Validation of Billewicz Scoring System for Detection of Overt Hypothyroidism During Pregnancy

Sima Nazarpour,[1,2] Fahimeh Ramezani Tehrani,[2,*] Maryam Rahmati,[2,3] Sonia Minooee,[2] Masoumeh Simbar,[4] Mahsa Noroozzadeh,[2] and Fereidoun Azizi[5]

[1]Department of Midwifery, Varamin-Pishva Branch, Islamic Azad University, Tehran, IR Iran
[2]Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, IR Iran
[3]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, IR Iran
[4]Midwifery and Reproductive Health Research Center, Department of Midwifery and Reproductive Health, School of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences, Tehran, IR Iran
[5]Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, IR Iran

*Corresponding author: Fahimeh Ramezani Tehrani, Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Velenjak St., Shahid Chamran Highway, Tehran, IR Iran. Tel: +98-2122439982, Fax: +98-2122439784, E-mail: ramezani@endocrine.ac.ir

Abstract

Background: Currently, various clinical and laboratory diagnostic methods are used to detect overt hypothyroidism during pregnancy. The Billewicz scoring index, as a clinical scale for detection of hypothyroidism, has been applied in general populations; however, its application during pregnancy remains a controversial subject of ongoing research.

Objectives: The purpose of this study was to evaluate the diagnostic value of Billewicz scoring index for overt hypothyroidism in a general population of Iranian pregnant women.

Methods: This study was conducted on 1843 pregnant women. A comprehensive questionnaire, including Billewicz scoring items, was completed, and relevant clinical examinations were performed. The participants underwent serum measurements of thyroxine (T4), thyroid hormone uptake, thyroid-stimulating hormone (TSH), and thyroid peroxidase antibody (TPOAb). Using the receiver operating characteristic (ROC) curve analysis, the optimal sensitivity and specificity were determined as values with maximum yields on the Youden and Rsquo's Index (sensitivity + specificity-1).

Results: The prevalence of overt hypothyroidism and subclinical hypothyroidism was 3.3% and 28.6%, respectively. Overall, 3.6%, 18.9%, and 50% of euthyroid, subclinical hypothyroid, and overt hypothyroid women were TPOAb-positive, respectively. The mean Billewicz scores of euthyroid, overt hypothyroid, and subclinical hypothyroid women were -41.16 (11.16), -17.11 (13.63), and -40.1 (11.2), respectively. Based on the Billewicz scale, at least one sign of hypothyroidism was reported in 38.84% (n, 491) of euthyroid women. Scores ≤ -26.5 (sensitivity, 100%; specificity, 90.82%) were considered as the optimal scores for predicting overt hypothyroidism (Ldquo, Norisk, and Rsquo).

Conclusions: The Billewicz clinical scoring system, as a reliable and inexpensive clinical tool, used along with laboratory measurements, can help screen overt hypothyroidism during pregnancy, primarily in low-resource settings.

Keywords: Hypothyroidism, Signs and Symptoms, Pregnancy

1. Background

Thyroid disorders are the second most frequent endocrine disorder in women of childbearing age (1). Hypothyroidism, with a prevalence of 2% - 5%, is the most common gestational thyroid dysfunction (2, 3). According to previous reports, the prevalence of hypothyroidism (both overt and subclinical) is estimated at 4.7% in Iranian pregnant women (4). Numerous studies have addressed the adverse effects of overt hypothyroidism (OH) on pregnancy outcomes (5-7), despite controversies about the adverse fetal-maternal effects of subclinical hypothyroidism

(SCH) (8). Proper screening can lead to early diagnosis of these disorders in pregnancy. However, universal screening for diagnosis of thyroid dysfunction in pregnancy is a controversial topic (9, 10).

Evidence shows that the selective high-risk case-finding approach may overlook a considerable number of gravid women with overt thyroid dysfunction (11, 12). As a result, having access to a convenient tool with an acceptable diagnostic value in detecting thyroid disorders during pregnancy has its merits. Apart from the universally advocated laboratory assessment of thyroid

function, several questionnaires have been designed for diagnosis of thyroid dysfunctions (13). These clinical tools are definitely not intended to replace blood measurements, but are reliable tools in low-income settings with limited access to costly tests; they can be also useful in posttreatment follow-ups.

Although validity of these clinical tools has been confirmed in general populations, no published study is available about their validation for pregnant women. Among various available tools, the Billewicz scoring index, as a cost-effective clinical index, was validated against thyroid iodine uptake and plasma protein-bound iodine in 1969 and against thyroid-stimulating hormone (TSH), triiodothyronine (T3), and free thyroxine (FT4) in 1997 (14, 15). Despite the convenience and accuracy of this tool in diagnosis of hypothyroidism in general populations, no information exists regarding Billewicz scale verification during pregnancy; therefore, the optimal pregnancy-related cutoff value has not been established yet.

In the present study, we aimed to assess the validity of Billewicz scale in a pregnant population, selected from Tehran thyroid and pregnancy study (TTAPS) to determine the optimal cut-off point of the Billewicz index with the highest sensitivity and specificity for predicting OH.

2. Methods

Data were collected from the first phase of TTAPS, which was carried out on pregnant women, attending the prenatal clinics of Shahid Beheshti Medical University. The details of TTAPS are published in the literature (16). All women were screened for thyroid dysfunction by collecting information from their medical history, clinical examinations, and serum thyroid peroxidase antibody (TPOAb), TSH, T4, and T-uptake measurements. They were then classified as follows: normal thyroid, OH, overt hyperthyroidism, TPOAb-positive euthyroidism, SCH, and subclinical hyperthyroidism. Subjects with twin pregnancies (n, 27) or history of chronic disorders were excluded from the study. Of a total of 1,843 pregnant women, 1,264, 38, and 541 cases exhibited euthyroidism, OH, and SCH, respectively.

In this study, the Billewicz scale was applied. This scale was developed based on seven symptoms and six signs, which were typically associated with hypothyroidism. According to this scale, higher positive scores indicate a higher level of clinical hypothyroidism. Items are weighed differently, based on their frequency of occurrence (see supplementary file, appendix 1) (13). A questionnaire related to the symptoms described in the Billewicz scale was completed, and a thorough physical examination of Billewicz signs (i.e., slow movement, coarse skin, cold skin,

periorbital puffiness, bradycardia, and delayed ankle reflex) was performed. After collection and centrifugation of fasting blood samples, the sera were sent to the Research institute of endocrine sciences of Shahid Beheshti University of Medical Sciences for laboratory assessments.

Using radioimmunoassay and immunoradiometric assay, T4 and TSH levels were measured, respectively by commercial kits (Izotop Kit, Budapest Co., Hungary) and a gamma counter (Dream Gamma-10, Goyang-si, Gyeonggi-do, South Korea). T-uptake and TPOAb were also determined by the enzyme immunoassay (Diaplus Kit, San Francisco, CA, USA) and immunoenzymometric assay (Monobind Kit, Costa Mesa, CA, USA), respectively, using a calibrated ELISA reader (Sunrise, Tecan Co., Salzburg, Austria). The inter- and intraassay coefficients of variation for T4, T-uptake, TSH, and TPOAb were 1.1%, 3.9%, 2.2%, and 4.3% and 1.9%, 4.7%, 1.0%, and 1.6%, respectively.

2.1. Definitions

Overt hyperthyroidism was defined as TSH level < 0.1 μIU/mL and FT4I > 4.5. OH was defined as TSH > 10 μIU/mL or TSH > 2.5 μIU/mL and FT4I < 1. SCH was defined as normal FT4I (1-4.5), despite elevated TSH (2.5-10 μIU/mL). Finally, subjects with TPOAb > 50 IU/mL were considered TPOAb-positive. It should be noted that when we first designed TTAPS and initiated data acquisition (from September 2013 to February 2016), the available guideline was the previous American thyroid association guideline (10). Accordingly, SCH was defined based on the common TSH cut-point of 2.5 μIU/mL instead of the recently recommended cut-point of 4.0 μIU/mL (17). In the Billewicz questionnaire, more than a 1.5-fold increase in weight (e.g., more than 8 kg in the second trimester for those with a normal prepregnancy BMI) was considered as abnormal weight gain. Also, bradycardia was defined as pulse rate < 75 bpm.

2.2. Ethics Approval

Written informed consents were obtained from all the participants, and the study was approved by the ethics committee of research institute of endocrine sciences (code, IR.SBMU.ENDOCRINE.REC.1396.383).

2.3. Statistical Analysis

Continuous variables were assessed for normality using one-sample Kolmogorov-Smirnov test. Categorical variables, expressed as percentages, were compared using Pearson's Chi square test. Normal continuous variables were compared between the three groups, using one-way ANOVA test and expressed as mean (standard deviation). Variables without a normal distribution were expressed as

median (interquartile) and compared using Kruskal-Wallis test.

First, we classified our participants into three groups, using the cut-off values reported by Billewicz and colleagues: score $\geq$ +25 for OH; score of -30 to +25 (above -30 and below +25) for SCH; and score $\leq$ –30 for excluding hypothyroidism. A receiver operating characteristics (ROC) curve analysis was performed, and the optimal cut-off value was identified, using the Youden index for differentiating OH from euthyroidism. In fact, the Youden index maximizes the difference between sensitivity and 1-specificity (Youden index: sensitivity + specificity-1). Therefore, by excluding SCH patients and maximizing sensitivity + specificity-1 across various cut-off points for the Billewicz index, the optimal cut-off point was calculated to distinguish between healthy subjects and patients with OH (18). Data were analyzed using SPSS version 22, and level of significance was set at P < 0.05.

3. Results

The mean (SD) age, BMI, and gestational age of the participants were 26.6 (5.3) years, 25.1 (4.6) kg/m^2, and 11.8 (4.1) weeks, respectively. Table 1 presents the characteristics of the study population according to their thyroid status. The median and interquartile range for TSH in euthyroid, OH, and SCH women were 1.46 (0.81 - 1.91), 7.92 (3.32 - 11.24), and 3.79 (3.06 - 4.96) μIU/mL, respectively. These values show that women with OH had significantly higher TSH levels, compared to the other two groups.

The mean±SD age of OH, euthyroid, and SCH women was 27.84 $\pm$ 5.01, 26.68 $\pm$ 5.30, and 26.32 $\pm$ 5.16 years, respectively. Euthyroid women had a significantly lower BMI, compared to OH and SCH groups (24.85 $\pm$ 4.53 versus 27.47 $\pm$ 5.47 and 25.62 $\pm$ 4.75). The prevalence of TPOAb positivity was 3.6%, 18.9%, and 50% in euthyroid, SCH, and OH groups, respectively. In addition, the median and interquartile range of TPOAb in euthyroid, OH, and SCH groups were 4 (2 - 8), 46 (5.5 - 445), and 7 (3 - 28) μIU/mL, respectively (Table 1).

Distribution of various risk factors for thyroid disorders showed that history of levothyroxine treatment and history of thyroid dysfunction were the two most common risk factors for OH in women. Also, age $\geq$ 30 years and thyroid antibodies, primarily TPOAb, were the most common risk factors for SCH (Table 2). Among euthyroid pregnant women, 491 (38.5%) exhibited at least one sign or symptom of the Billewicz index. Based on the findings, 257 (47.5%) and 34 (89.5%) SCH and OH women had more than (or equal to) one Billewicz sign or symptom, respectively. The mean $\pm$ SD Billewicz score was -17.11 $\pm$ 13.63 in the OH group,

which is significantly different from that of euthyroid (-41.16 $\pm$ 11.16; P < 0.001) and SCH (-40.1 $\pm$ 11.2; P < 0.001) women.

Weakness/fatigue and laziness/sleepiness were the two most common symptoms in all the groups (Table 3). In total, 34 out of 38 women with OH showed delayed ankle reflexes; this sign was observed in six out of 541 women with SCH and 10 out of 1,264 euthyroid pregnant women. In addition, pulse rate < 75/minute was observed in 23.7%, 21.8%, and 15.4% of women with OH, SCH, and euthyroidism, respectively (Table 3).

Using the cut-off point suggested by Billewicz et al. (score $\geq$ +25 for OH; -30 to +25 [> -30 and < +25] for SCH; and $\leq$ -30 for excluding hypothyroidism), we observed a significant difference in the TSH level of OH patients and those with SCH and euthyroidism. Based on these cut-off values, 474 (29.4%) out of 1613 participants were categorized as the normal group, while based on the TSH levels, they were categorized as SCH. Also, 125 (54.3%) and 38 (16.5%) out of 230 women, who were categorized as euthyroid and OH based on the TSH values, respectively, were diagnosed with SCH. On the other hand, no subject was diagnosed with OH based on the Billewicz score, as the maximum score of Billewicz index is +24.

The mentioned values corresponded to sensitivities of 12.38%, 100%, and 100% and specificities of 90.11%, 90.11%, and 87.61% for distinguishing euthyroidism from SCH, euthyroidism from OH, and SCH from OH, respectively. These values yielded positive predictive values of 34.89%, 23.31%, and 36.19% and negative predictive values of 70.61%, 100%, and 100%, respectively in discrimination between euthyroidism and SCH, euthyroidism and OH, and lastly SCH and OH.

By excluding the data of SCH patients and maximizing sensitivity + specificity-1 across various cut-off points of the Billewicz index, an optimal cut-off point of -26.5 was obtained for distinguishing between normal individuals and patients with OH. The AUC of the Billewicz index for predicting the absence of risk for OH was approximately 0.93 (95% CI, 0.92 - 0.95; P < 0.001). The Billewicz score $\leq$ -26.5 corresponded to a sensitivity of 100% and specificity of 90.82% and yielded a positive predictive value of 24.67% for distinguishing euthyroidism from OH (this cut-off point corresponded to a positive likelihood ratio of 10.89 and a negative likelihood ratio of 0) (Figure 1).

4. Discussion

The results of this study suggest that despite confusions in identifying thyroid disorders during pregnancy due to physiological changes, the Billewicz scoring index is a valid scale for diagnosis of gravid women with

Table 1. Characteristics of the Study Population According to Their Thyroid Status[a]

Characteristics	Euthyroidism (n, 1264)	Subclinical Hypothyroidism (n, 541)	Overt Hypothyroidism (n, 38)	P Value
Age, y[b]	26.68 ± 5.30	26.32 ± 5.16	27.84 ± 5.01	0.140
Body mass index, kg/m²[b]	24.85 ± 4.53[c]	25.62 ± 4.75[d]	27.47 ± 5.47	< 0.001
Gestational age, wk[e]	11 (8 - 14)[d]	12 (8 - 16)	10 (7 - 14)	< 0.001
Gestational age < 14 wk[f]	880 (69.81)	332 (61.5)[d]	25 (65.8)	0.003
Primigravida[f]	442 (35)	213 (39.4)	16 (42.1)	0.156
TSH[c]	1.46 (0.81 - 1.91)[d]	3.79 (3.06 - 4.96)[g]	7.92 (3.32 - 11.24)[c]	< 0.001
T4[e]	10.5 (8.8 - 12.5)	10.3 (8.7 - 12.4)	10.2 (8.4 - 13.02)	0.320
FTI[e]	2.9 (2.5 - 3.4)	2.8 (2.3 - 3.2)[d]	2.8 (2.3 - 3.22)	0.420
TPOAb[e]	4 (2 - 8)[d]	7 (3 - 28)[g]	46 (5.5 - 445)[c]	< 0.001
T3 uptake[e]	28 (25 - 30)[c]	27 (24 - 29)[d]	26.5 (23 - 29)	< 0.001

Abbreviations: FTI, free thyroxine index; TPOAb, thyroid peroxidase antibody; TSH, thyroid stimulating hormone.
[a]Values are expressed as mean ± SD, Median (IQR), or No. (%).
[b]ANOVA.
[c]Significant difference between overt hypothyroidism and euthyroidism.
[d]Significant difference between subclinical hypothyroidism and euthyroidism.
[e]Kruskal-Wallis test was used to compare the groups whenever needed.
[f]Chi square.
[g]Significant difference between subclinical hypothyroidism and overt hypothyroidism.

Table 2. Risk Factors for Thyroid Disorders in the Participants According to Their Thyroid Status[a]

Risk Factors	Euthyroidism (n, 1264)	Subclinical Hypothyroidism (n, 541)	Overt Hypothyroidism (n, 38)	P Value
History of thyroid dysfunction	26 (2.1)[b]	35 (6.5)[c]	22 (57.9)[d]	< 0.001
History of levothyroxine treatment	7 (0.6)[b]	1 (0.2)[c]	23 (60.5)[d]	< 0.001
History of treatment with radioactive iodine	2 (0.2)	0 (0.0)	0 (0.0)	-
History of therapeutic head or neck irradiation or thyroid surgery	0 (0.0)[b]	2 (0.4)	1 (2.6)[d]	-
Family history of autoimmune thyroid disease or thyroid	127 (10.1)[b]	76 (14)[c]	11 (28.9)[d]	< 0.001
Goiter	25 (2)[b]	6 (1.1)[c]	4 (10.5)	< 0.001
TPOAb	45 (3.6)[b]	102 (18.9)[c]	19 (50)[d]	< 0.001
Type I diabetes mellitus or other autoimmune disorders	13 (1)[b]	0 (0.0)	0 (0.0)	-
History of miscarriage	200 (15.9)	89 (16.5)	6 (15.8)	0.951
History of preterm delivery	21 (3.0)	8 (3.0)	0 (0.0)	0.758
Infertility	62 (4.9)	23 (4.3)[c]	8 (21.1)[d]	< 0.001
Morbid obesity (BMI ≥ 40)	7 (0.6)	1 (0.2)[c]	2 (6.1)[d]	-
Age ≥ 30 y	387 (30.7)	144 (26.6)	16 (42.1)	0.054

Abbreviation: TPOAb, thyroid peroxidase antibody.
[a]Values are expressed as No. (%). Chi square test was used to compare the groups. Significance level was set at P < 0.05.
[b]Significant difference between subclinical hypothyroidism and euthyroidism.
[c]Significant difference between subclinical hypothyroidism and overt hypothyroidism.
[d]Significant difference between overt hypothyroidism and euthyroidism.

OH. The optimal cut-off point was -26.5, with sensitivity of 100%, specificity of 90.82%, and positive predictive value of 24.67%. Our findings, validating the accuracy of Billewicz index during pregnancy, are novel, and to the best of our knowledge, no study has been published in this field so far.

We noted that delayed ankle reflex, fatigue, weakness, sleepiness, laziness, pulse rate < 75/min, constipation, and cold intolerance were the most common signs and symp-

Table 3. Signs and Symptoms of the Patients According to the Billewicz Index Based on Their Thyroid Status[a]

Symptoms and Signs	Euthyroidism (n, 1264)	Subclinical Hypothyroidism (n, 541)	Overt Hypothyroidism (n, 38)	P Value
Symptoms				
Diminished sweating	22 (1.7)	4 (0.7)	-	0.193
Dry skin	85 (6.7)	48 (8.9)	3 (7.9)	0.276
Cold intolerance	130 (10.3)	62 (11.5)	8 (21.1)	0.095
Weight gain	34 (2.7)	23 (4.3)	2 (5.3)	0.173
Constipation	159 (12.6)	91 (16.80)[b]	8 (21.1)[c]	0.026
Hoarseness	12 (3)	4 (2.7)	-	0.849
Hearing impairment	8 (2)	10 (6.8)[b]	-	0.015
Weakness/fatigue	407 (32.2)	161 (29.8)	14 (36.8)	0.463
Laziness/sleepiness	264 (20.9)	156 (28.8)[b]	14 (36.8)[c]	< 0.001
Signs				
Slow movement	36 (9.0)	13 (8.9)	3 (30)	0.077
Coarse skin	85 (6.7)	48 (8.9)	3 (7.9)	0.276
Periorbital puffiness	53 (12.7)	26 (16.3)	4 (30.8)	0.121
Cold skin	85 (6.7)	48 (8.9)	3 (7.9)	0.276
Pulse rate (< 75/min)	195 (15.4)	118 (21.8)[b]	9 (23.7)[c]	0.003
Delayed ankle reflex	10 (0.8)	6 (1.1)	34 (89.5)[c]	< 0.001

[a] Values are expressed as No. (%).Chi square test was used to compare the groups. The significance level was set at $P < 0.05$.

[b] Significant difference between subclinical hypothyroidism and euthyroidism.

[c] Significant difference between overt hypothyroidism and euthyroidism.
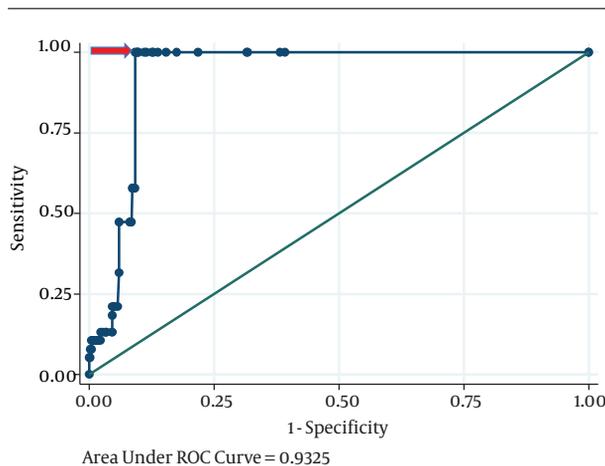
Figure 1. The ROC curve analysis for diagnosis of overt hypothyroidism (OH)

toms of OH. In accordance with our findings, in a study by Galia et al. the most frequent symptoms in hypothyroid non-pregnant patients were easy fatigue (64%), dyspnea on effort (52%), weight gain (44%), and constipation (44%), while the most common signs were rough and dry skin (36%), thyroid enlargement (32%), and sluggish movement (32%) (19). Almost similarly, Zulewski et al. reported weight gain (54%), constipation (48%), dry skin (76%), and bradycardia (58%) as the most common manifestations of hypothyroidism, which are comparable to the current study (15).

We also noticed that all signs and symptoms, except for three (diminished sweating, hoarseness, and hearing impairment), were present in both euthyroid and hypothyroid women. However, only four items, including laziness/sleepiness, constipation, ankle jerk reflex, and pulse rate < 75/m, were significantly higher in hypothyroid women, compared to their euthyroid counterparts; therefore, there is a lower risk of overlap with pregnancy-related physiological changes. Since physiological changes of pregnancy may mimic thyroid dysfunction signs, accurate diagnosis of hypothyroidism, based on the common signs and symptoms used in non-pregnant populations, is problematic (20). Meanwhile, laboratory measurements of thyroid function need to be interpreted cautiously during pregnancy, compared to non-pregnant women (17).

Screening of thyroid disorders during pregnancy has been a long-disputed issue. Universal screening and targeted high-risk case-finding have their own advantages and disadvantages (17). Overall, universal thyroid screening is a common and cost-effective approach; however,

there are still controversies about its application, since the prevalence of thyroid dysfunctions varies in different racial and ethnic populations (21); therefore, the significance of universal screening may be diminished among low-risk groups. In addition, there are still areas of uncertainty and disagreement among experts about the beneficial impact of maternal SCH diagnosis and management of poor pregnancy outcomes (22, 23).

A recent guideline of the American thyroid association declared that there is insufficient information to recommend or argue against universal screening during pregnancy, and TSH tests are recommended for women with the risk factors (17). On the other hand, the selective case-finding approach fails to identify about one-third of women with OH or SCH (24), which is considered a significant disadvantage for a screening test. Furthermore, challenging limitations, such as high cost of laboratory tests and endocrinologist visits, maternal anxiety, and absence of trimester-specific reference ranges, may necessitate validation of a preexisting reliable scale.

The Billewicz scale, appraised as one of the most cited thyroid scoring indices worldwide (25), has the potential to fill the mentioned gaps. This index serves as an inexpensive and accurate tool, basically for high-risk populations from low- and middle-income countries with limited access to costly laboratory tests. We agree that despite the strong association between clinical symptoms and abnormal TSH level (26), the accuracy of clinical diagnosis is limited, and laboratory diagnosis is also needed. Yet, we emphasize that the Billewicz index is not intended to replace sensitive biochemical measurements, but is a useful tool in detecting high-risk women prior to serum testing and can be applicable for measuring the patient's response to medical treatment.

While the consequences of OH for pregnancy outcomes and neonatal/child development are well-established (5, 6), findings regarding the relationship between maternal SCH and adverse pregnancy outcomes in both mothers and neonates are contradictory. Current evidence does not support the lower level of intellectual development in children born to mothers with SCH (27), and data on the beneficial effects of thyroxine supplementation on SCH women are inconsistent (28). Therefore, identifying OH during pregnancy screening is a principal issue. Undoubtedly, undetected OH is associated with increased maternal and neonatal complications (5, 6, 29). However, even in the event of severe maternal hypothyroidism, neurocognitive deficiencies in the child may not occur once the condition is managed immediately in the first half of pregnancy; therefore, use of the Billewicz scale can decrease undiagnosed OH during pregnancy screening.

This study has one limitation. Instead of using the newly recommended limit of 4.0 μIU/mL, we defined SCH based on the TSH cut-off point of 2.5 μIU/mL according to the available guideline during study design and conception (September 2013 to February 2016). Considering the difference between the previous (12) and new (17) guidelines, a number of our euthyroid subjects might be classified as SCH patients. Euthyroid women had a significantly lower BMI, compared to the OH and SCH groups. However, this difference is not significant enough to influence the cut-off points; accordingly, there is no need to present age-BMI specific cut-off points.

4.1. Conclusions

In summary, this is a novel study, validating the Billewicz scale and identifying the cut-off points during pregnancy. We found that the Billewicz index could be simply used as an auxiliary diagnostic tool in resource-constrained settings. However, future studies should include pregnant women from different ethnic or geographic (in terms of iodine sufficiency) populations.

Supplementary Material

Supplementary material(s) is available here [To read supplementary materials, please refer to the journal website and open PDF/HTML].

Acknowledgments

References

1. Carney LA, Quinlan JD, West JM. Thyroid disease in pregnancy. *Am Fam Physician*. 2014;**89**(4):273–8. [PubMed: 24695447].
2. Cleary-Goldman J, Malone FD, Lambert-Messerlian G, Sullivan L, Canick J, Porter TF, et al. Maternal thyroid hypofunction and pregnancy outcome. *Obstet Gynecol*. 2008;**112**(1):85–92. doi: 10.1097/AOG.0b013e3181788dd7. [PubMed: 18591312]. [PubMed Central: PMC4949950].
3. Lazarus JH. Epidemiology and prevention of thyroid disease in pregnancy. *Thyroid*. 2002;**12**(10):861–5. doi: 10.1089/105072502761016485. [PubMed: 12487768].
4. Yassaee F, Farahani M, Abadi AR. Prevalence of subclinical hypothyroidism in pregnant women in tehran-iran. *Int J Fertil Steril*. 2014;**8**(2):163–6. [PubMed: 25083181]. [PubMed Central: PMC4107690].
5. Carty DM, Doogan F, Welsh P, Dominiczak AF, Delles C. Thyroid stimulating hormone (TSH) >/=2.5mU/l in early pregnancy: Prevalence and subsequent outcomes. *Eur J Obstet Gynecol Reprod Biol*. 2017;**210**:366–9. doi: 10.1016/j.ejogrb.2017.01.048. [PubMed: 28153744].
6. Mannisto T, Mendola P, Grewal J, Xie Y, Chen Z, Laughon SK. Thyroid diseases and adverse pregnancy outcomes in a contemporary US cohort. *J Clin Endocrinol Metab*. 2013;**98**(7):2725–33. doi: 10.1210/jc.2012-4233. [PubMed: 23744409]. [PubMed Central: PMC3701274].

7. Su PY, Huang K, Hao JH, Xu YQ, Yan SQ, Li T, et al. Maternal thyroid function in the first twenty weeks of pregnancy and subsequent fetal and infant development: a prospective population-based cohort study in China. *J Clin Endocrinol Metab*. 2011;**96**(10):3234–41. doi: 10.1210/jc.2011-0274. [PubMed: 21832110].

8. Maraka S, Ospina NM, O'Keeffe DT, Espinosa De Ycaza AE, Gionfriddo MR, Erwin PJ, et al. Subclinical Hypothyroidism in Pregnancy: A Systematic Review and Meta-Analysis. *Thyroid*. 2016;**26**(4):580–90. doi: 10.1089/thy.2015.0418. [PubMed: 26837268]. [PubMed Central: PMC4827301].

9. Kahric-Janicic N, Soldin SJ, Soldin OP, West T, Gu J, Jonklaas J. Tandem mass spectrometry improves the accuracy of free thyroxine measurements during pregnancy. *Thyroid*. 2007;**17**(4):303–11. doi: 10.1089/thy.2006.0303. [PubMed: 17465859]. [PubMed Central: PMC3641793].

10. Stagnaro-Green A, Abalovich M, Alexander E, Azizi F, Mestman J, Negro R, et al. Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and postpartum. *Thyroid*. 2011;**21**(10):1081–125. doi: 10.1089/thy.2011.0087. [PubMed: 21787128]. [PubMed Central: PMC3472679].

11. Horacek J, Spitalnikova S, Dlabalova B, Malirova E, Vizda J, Svilias I, et al. Universal screening detects two-times more thyroid disorders in early pregnancy than targeted high-risk case finding. *Eur J Endocrinol*. 2010;**163**(4):645–50. doi: 10.1530/EJE-10-0516. [PubMed: 20682632].

12. Nazarpour S, Tehrani FR, Simbar M, Tohidi M, AlaviMajd H, Azizi F. Comparison of universal screening with targeted high-risk case finding for diagnosis of thyroid disorders. *Eur J Endocrinol*. 2016;**174**(1):77–83. doi: 10.1530/EJE-15-0750. [PubMed: 26510839].

13. Kalra S, Khandelwal SK, Goyal A. Clinical scoring scales in thyroidology: A compendium. *Indian J Endocrinol Metab*. 2011;**15**(Suppl 2):S89–94. doi: 10.4103/2230-8210.83332. [PubMed: 21966660]. [PubMed Central: PMC3169861].

14. Billewicz WZ, Chapman RS, Crooks J, Day ME, Gossage J, Wayne E, et al. Statistical methods applied to the diagnosis of hypothyroidism. *Q J Med*. 1969;**38**(150):255–66. [PubMed: 4181088].

15. 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**(3):771–6. doi: 10.1210/jcem.82.3.3810. [PubMed: 9062480].

16. Nazarpour S, Ramezani Tehrani F, Simbar M, Tohidi M, Azizi F. Thyroid and Pregnancy in Tehran, Iran: Objectives and Study Protocol. *Int J Endocrinol Metab*. 2016;**14**(1). e33477. doi: 10.5812/ijem.33477. [PubMed: 27279833]. [PubMed Central: PMC4895998].

17. Alexander EK, Pearce EN, Brent GA, Brown RS, Chen H, Dosiou C, et al. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum. *Thyroid*. 2017;**27**(3):315–89. doi: 10.1089/thy.2016.0457. [PubMed: 28056690].

18. Christensen E. Methodology of diagnostic tests in hepatology. *Ann Hepatol*. 2009;**8**(3):177–83. [PubMed: 19841495].

19. Galia AM, DPCP AA, FPCP F KSA, FPCP F, San Luis Jr TO. Validation of the UST thyroid scoring index against ultrasensitive assays for thyroid-stimulating hormone and free thyroxine. *Philp J Int Med*. 2010;**48**(1):15–23.

20. Casey BM, Leveno KJ. Thyroid disease in pregnancy. *Obstet Gynecol*. 2006;**108**(5):1283–92. doi: 10.1097/01.AOG.0000244103.91597.c5. [PubMed: 17077257].

21. Surks MI, Boucai L. Age- and race-based serum thyrotropin reference limits. *J Clin Endocrinol Metab*. 2010;**95**(2):496–502. doi: 10.1210/jc.2009-1845. [PubMed: 19965925].

22. Maraka S, Mwangi R, McCoy RG, Yao X, Sangaralingham LR, Ospina NMS, et al. Thyroid hormone treatment among pregnant women with subclinical hypothyroidism: US national assessment. *BMJ*. 2017;**356**. i6865.

23. Negro R, Stagnaro-Green A. Diagnosis and management of subclinical hypothyroidism in pregnancy. *BMJ*. 2014;**349**:g4929. doi: 10.1136/bmj.g4929. [PubMed: 25288580].

24. Jouyandeh Z, Hasani-Ranjbar S, Qorbani M, Larijani B. *Universal screening versus selective case-based screening for thyroid disorders in pregnancy*. Springer; 2015.

25. O'Reilly DS. Thyroid function tests-time for a reassessment. *BMJ*. 2000;**320**(7245):1332–4. [PubMed: 10807634]. [PubMed Central: PMC1127319].

26. Canaris GJ, Steiner JF, Ridgway EC. Do traditional symptoms of hypothyroidism correlate with biochemical disease? *J Gen Intern Med*. 1997;**12**(9):544–50. [PubMed: 9294788]. [PubMed Central: PMC1497160].

27. Behrooz HG, Tohidi M, Mehrabi Y, Behrooz EG, Tehranidoost M, Azizi F. Subclinical hypothyroidism in pregnancy: intellectual development of offspring. *Thyroid*. 2011;**21**(10):1143–7. doi: 10.1089/thy.2011.0053. [PubMed: 21943136]. [PubMed Central: PMC3179614].

28. Reid SM, Middleton P, Cossich MC, Crowther CA, Bain E. Interventions for clinical and subclinical hypothyroidism pre-pregnancy and during pregnancy. *Cochrane Database Syst Rev*. 2013;(5). CD007752. doi: 10.1002/14651858.CD007752.pub3. [PubMed: 23728666].

29. Nazarpour S, Ramezani Tehrani F, Simbar M, Azizi F. Thyroid dysfunction and pregnancy outcomes. *Iran J Reprod Med*. 2015;**13**(7):387–96. [PubMed: 26494985]. [PubMed Central: PMC4609317].