patient reports a recorded increase in weight or complains of tightness of clothing. Similarly, constipation is scored as present if the patient reports a change in bowel habit or use of laxative. Hoarseness is assessed in both speaking voice and singing voice, while paresthesia are scored based on subjective sensations. Deafness is defined as progressive improvement of hearing.

Slow movements are noted while observing the patient removing and replacing a buttoned garment. Correction of skin is assessed over the hands, forearms, and elbows, with the examiner checking for roughness and thickening of skin. Cold skin is assessed by comparing the patient's

Table 4: NOSPECS classification

Score	Grade	Change
0		No signs and symptoms
1		Only Signs
2		Soft tissue involvement, with symptoms and signs
o	Absent	
a	Minimal	
b	Moderate	
c	Marked	
3		Proptosis
o	<23 mm	
a	23-24 mm	
b	25-27 mm	
c	≥28 mm	
4		Extraocular muscle involvement
o	Absent	
a	Limitation of motion in extremes of gaze	
b	Evident restriction of movement	
c	Fixed eye ball	
5		Corneal involvement
o		Absent
a	Stippling of cornea	
b	Ulceration	
c	Clouding	
6		Sight loss
o	Absent	
a	20/20- 20/60	
b	20/70- 20/200	
c	<20/200	

Table 5: TEARS mnemonic for remembering initial management

- T- Tobacco abstinence is immensely important.
- E- Euthyroidism must be achieved and maintained.
- A- Artificial tears are helpful for the majority of the patients and can afford rapid relief from symptoms of corneal exposure.
- R- Referral to a specialist center with experience and expertise in treating thyroid eye disease is indicated in all but the mildest of cases.
- S- Self-help groups can provide valuable additional support.

Table 6: Billewicz diagnostic index^[6]

	Present	Absent
Symptom		
Diminished sweating	+ 6	- 2
Dry skin	+ 3	- 6
Cold intolerance	+ 4	- 5
Weight increase	+ 1	- 1
Constipation	+ 2	- 1
Hoarseness	+ 5	- 4
Deafness	+ 2	0
Signs		
Slow movements	+ 11	- 3
Coarse skin	+ 7	- 7
Cold skin	+ 3	- 2
Periorbital puffiness	+ 4	- 6
Pulse rate	+ 4	- 4
Ankle jerk	+ 15	- 6

hand temperature with that of the examiners. Periorbital puffiness is defined if it obscures the curve of the malor bone. Billewicz *et al.* recommend counting the pulse for a 30-s period, and report bradycardia of the pulse <75/min. They elicit the ankle jerk with the patient kneeling on a chair, grasping its back.^[8]

A score of +25 or more suggests hypothyroidism, while a score of -30 or less excludes the disease.^[8] The score can be used as a screening tool, especially in psychiatric patients. The utility of the Billewicz score lies in its ability to teach students about the clinical features of the disease.

A simple postal questionnaire has also been used to assess in the follow-up of patients treated with radioactive iodine.^[9] This is a cost-effective method of screening for hypothyroidism, and can help increase the rate of detection of this condition in patients at risk [Table 7].

Initially designed as a seven-question list, it was later expanded to nine questions, all to be answered as yes or no. Patients with symptoms are investigated for thyroid function, while asymptomatic ones may have investigations deferred. This leads to considerable cost-saving.^[9]

The Billewicz score and the postal questionnaire mentioned above were studied at a time when modern methods of diagnosing hypothyroidism were not available.

Zulewski *et al.* [Table 8] set out to reevaluate the classical signs and symptoms of hypothyroidism in the light of modern laboratory tests. They measured clinical scores, thyroid function, and tissue thyroid status (using ankle reflex relaxation time [ART] and total cholesterol, in 50 hypothyroid, 93 subclinically hypothyroid, 67 treated hypothyroid, and 189 euthyroid female adults.^[10]

The 14 symptoms and signs identified by Billewicz *et al.*

Table 7: Postal formulary used to elicit symptoms of hypothyroidism

Please answer yes or no to the following questions:

1. Do you feel as well as you did a year ago?
2. Do you now feel the cold more than ever before, so that you cannot get properly warm?
3. Is your appetite as good as it was a year ago?
4. Do you feel less energetic than usual?
5. Do you think you have put on weight in the last year?
6. Have you, or any of your family or friends, noticed that your voice has recently become huskier or weaker?
7. Are you getting fuller in the face?
8. Has the skin of your arms or legs becomes more dry or rough during the past year?
9. Has your hair recently become unruly or more difficult to manage?

were evaluated. Two features, i.e., pulse rate and cold intolerance, had positive and negative predictive values below 70%, and were excluded.

The most sensitive features were delayed ART (77%) and dry skin (76%), while the most specific were slow movements (98.7%) and diminished hearing (97.5%). A positive predictive value was highest for slow movements (96.5%) and puffiness (94.2%). On the other hand, a negative predictive value was highest for ART (80.3%) and dry skin (72.7%).

As women aged > 55 years also complained of “hypothyroid” symptoms, especially constipation and dry skin, an age-correcting factor was added. One point was added to the sum of symptoms and signs in younger women (aged < 55 years).^[10]

A score >5 points defined hypothyroidism, while a score of 0-2 points defined euthyroidism. Sixty two percent of all overt hypothyroidism was detected by the new score (as compared to 42% with the Billewicz score).

The new score was higher in overt hypothyroid smokers than in nonsmokers (7.6 ± 3.2 vs. 5.5 ± 2.7 ; $P = 0.024$). The score demonstrated excellent correlation with tests

Table 8: Zulewski’s clinical score for hypothyroidism

On the basis of			New score	
			Present	Absent
Symptoms				
1	Diminished sweating	Sweating in the warm room or a hot summer day	1	0
2	Hoarseness	Speaking voice, singing voice	1	0
3	Paraesthesia	Subjective sensation	1	0
4	Dry Skin	Dryness of skin, noticed spontaneously, requiring treatment	1	0
5	Constipation	Bowel habit, use of laxative	1	0
6	Impairment of hearing	Progressive impairment of hearing	1	0
7	Weight increase	Recorded weight increase, tightness of clothes	1	0
Physical signs				
1	Slow movements	Observe patient removing his clothes	1	0
2	Delayed ankle reflex	Observe the relaxation of the reflex	1	0
3	Coarse Skin	Examine hands, forearms, elbow for roughness and thickening of skin	1	0
4	Peritorbital puffiness	This should obscure the curve of the malar bone	1	0
5	Cold skin	Compare temperature of hands with examiner’s	1	0
Sum of all symptoms and signs present			12	0

of tissue hypothyroidism ($r = 0.76$ for ART, $P < 0.0001$; $r = 0.060$ for total cholesterol, $P < 0.00001$; $r = 0.55$ for creatine kinase). Surprisingly, no correlation was observed with serum TSH levels ($r = 0.01$; $P = \text{NS}$).^[10]

The clinical score for tissue hypothyroidism deserves wider usage and attention. It has highlighted the common clinical features of the condition, while emphasizing the changes in presentation that have occurred over the past few decades. The score has underscored the concept of tissue hypothyroidism, given an easy method of assessing its severity. It can be used to evaluate patients with discordant laboratory results, and to monitor effects of therapy.^[10]

Another simple questionnaire known as the Thyroid Symptom Questionnaire (TSQ) has been used to detect^[11] how patients feel on medication. The TSQ questions are derived from symptoms reported by patients to the British Thyroid Foundation Newsletter.^[12] The 12 questions are listed in Table 9.

In a group of 597 hypothyroid subjects, and 551 euthyroid controls, significant differences were seen in TSQ screening. A total of 35.0% of controls, 46.8% of patients, and 48.6% of euthyroid patients scored ≥ 3 on the TSQ ($P < 0.001$ for patients vs. controls, and for euthyroid patients vs. controls).

This community-based study showed that the TSQ can be used to assess well-being in hypothyroid patients on thyroxine replacement. A larger number of patients report dissatisfaction while on treatment. A total of 46.8% of all patients had a TSQ ≥ 3 , while 37.2% and 25.9% scored ≥ 4 and ≥ 5 respectively. These values were significantly higher than the 35.0%, 24.9%, and 17.6% of controls who scored similar grades on the TSQ.

Symptoms with highest differences between patients and

controls were “remembering things” ($P = 0.014$ all patients vs. controls), “can’t think of the right word” ($P = 0.009$), “felt tired and lethargic” ($P = 0.001$). “Able to think clearly” ($P = 0.007$), and “clumsy, bumped into things and fallen over” ($P = 0.034$).^[11]

THYROID MALIGNANCY

A grading system for ultrasonologically detected thyroid nodules has recently been reported.

The TIRADS (Thyroid Imaging Reporting and Data System) grades thyroid nodules into six categories (TIRADS 2 to 6).^[13]

The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were 88, 49, 49, 88, and 94% respectively, when compared with fine-needle aspiration biopsy results.

This grading system can be used to avoid unnecessary interventional procedures. TIRADS 2 represents benign findings; TIRADS 3 is probably benign, while TIRADS 4A and 4B represent undetermined and suspicious findings respectively. An imaging picture consistent with malignancy is graded as TIRADS 5, while TIRADS 6 represents confirmed malignancy [Table 10].

The authors feel that patient management and cost-effectiveness improved with the TIRADS,^[13] which has established standard codes to be used in both endocrinology and radiology.

CONCLUSION

This compendium has tried to bring together clinical grading and scoring systems for the diagnosis of goiter, hypothyroidism, hyperthyroidism, and Grave’s ophthalmopathy. It highlights the old, as well as relatively modern, methods of clinical diagnosis in thyroidology. The compendium should be useful for medical students and practitioners, as well as physicians and endocrinologists.

The first thyroid score (based on thyroid function tests) was established by Schultz and Ziene in 1956 to predict relapse of thyrotoxic patients after radioactive iodine.^[14] The first clinical thyroid score was the index of Crooks, Murray and Wayne, created by trial and error in 1959,^[4] which we now know as Wayne’s Index. Thyroidology has grown immensely since then. Most of the diagnostic tools and clinical investigations which we use today will not be recognized by those who practised thyroid medicine long ago.

Table 9: Thyroid symptom questionnaire

Remember things?
Can't think of the right word?
Making careless mistakes?
Had palpitations?
Felt tired and lethargic?
Felt cold?
Putting on weight?
Having headaches?
Aches and pains all over the body?
Able to think clearly?
Clumsy, bumped into things and fallen over?
Difficulty in reading and spelling?
Answer: "better than usual" to "much less than usual" on a four point scale. Scores 0, 0, 1, 1.
Maximum score: 12

Table 10: US characteristics of thyroid nodules, 10 US with their malignancy risk, and TIRADS category

Description of US pattern	US patterns	Malignancy	TIRADS
Anechoic with hyperechoic spots, nonvascularized lesion	Colloid type 1		
Nonencapsulated, mixed, nonexpansile with hyperechoic spots, vascularized lesion, "grid" aspect (spongiform nodule)	Colloid type 2	0%	TIRADS 2: benign findings
Nonencapsulated, mixed with solid portion, isoechoic, expansile, vascularized nodule with hyperechoic spots	Colloid type 3		
Hyper-, iso-, or hypoechoic, partially encapsulated nodule with peripheral vascularization, in Hashimoto's thyroiditis	Hashimoto pseudonodule	5%	TIRADS 3: probably benign
Solid or mixed hyper-, iso-, or hypoechoic nodule, with a thin capsule	Simple neoplastic pattern	5-10%	TIRADS 4A: undetermined
Hypoechoic lesion with ill-defined borders, without calcifications	de Quervain pattern		
Hyper-, iso-, or hypoechoic, hypervascularized, encapsulated nodule with a thick capsule, containing calcifications (coarse or microcalcifications)	Suspicious neoplastic pattern		
Hypoechoic, nonencapsulated nodule, with irregular shape and margins, penetrating. Vessels, with or without calcifications	Malignant pattern A	10-80%	TIRADS 4B: suspicious
Iso- or hypoechoic, nonencapsulated nodule with multiple peripheral microcalcifications and hypervascularization	Malignant pattern B	>80%	TIRADS 5: consistent with malignancy
Nonencapsulated, isoechoic mixed hypervascularized nodule with or without calcifications, without hyperechoic spots	Malignant pattern C cancer, confirmed by the previous biopsy	100%	TIRADS 6: malignant

Some things, however, do not change. The need for clinical skills will remain paramount in the assessment and management of thyroid disease, and in fact, all medical conditions. As we complete 55 years of use of the first score in thyroid disease, this is a fitting tribute to the doyens of clinical thyroidology.

REFERENCES

- WHO/UNICEF/ICCIDD. Chapter 2: Selecting target groups and Chapter 5: Selecting appropriate indicators: Biochemical indicators. In: Indicators for Assessing Iodine Deficiency Disorder and their Control Through Salt Iodination. Geneva: World Health Organization; WHO/NUT/94.6, 1994.
- Delange F, Bastani S, Benmiloud M. Definitions of endemic goiter and cretinism, classification of goiter size and severity of endemias, and survey techniques. In: Dunn JT, Pretell EA, Daza CH, Viteri FE, editors. Towards the eradication of endemic goiter, cretinism, and iodine deficiency. Washington, DC: Pan American Health Organization; 1986. p. 373-6.
- Kendall-Taylor P. Hyperthyroidism. *BMJ* 1972;2:337-41.
- Crooks J, Murray IPC, Wayne EJ. Statistical methods applied to the clinical diagnosis of thyrotoxicosis. *Q J Med* 1959;28:211-34.
- Mourits MP, Koornneef L, Wiersinga WM, Prummel MF, Berghout A, van der Gaag R. Clinical criteria for the assessment of disease activity in Graves' Ophthalmopathy: A novel approach. *Br J Ophthalmol* 1989;73:639-44.
- Werner SC. Classification of the eye changes of Graves' disease. *Am J Ophthalmol* 1969;68:646-8.
- Krassas GE, Heufelder AE. Immunosuppressive therapy in patients with thyroid eye disease: An overview of current concepts. *Eur J Endocrinol* 2001;144:311-8.
- Billewicz WZ, Chapman RS, Crooks J, Day ME, Gossage J, Wayne E, *et al.* Stastical Methods applied to the diagnosis of hypothyroidism. *Q J Med* 1969;38:255-66.
- Barker DJ, Bishop JM. Computer- based screening system for patient at risk of hypothyroidism. *Lancet* 1969;2:835-8.
- Zulewski H, Muller B, Exer P, Miserez AR, Staub JJ. Estimation of tissue hypothyroidism by a new clinical score: Evaluation of patients with various grades of hypothyroidism and controls. *J Clin Endocrinol Metab* 1997;82:771-6.
- Saravanan P, Chau WF, Vedhra K, Greenwood R, Dayan CM. Psychological well-being in patients on 'adequate' doses of L-thyroxine: Results of a large, controlled community-based questionnaire study. *Clin Endocrinol* 2002;57:577-85.
- Roberts ND. Psychological problems in thyroid disease. *Br Thyroid Found News* 1996;18:3.
- Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, *et al.* An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab* 2009;94:1748-51.
- Schultz AL, Ziene L. Alterations in thyroid I-131 uptake, basal metabolic rate and serum cholesterol following treatment of hyperthyroidism with radioactive iodine: Value in early prediction of success or failure of therapy. *Am J Med* 1956;20:30-41.

Cite this article as: Kalra S, Khandelwal SK, Goyal A. Clinical scoring scales in thyroidology: A compendium. *Indian J Endocr Metab* 2011;15:S89-94.

Source of Support: Nil, **Conflict of Interest:** None declared.