Clinical Biochemistry 74 (2019) 36–41

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com/locate/clinbiochem

Establishing thresholds and effects of gender, age, and season for thyroglobulin and thyroid peroxidase antibodies by mining real-world big data

Chaochao Ma[1], Dandan Li[1], Yicong Yin[1], Jie Wu, Xiuzhi Guo, Ruili Zhang, Yingying Hu, Yutong Zou, Wenzheng Li, Danchen Wang, Xinqi Cheng*, Ling Qiu*

Department of Clinical Laboratory, Peking Union Medical College Hospital, Peking Union Medical College & Chinese Academy of Medical Science, Beijing 100730, PR China

ARTICLE INFO

Keywords:
Thyroglobulin antibody
Thyroid peroxidase antibody
Threshold
Data mining

ABSTRACT

Background: Thyroglobulin antibody (TG-Ab) and thyroid peroxidase antibody (TPO-Ab) are cornerstone biomarkers for autoimmune thyroid diseases, and establishment of appropriate thresholds is crucial for physicians to appropriately interpret test results. Therefore, we established the thresholds of TG-Ab and TPO-Ab in the Chinese population through analysis of real-world big data, and explored the influence of age, gender, and seasonal factors on their levels.

Methods: The data of 35,869 subjects downloaded from electronic health records were analyzed after filtering based on exclusion criteria and outliers. The influence of each factor on antibody levels was analyzed by stratification. Thresholds of TG-Ab and TPO-Ab were established through Clinical Laboratory Standards Institute document C28-A3 and National Academy of Clinical Biochemistry (NACB) guidelines, respectively.

Results: There were significant differences according to gender after age stratification; the level of TG-Ab gradually increased with age in females. There were significant differences in TG-Ab and TPO-Ab distributions with respect to age after gender stratification. Moreover, differences were observed between seasons for TG-Ab and TPO-Ab. The thresholds of TG-Ab and TPO-Ab were 107 [90% confidence interval (CI):101–115] IU/mL and 29 (90% CI: 28–30) IU/mL, respectively, using C28-A3 guidelines, but were 84 (90% l CI: 50–126) IU/mL and 29 (90% CI: 27–34) IU/mL, respectively, using NACB guidelines.

Conclusion: The levels of TG-Ab and TPO-Ab were significantly affected by gender, age, and season. The thresholds for TG-Ab and TPO-Ab for the Chinese population were established by big data analysis.

1. Introduction

Thyroglobulin antibody (TG-Ab) and thyroid peroxidase antibody (TPO-Ab) play a vital role in the diagnosis of autoimmune thyroid disease. These two antibodies have also been used to predict hypothyroidism. Some studies indicated that TG-Ab production is relatively common (~15–25%) in patients with differentiated thyroid cancer, and should be considered when appraising thyroid globulin results [1,2].

Determining appropriate thresholds of TG-Ab and TPO-Ab is essential for clinical decision-making when physicians are interpreting test results. However, currently, the TG-Ab and TPO-Ab thresholds used in clinical laboratories in China are provided by the manufacturers of test kits, which were established based on foreign populations. Given known influences of population and genetic background, these thresholds may not apply to the Chinese population, resulting in incorrect diagnosis and prognosis of thyroid-related diseases. Furthermore, the C28-A3 file [3] published by the Clinical Laboratory Standard Institute (CLSI), and the reagent specifications for these two antibodies, all recommend that clinical laboratories establish thresholds for their own laboratories. This is mainly because the apparent thresholds for healthy subjects are significantly affected by the analytical system used and the applicable population. Therefore, the specific thresholds of TG-Ab and TPO-Ab for the Chinese population should be established.

Moreover, current reference intervals or thresholds for most clinical laboratory indicators are mainly established based on data from healthy

* Corresponding authors.
E-mail addresses: chengxq@pumch.cn (X. Cheng), lingqiubj@163.com (L. Qiu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.clinbiochem.2019.08.011
Received 5 August 2019; Received in revised form 23 August 2019; Accepted 23 August 2019
Available online 30 August 2019

subjects, which is referred to as the “direct method” in the C28-A3 file [3]. However, the tedious and time-consuming process of the direct method makes it difficult to obtain fully defined data of healthy subjects belonging to specific demographic groups such as the elderly, children, and pregnant women. The high cost of establishing population-specific reference intervals or thresholds by the direct method limits the feasibility of this approach for small clinical laboratories. Especially for TG-Ab and TPO-Ab, the establishment of the thresholds by the direct method requires thyroid ultrasound and other imaging examinations for the subjects. In addition, use of the direct method to establish a reference interval or threshold specific to age, sex, or season would require analysis of a large number of apparently healthy people, necessitating a large amount of labor and financial resources.

The U.S. Food and Drug Administration and China's State Food and Drug Supervision and Administration respectively released the “Use Real-World Evidence to Support Medical Device Registration Laws and Regulations” [4] and “Opinions on Deepening Reform of the Review and Approval System to Encourage Innovation of Drugs and Medical Devices” [5] on August 31, 2017 and October 8, 2017, respectively, both of which encouraged the use of evidence from real-world research to support medical device approval decisions. Clinical laboratory data constitute an important real-world data source, and thus their standardized utilization is an important link in the production of real-world evidence. Accordingly, real-world big data research in clinical laboratories provides the possibility of establishing personalized reference intervals and thresholds.

A previous Chinese multicenter study conducted by our team [6] showed differences in TG-Ab and TPO-Ab levels between females and males by direct method using analytical system of Beckman Coulter. Therefore, in the present study, we used the real world data from the clinical laboratory to analyze the differences between TG-Ab and TPO-Ab in gender, age and season and provide useful information for establishing appropriate thresholds for these two antibodies using analytical system of Roche. To our knowledge, this is the first study to establish thresholds for TG-Ab and TPO-Ab by mining real-world data from clinical laboratories. We further adopted a bootstrap approach with 1000 re-samples to verify the feasibility of using real-world big data to establish thresholds.

2. Materials and methods

2.1. Study population

This real-world study was based on the clinical laboratory big data of 150,431 subjects who underwent health checkups at Peking Union Medical College Hospital between January 1, 2014 and December 29, 2018. We excluded subjects according to the following criteria based on the IFCC/C-RIDL protocol [7]: (1) history of acute or chronic diseases, including respiratory diseases, circulatory diseases, digestive diseases, urinary diseases, autoimmune disease, acute and chronic infections, metabolic and nutritional diseases, blood system diseases, endocrine diseases, and cancers; (2) body mass index (BMI) $\geq 28\ kg/m^2$ or $\leq 18.5\ kg/m^2$; (3) systolic blood pressure ≥ 160 mmHg or diastolic blood pressure ≥ 100 mmHg; (4) goiter; (5) thyroid hormone levels [thyroid stimulating hormone (TSH), free thyroxine (FT4), free triiodothyronine (FT3), total thyroxine (TT4), total triiodothyronine (TT3)] and antibody levels (TPO-Ab and TG-Ab) were not completely determined; and (6) missing values of categorical variables or invalid information in baseline demographic data. Subjects will not be included in the study if the their detection value of TG-Ab or TPO-Ab is outside mean + 3 × the standard deviation. After removing the subjects with outliers in TG-Ab or TPO-Ab levels by the statistical method [8], a total of 35,869 subjects were included in the study.

2.2. Data collection and cleaning

Thyroid antibodies (TG-Ab, TPO-Ab), thyroid hormones (TSH, FT4, FT3, TT4, TT3), and baseline demographic and clinical data [unique identifier, name, age, sex, medical history, BMI, systolic blood pressure, diastolic blood pressure, heart rate, total protein (TP), albumin (Alb), total bilirubin (TBil), direct bilirubin (DBil), alanine aminotransferase (ALT), glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), aspartate aminotransferase (AST), lactate dehydrogenase (LD), urea, uric acid (UA), creatinine (Cr), glucose (Glu), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), potassium, sodium, chloride] were collected from electronic health records. These data were stored in a custom database for later analysis.

All data were filtered and analyzed after identifying information was removed for maintaining anonymity. Finally, unqualified fields such as blank spaces behind records were removed.

2.3. Ethics approval

The Ethics Committee of Peking Union Medical College & Chinese Academy of Medical Sciences, Peking Union Medical College Hospital, approved this study (S-K802).

2.4. Laboratory measurements

Preparation before sampling, blood samples collection, transportation, and processing were conducted according to the Guidelines for the Collection and Transportation of Samples for Testing (PUMCHL-L-2-Q25b-04) program. The samples were centrifuged (3000 rpm for 10 min) to separate the serum before testing. All samples with hemolysis, icterus, and lipemia were considered to be unqualified and excluded from the analysis. The TG-Ab and TPO-Ab measurements were performed using Roche kits and included reagents (Cobas e601, Roche Diagnostics, Mannheim, Germany) according to the manufacturer instructions. The thyroid function tests, including TSH, FT3, FT4, TT4 and TT3 levels, were measured with an ADVIA Centaur XP automatic chemiluminescence immunoassay analyzer (Siemens Healthineers, Erlangen, Germany) using the supplied reagents and calibrators. TP, Alb, TBil, DBil, ALT, GGT, ALP, AST, LD, urea, UA, Cr, Glu, TC, TG, HDL-C, LDL-C, potassium, sodium, and chloride were evaluated on the Cobas 8000 system (Roche Diagnostics, Mannheim, Germany). The threshold used to discriminate between TG-Ab-positive and -negative subjects was 115 IU/mL, according to the manufacturer's guideline. The coefficients of variation of TG-Ab provided by manufacturer were 4.9%, 1.3%, and 1.3% at the level of 47.2 IU/mL, 588 IU/mL, and 3289 IU/mL, respectively. The limit of detection was 10 IU/mL. According to the reagent manufacturer's instruction, 34 IU/mL was used as the threshold to distinguish between TPO-positive and -negative individuals. The coefficients of variation of TPO-Ab provided by manufacturer were 6.3%, 5.1%, and 2.7% at the level of 21.3 IU/mL, 51.2 IU/mL, and 473 IU/mL, respectively; the limit of detection was 5 IU/mL. Characteristics of the assays for other basic measurements are provided in Supplementary Table 1.

2.5. Quality assurance

Quality controls before the analyses and appropriate calibration were performed to ensure the accuracy of test results and monitor the precision of the instruments. The instruments were calibrated and maintained annually. Furthermore, our laboratory participated in external quality assessments conducted by the National Center for Clinical Laboratories every year and by the College of American Pathologists once every two years to ensure that our test results are accurate and reliable [9]. TG-Ab and TPO-Ab were assessed by the College of American Pathologists three times a year.

2.6. Statistical analysis

Excel 2016 (Microsoft, Redmond, WA, USA) was used to store and clean the data. Data were statistically analyzed using SPSS 25.0 software (IBM Inc., Armonk, NY, USA) and Medcalc Statistical software 18.116.6 (Mariakerke, Belgium). Levels of TG-Ab or TPO-Ab that were lower than the limit of detection were considered to be equal to the limit of detection value for statistical analysis. The data distributions were evaluated by Kolmogorov-Smirnov tests and visualization of the quantile-quantile plot. Data with a Gaussian distribution are presented as mean ± standard deviation, while non-Gaussian distributed data are presented as median (1st quartile, 3rd quartile). The effects of age and gender on TG-Ab and TPO-Ab levels were analyzed by stratification; the Mann-Whitney *U* test and Kruskal–Wallis test for independent samples were used to compare differences between groups. Age - and sex-specific thresholds were then established. These thresholds were determined as the 95th percentiles of the distribution for reference subjects (Method 1 in the study). Moreover, criteria of the National Academy of Clinical Biochemistry (NACB) guidelines [10] were also referenced to select the reference population among the total 150,431 subjects, and corresponding thresholds for TG-Ab and TPO-Ab were also established (Method 2 in the study). These thresholds were used as the 97.5th percentiles of the distribution for the reference subjects. A bootstrap procedure through random resampling of the same database for 1000 replicates and 20% of the data extracted at a time was performed to calculate the 90% confidence intervals (90% CIs) for the threshold values. *P* values < .05 indicated that a difference was statistically significant.

3. Results

3.1. Baseline characteristics of the subjects

A total of 35,869 subjects (15,864 females and 20,005 males) with a median age and BMI of 38 (31, 47) years and 23.34 (21.34, 25.26) kg/m^2 were ultimately included in the analysis (Table 1). The median TG-Ab and TPO-Ab levels were 11.98 (10.0, 16.04) and 11.49 (8.63, 14.90), respectively. The study subjects were divided into six groups according to age: ≤30, 31–40, 41–50, 50–60, 60–70, and ≥71 years.

3.2. Effect of gender and age on TG-Ab and TPO-Ab levels

The levels of TG-Ab and TPO-Ab in the subjects stratified by gender and age are presented in Table 2. There were significant differences according to gender ($p < .001$) after age stratification. The level of TG-Ab gradually increased with increasing age in females, while TG-Ab levels reached their peak in males of the age group 50–59 years (Fig. 1). Similarly, significant differences in the distribution of TG-Ab ($p < .001$) and TPO-Ab ($p < .001$) with regard to age were observed after gender stratification. Pairwise comparisons of the age groups after gender stratification are shown in Supplementary Table 2. The thresholds (95% percentile) for TG-Ab and TPO-Ab in females were higher than the males ($p < .001$) after age stratification. Furthermore, the threshold (95% percentile) for these two antibodies increased with age both in females and males.

3.3. Influence of season on TG-Ab and TPO-Ab levels

Differences in TG-Ab and TPO-Ab levels were also observed between the different seasons ($p < .05$). The median serum TG-Ab levels were higher in summer and autumn, and were lower in spring and winter in females and males (Supplementary Table 3). Although the TG-Ab levels did not differ between spring and winter ($p = .967$) or between summer and autumn ($p = .729$) for females, the difference of TG-Ab levels between each season was statistically significant for males ($p < .0001$). TG-Ab levels were the highest in summer and the lowest in winter for both males and females (Fig. 2). The median TPO-Ab levels were similar across seasons for females, with the highest level in summer and the lowest level in winter, whereas the TPO-Ab levels in males were the highest in summer but the lowest in autumn; the TPO-Ab levels did not differ between spring and winter ($p = .516$), spring and autumn ($p = 1.000$), or autumn and winter ($p = 1.000$) in females and did not differ between spring and winter in males ($p = .967$).

3.4. Establishment and validation of thresholds for TG-Ab and TPO-Ab for the Chinese population

Finally, thresholds of TG-Ab and TPO-Ab were established based on different reference populations (Table 3). The inclusion and exclusion criteria and data filtering methods were used to select the reference population for comparison, whose 95th quantile of TG-Ab and TPO-Ab was 107 (90% CI:101–115) IU/mL and 29 (90% CI: 28–30) IU/mL, respectively. The other reference population was selected according to the guidelines of the NACB, and the 97.5th quantile of TG-Ab and TPO-Ab was 84 (90% CI: 50–126) IU/mL and 29 (90% CI: 27–34) IU/mL, respectively.

4. Discussion

To our knowledge, this is the first “big data” study to analyze the factors affecting TG-Ab and TPO-Ab levels. Here, we focused on three factors that could potentially affect antibody levels: gender, age, and season. Furthermore, we established the thresholds of these two antibodies based on big data obtained from a clinical laboratory, and adopted a bootstrap approach with 1000 resamples to verify the feasibility of using real-world big data to establish thresholds.

Quality control measures, including external quality assessment, and changes in methodology, instruments, and reagents during the five years of data collection were reviewed prior to data cleaning to ensure high-quality data for analysis. Moreover, data weight and observation number were used as evaluation indices to ensure consistency before and after data cleaning. Meanwhile, an equidistant sampling method was adopted to check the data before and after data cleaning to ensure the quality of the data cleaning process.

A previous study [11] including 120 males established the reference intervals for TG-Ab and TPO-Ab, revealing a difference in the prevalence of TG-Ab-positive and TPO-Ab-positive individuals between females and males. Another Chinese multi-center study conducted by our team [6] also showed a clear difference in the frequencies of positivity of these antibodies between females and males, based on the immunoassay method using an automated analyzer (Beckman Coulter DXI 800, Beckman Coulter; Brea, CA, USA). Our present results confirmed these findings, showing significantly higher TG-Ab and TPO-Ab levels as well as wider distributions of TG-Ab in females. Furthermore, this is the first study to demonstrate a change in TG-Ab and TPO-Ab levels with age; after gender stratification, significant differences between some age groups were found based on the Kruskal–Wallis test. This finding suggests the relevance of establishing gender- and age-specific thresholds of TG-Ab and TPO-Ab.

Previous various big data studies [8,12–14] found that the TSH level rose in the winter and declined in the summer, which was confirmed in the present study after stratification by gender. However, the seasonal fluctuations of the TG-Ab level were opposite to the direction of TSH fluctuations reported previously [15]. However, TG-Ab- and TPO-Ab positive patients were excluded from previous studies [11,16,17] that established TSH reference intervals. This suggests that the presence of TG-Ab and TPO-Ab may affect TSH levels. Therefore, the correlation between TSH and TG-Ab levels warrants further study.

To the best of our knowledge, few previous studies have paid attention to the specific thresholds for TG-Ab and TPO-Ab; thus, further effort is needed to develop such clinically valuable reference data. If the 95% confidence interval of the newly established threshold contains the

Table 1
Characteristics of the study population.

Characteristic	Unit	Total	Females	Males
n		35,869	15,864	20,005
Age		38 (31, 47)	38 (30, 47)	39 (32, 47)
BMI	kg/m^2	23.34 (21.34, 25.26)	21.91 (20.32, 23.83)	24.30 (22.6, 25.86)
Systolic pressure	mmHg	114 (105, 125)	108 (100, 117)	119 (110, 130)
Diastolic pressure	mmHg	73 (66, 80)	72 (66, 79)	74 (67, 80)
Heart rate	bmp	68 (62, 74)	68 (63, 74)	68 (62, 74)
TSH	μIU/mL	1.916(1.374, 2.669)	2.067 (1.481, 2.909)	1.809 (1.302, 2.484)
FT3	pg/mL	3.20 (2.97, 3.44)	3.01 (2.83, 3.2)	3.35 (3.16, 3.56)
FT4	ng/dL	1.23 (1.13, 1.33)	1.17 (1.08, 1.26)	1.27 (1.17, 1.38)
TT3	ng/mL	1.06 (0.96, 1.18)	1.04 (0.94, 1.14)	1.09 (0.98, 1.20)
TT4	μg/dL	7.74 (6.82, 8.68)	7.68 (6.81, 8.57)	7.80 (6.83, 8.75)
TP	g/L	73 (70, 75)	73 (70, 75)	73 (71, 75)
Alb	g/L	47 (45, 48)	46 (44, 47)	47 (46, 49)
TBil	μmol/L	10.8 (8.1, 14.4)	9.3 (7.2, 12.2)	12.2 (9.3, 15.8)
DBil	μmol/L	4.1 (3.2, 5.1)	3.7 (2.9, 4.6)	4.4 (3.5, 5.5)
ALT	U/L	17 (12, 25)	13 (10, 18)	21 (16, 30)
GGT	U/L	18 (13, 29)	13 (10, 18)	24 (17, 38)
ALP	U/L	61 (51, 73)	55 (46, 66)	65 (56, 76)
AST	U/L	18 (15, 22)	17 (14, 20)	19 (17, 23)
LD	U/L	166 (150, 183)	162 (147, 180)	169 (153, 186)
Urea	mmol/L	4.48 (3.78, 5.30)	4.08 (3.46, 4.85)	4.79 (4.10, 5.57)
UA	μmol/L	311 (255,373)	254 (221, 292)	360 (316, 410)
Cr	μmol/L	73 (61,84)	60 (55, 66)	82 (76, 89)
Glu	mmol/L	5.0 (4.7, 5.4)	4.9 (4.7, 5.2)	5.1 (4.8, 5.5)
TC	mmol/L	4.61 (4.08, 5.21)	4.52 (4.01, 5.13)	4.67 (4.13, 5.27)
TG	mmol/L	1.06 (0.76, 1.56)	0.86 (0.66, 1.21)	1.26 (0.90, 1.83)
HDL-C	mmol/L	1.29 (1.09, 1.53)	1.46 (1.25, 1.69)	1.17 (1.01, 1.36)
LDL-C	mmol/L	2.87 (2.40, 3.40)	2.70 (2.26, 3.23)	3.0 (2.53, 3.52)
Potassium	mmol/L	4.2 (4.0, 4.4)	4.2 (4.0, 4.4)	4.2 (4.0, 4.4)
Sodium	mmol/L	141 (140, 142)	140 (139, 142)	141 (140, 142)
Chloride	mmol/L	102 (100, 104)	102 (101, 104)	102 (100, 103)

The data are presented as median with quartiles (25th, 75th).
Abbreviations: BMI, body mass index; TP, total protein; Alb, albumin; TBil, total bilirubin; DBil, direct bilirubin; ALT, alanine aminotransferase; GGT, glutamyl transpeptidase; ALP, alkaline phosphatase; AST, aspartate aminotransferase; LD, lactate dehydrogenase; UA, uric acid; Cr, creatinine; Glu, glucose; TC, total cholesterol; TG, triglycerides; HDL-C, high density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

manufacturer's threshold, it is considered that our method of establishing the threshold through big data is feasible. Here, we used two methods based on CLSI and NACB guidelines, respectively, to establish the thresholds for these two antibodies. Under bootstrap validation with 1000 resamples, the 90% CI values corresponded to the thresholds provided by the manufacturers, except for method 1 used to establish the threshold for TPO-Ab in this study, which was 29 IU/L and still very similar to that provided by the manufacturers. In addition, there were differences in the results of establishing antibody thresholds between the two data screening methods for TG-Ab. The 90 confidence interval of method 1 is narrower than that of method 2, and the accuracy is better. These findings thus demonstrate the feasibility of using big data to establish a one-side threshold of the test index to some extent. The different data analysis and mining methods should be selected for different index.

Our study has certain advantages compared to previous studies that established thresholds or reference intervals for TG-Ab and TPO-Ab. First, a large cohort of big clinical laboratory data was used in the present study. Second, we used a single assay for the determination of the two antibodies to prevent inter-analytical variation. Third, although this was a big data study, the clinical baseline information of each patient was very complete. Moreover, we filtered the data according to strict inclusion and exclusion criteria and used statistically based outlier removal methods. Fourth, we made great efforts to ensure the quality of data by reviewing the quality control data, and checked the data weight during filtering. Finally, but most importantly, we adopted a bootstrap

Table 2
Tg-Ab and TPO-Ab levels by age and gender.

Index	Females						Males					
	Age group (years)	N	Median	25th percentile	75th percentile	95 percentile	N	Median	25th percentile	75th percentile	95 percentile	p
TG-Ab (IU/mL)	≤29	3363	11.74	10	15.69	141.78	3298	11.14	10	14.43	33.94	< 0.001
	30–39	5508	12.23	10	16.98	198.26	7170	11.44	10	14.8	33.10	< 0.001
	40–49	3966	12.74	10	18.83	263.79	5455	11.94	10	15.58	40.64	< 0.001
	50–59	2191	13.29	10	21.57	272.12	2770	12.33	10	16.17	53.68	< 0.001
	≥60	836	13.3	10	21.72	302.69	1312	12.11	10	15.82	54.40	< 0.001
TPO-Ab (IU/mL)	≤29	3363	12.09	9.29	15.32	27.81	3298	10.76	8.05	13.77	20.87	< 0.001
	30–39	5508	11.85	9.08	15.42	34.20	7170	11.17	8.4	14.39	22.93	< 0.001
	40–49	3966	11.88	8.94	15.41	42.98	5455	11.12	8.23	14.51	24.34	< 0.001
	50–59	2191	12.29	9.32	16.49	54.42	2770	11.3	8.3	14.77	30.43	< 0.001
	≥60	836	12.43	9.38	16.04	60.16	1312	11.16	8.11	15.26	31.04	< 0.001

P: Males and Females in the same age group were compared.

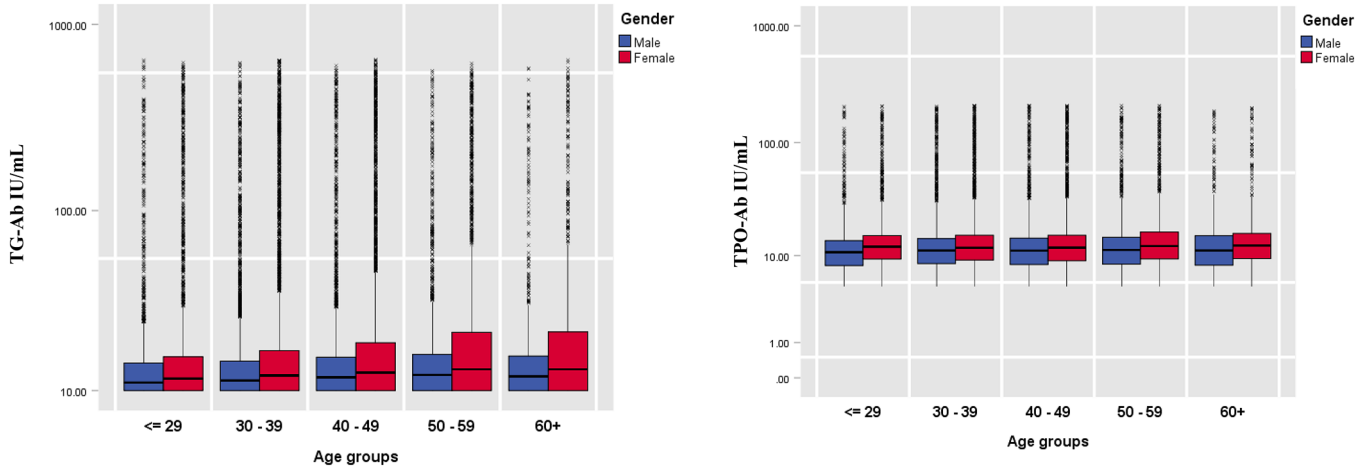

Fig. 1. Distribution of TG-Ab and TPO-Ab by age and gender.

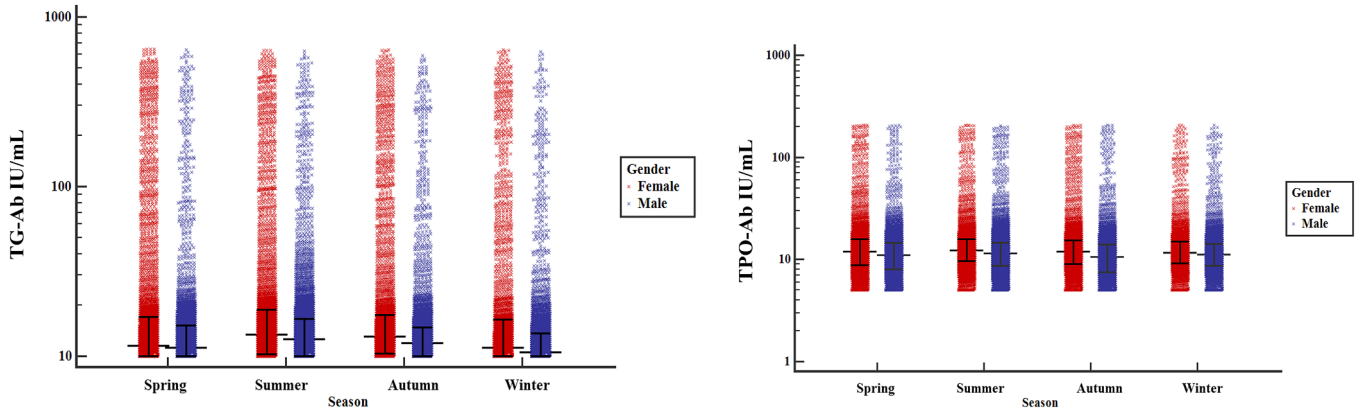

Fig. 2. Distribution of TG-Ab and TPO-Ab by season.

Table 3
Thresholds of Tg-Ab and TPO-Ab determined by different data filtering methods.

	Tg-Ab (IU/mL)		Manufacturer	TPO-Ab (IU/mL)		Manufacturer
	Method 1	Method 2		Method 1	Method 2	
N	35,869	2581	392	35,869	2581	208
Threshold	95th percentile: 107	97.5th percentile: 84	94th percentile: 115	95th percentile: 29	97.5th percentile: 29	95th percentile: 34
90% CI	101–115	50–126	/	28–30	27–34	/

N: number of subjects; CI: confidence interval; Method 1: The inclusion and exclusion criteria and data cleaning methods referenced the CLSI criteria/cleaning methods; Method 2: Reference criteria according to the guidelines of the National Academy of Clinical Biochemistry (NACB): 1. males; 2. age < 30 years old; 3. no goiter; 4. Serum TSH 0.5–2 mIU/mL.

method with 1000 resamples to verify the feasibility of using real-world big data to establish thresholds.

These results provide an analytical framework for clinical laboratories to establish reference intervals or thresholds based on real-world big data. However, our study has several limitations. First, the method used in the validation could have influenced the results; thus, use of an independent, relatively homogeneous validation cohort could better verify the feasibility of big data to establish thresholds. Second, use of a more specific age grouping method could provide more accurate results, including classification algorithms such as a decision tree or discriminant analysis [18]. Nevertheless, the results obtained by the age grouping method adopted in our study are more relevant for clinical practice.

5. Conclusion

We detected a significant difference in TG-Ab and TPO-Ab levels between females and males and according to season, with higher TG-Ab levels in the summer and lower levels in the winter. And age- and gender-specific thresholds were established. The thresholds of all health subjects for TG-Ab and TPO-Ab were established and verified by mining big data. These findings demonstrate the feasibility of establishing one-side thresholds for antibodies using real-world big data from clinical laboratories.

Author contributions

All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding

None declared.

Employment or leadership

None declared.

Honorarium

None declared.

Declaration of Competing Interest

The funding organization(s) played no role in the study design; in collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Acknowledgements

The authors are grateful to Deepwise (China) Co., Ltd., for providing technical support. The authors also thank Libiao Peng and Wei Liu from Deepwise Inc. for their kind cooperation, which helped implement this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinbiochem.2019.08.011.

References

[1] S. Rinaldi, M. Plummer, C. Biessy, K.K. Tsilidis, J.N. Østergard, K. Overvad, et al., Thyroid-stimulating hormone, thyroglobulin, and thyroid hormones and risk of differentiated thyroid carcinoma: the EPIC study, J. Natl. Cancer Inst. 106 (2014) dju097.

[2] C.A. Spencer, M. Takeuchi, M. Kazarosyan, C.C. Wang, R.B. Guttler, P.A. Singer, et al., Serum thyroglobulin autoantibodies: prevalence, influence on serum thyroglobulin measurement, and prognostic significance in patients with differentiated thyroid carcinoma, J. Clin. Endocrinol. Metab. 83 (1998) 1121–1127.

[3] Clinical Laboratory Standard Institute, Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approval Guideline-Third Edition, CLSI, Geneva, 2010.

[4] Food and Drug Administration, Use of Real-World Evidence to Support Regulatory Decisions-Making for Medical Devices: Guidance for Industry and Food and Drug Administration Staff, Available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices.

[5] China Food and Drug Administration, Opinions on Deepening the Reform of the Review and Approval System to Encourage Innovation of Drugs and Medical Devices, Available at: http://www.gov.cn/zhengce/2017-10/08/content_5230105.htm.

[6] D. Wang, S. Yu, X. Cheng, L. Cao, H. Zhang, L. Liu, et al., Nationwide Chinese study for establishing reference intervals for thyroid hormones and related tests, Clin. Chim. Acta (2019), https://doi.org/10.1016/j.cca.2019.06.011 SOOO9-898:31903–5.

[7] Y. Ozarda, K. Ichihara, J.H. Barth, G. Klee, Committee on Reference Intervals and Decision Limits (C-RIDL), International Federation for Clinical Chemistry and Laboratory Medicine, Protocol and standard operating procedures for common use in a worldwide multicenter study on reference values, Clin. Chem. Lab. Med. 51 (2013) 1027–1040.

[8] D. Wang, S. Yu, C. Ma, H. Li, L. Qiu, X. Cheng, et al., Reference intervals for thyroid-stimulating hormone, free thyroxine, and free triiodothyronine in elderly Chinese persons, Clin. Chem. Lab. Med. 57 (2019) 1044–1052.

[9] D. Li, D. Wang, D. Wang, C. Ma, J. Wu, P. Li, et al., Data mining: biological and temporal factors associated with blood cardiac troponin I concentration in a Chinese population, Clin. Chim. Acta 495 (2019) 8–12.

[10] Z. Baloch, P. Carayon, B. Conte-Devolx, L.M. Demers, U. Feldt-Rasmussen, J.F. Henry, et al., Laboratory medicine practice guidelines. Laboratory support for the diagnosis and monitoring of thyroid disease, Thyroid 13 (2003) 3–126.

[11] C. Sriphrapradang, S. Pavarangkoon, W. Jongjaroenprasert, L.O. Chailurkit, B. Ongphiphadhanakul, W. Aekplakorn, Reference ranges of serum TSH, FT4 and thyroid autoantibodies in the Thai population: the national health examination survey, Clin. Endocrinol. 80 (2014) 751–756.

[12] J. Ehrenkranz, P.R. Bach, G.L. Snow, A. Schneider, J.L. Lee, S. Ilstrup, et al., Circadian and circannual rhythms in thyroid hormones: determining the TSH and free T4 reference intervals based upon time of day, age, and sex, Thyroid 25 (2015) 954–961.

[13] L.A. De Grande, K. Goossens, K. Van Uytfanghe, I. Halsall, J. Yoshimura Noh, K. Hens, et al., Using “big data” to describe the effect of seasonal variation in thyroid-stimulating hormone, Clin. Chem. Lab. Med. 55 (2017) e34–e36.

[14] D. Wang, X. Cheng, S. Yu, L. Qiu, X. Lian, X. Guo, et al., Data mining: seasonal and temperature fluctuations in thyroid-stimulating hormone, Clin. Biochem. 60 (2018) 59–63.

[15] D. Wang, D. Li, X. Guo, S. You, L. Qiu, X. Cheng, et al., Effects of sex, age, sampling time, and season on thyroid-stimulating hormone concentrations: a retrospective study, Biochem. Biophys. Res. Commun. 506 (2018) 450–454.

[16] A. Clerico, T. Trenti, R. Aloe, R. Dittadi, S. Rizzardi, M. Migliardi, et al., A multicenter study for the evaluation of the reference interval for TSH in Italy (ELAS TSH Italian study), Clin. Chem. Lab. Med. 57 (2018) 259–267.

[17] R. Tozzoli, F. D’Aurizio, P. Metus, A. Steffan, C. Mazzon, M. Bagnasco, Reference intervals for thyrotropin in an area of Northern Italy: the Pordenone thyroid study (TRIPP), J. Endocrinol. Investig. 41 (2018) 985–994.

[18] X. Peng, Y. Lv, G. Feng, Y. Peng, Q. Li, W. Song, et al., Algorithm on age partitioning for estimation of reference intervals using clinical laboratory database exemplified with plasma creatinine, Clin. Chem. Lab. Med. 56 (2018) 1514–1523.