RESEARCH ARTICLE

Status of internal quality control for thyroid hormones immunoassays from 2011 to 2016 in China

Shishi Zhang^{1,2} | Wei Wang¹ | Haijian Zhao¹ | Falin He¹ | Kun Zhong¹ | Shuai Yuan¹ | Zhiguo Wang¹

¹National Center for Clinical Laboratories/ Beijing Engineering Research Center of Laboratory Medicine, Beijing Hospital, National Center of Gerontology, Beijing, China

²Graduate School of Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

Correspondence

Zhiguo Wang, National Center for Clinical Laboratories/Beijing Engineering Research Center of Laboratory Medicine, Beijing Hospital, National Center of Gerontology, Dongdan, Beijing, China.
Email: zgwang@nccl.org.cn

Funding information

1. Beijing Natural Science Foundation in 2014 (Grant number: No. 7143182). 2. Beijing Hospital Foundation in 2015 (Grant number: BJ-2015-025).

Background: Internal quality control (IQC) plays a key role in the evaluation of precision performance in clinical laboratories. This report aims to present precision status of thyroid hormones immunoassays from 2011 to 2016 in China.

Methods: Through Clinet-EQA reporting system, IQC information of Triiodothyronine and Thyroxine in the form of free and total (FT3, TT3, FT4, TT4), as well as Thyroid Stimulating Hormone (TSH) were collected from participant laboratories submitting IQC data in February, 2011-2016. For each analyte, current CVs were compared among different years and measurement systems. Percentages of laboratories meeting five allowable imprecision specifications (pass rates) were also calculated. Analysis of IQC practice was conducted to constitute a complete report.

Results: Current CVs were decreasing significantly but pass rates increasing only for FT3 during 6 years. FT3, TT3, FT4, and TT4 had the highest pass rates comparing with 1/3TEa imprecision specification but TSH had this comparing with minimum imprecision specification derived from biological variation. Constituent ratios of four mainstream measurement systems changed insignificantly. In 2016, precision performance of Abbott and Roche systems were better than Beckman and Siemens systems for all analytes except FT3 had Siemens also better than Beckman. Analysis of IQC practice demonstrated wide variation and great progress in aspects of IQC rules and control frequency.

Conclusion: With change of IQC practice, only FT3 had precision performance improved in 6 years. However, precision status of five analytes in China was still unsatisfying. Ongoing investigation and improvement of IQC have yet to be achieved.

KEYWORDS

internal quality control, performance specification, precision, thyroid hormones immunoassays

1 | INTRODUCTION

Thyroid dysfunction, both clinically significant and subclinical, affects up to 5%-8% of the global population.¹ A recent epidemiological survey of thyroid diseases in 10 Chinese cities showed that the total prevalence was 1.1% for clinical hyperthyroidism, 2.6% for subclinical hyperthyroidism, 0.9% for hypothyroidism, 5.6% for subclinical hypothyroidism, 2.4% for goiter, 11.6% for thyroid solitary nodule and 7.0% for multiple thyroid nodules, as well as 11.6% and 12.6% positive rates

for TPOAb and TgAb, respectively.² Moreover, the incidence of thyroid cancer (TC) in China was increased 2.36 times from 1988 to 2009, with an average annual increase of 5.92%.³ Compared with the data displayed in the American Thyroid Association's (ATA's) guidelines^{4,5} which had prevalence of hyperthyroidism approximately 1.2% (0.5% overt and 0.7% subclinical) and yearly incidence of TC nearly tripled from 4.9 per 100 000 in 1975 to 14.3 per 100 000 in 2009 among American population, there was no doubt that the situation of thyroid diseases in the whole world is worrying.

Analyte	Allowable imprecision specification based on biological variation and CLIA'88				
	Minimum	Desirable	Optimum	1/3 TEa	1/4 TEa
FT3	5.925	3.95	1.975	8.33	6.25
TT3	5.175	3.45	1.725	8.33	6.25
FT4	4.275	2.85	1.425	8.33	6.25
TT4	3.675	2.45	1.225	6.67	5.00
TSH	14.475	9.65	4.825	8.33	6.25

TABLE 1 Allowable imprecision specifications (%) for each analyte

In the subsequent section of this article, five allowable imprecision specifications were respectively abbreviated as minimum specification, desirable specification, optimum specification, 1/3TEa specification, and 1/4TEa specification.

As one of the most common endocrine disorders, thyroid diseases can be classified into different categories, such as hypothyroidism, hyperthyroidism, thyroid nodule, etc., through the concentration detection of main thyroid hormones triiodothyronine (T3) and thyroxine (T4) in the form of either total or free, as well as TSH in the serum. So, analytical performance of thyroid hormones immunoassays is worthy of more attention. Total error (TE), introduced by Westgard,⁶ was composed of systematic error and random error. The TE model can be expressed as equation: $TE = \text{average bias} + Z \times \text{imprecision}$ (Z defines the level of confidence, usually 1.65).⁷ TE, bias, and imprecision (expressed as CV or SD) were three important indicators used to reflect analytical accuracy, trueness, and precision. Using immunoassays with better precision (smaller CV) is one of important aspects to decrease TE and improve accuracy of diagnosis and treatment.

Internal quality control (IQC), as an effective tool to discover detection errors in the analytical phase, plays a key role in improving precision level of clinical laboratories.^{8,9} Through collecting 2011–2016 IQC results in the national External Quality Assessment (EQA) programs and assessing percentages of laboratories meeting five different performance specifications of precision, this study was to make a conclusion for FT3, TT3, FT4, TT4, and TSH whether the IQC practice had changed and whether the precision performance of Chinese clinical laboratories had made great progress, also help laboratories select the most appropriate allowable imprecision specification and raise precision level.

2 | MATERIALS AND METHODS

2.1 | Subjects

Clinical laboratories came from all over the China which continuously participated in the national EQA programs of FT3, TT3, FT4, TT4, and TSH organized by National Center for Clinical Laboratories of China (CNCCL), the official Proficiency Testing provider in China, and submitted their IQC information from 2011 to 2016.

2.2 | Study design

In the annual EQA cycle of each analyte in China, when participant laboratories submit results of EQA samples via a specific software module

of Clinet-EQA reporting system version 1.5 developed by CNCCL (see <http://www.clinet.com.cn/shop/shop>), they are also asked for reporting their IQC information by another module of this system. If laboratories do not provide their IQC data, CNCCL will not prepare EQA reports to them. However, significantly, IQC control materials (control materials in this text all refer to IQC control material) are no longer provided by CNCCL. Instead, they are purchased from different vendors or prepared in-house by laboratories themselves. Through continuously measuring their own IQC control materials and monitoring current and long-time cumulative coefficient of variations (CVs), laboratories can evaluate the within-laboratory imprecision. Chinese current IQC survey is only to investigate the within-laboratory imprecision, different from studies in western countries devoting to estimate the within-method and between-laboratories imprecision.^{10,11}

The IQC information of FT3, TT3, FT4, TT4, and TSH in clinical laboratories was collected annually in each February from 2011 to 2016. All participant laboratories were asked for returning the following information, including vendors or producers of control materials, lot number of control materials, control rules, control frequency, mean value and standard deviation of control materials, current CVs, cumulative CVs, principle of assay, as well as the manufacturer of instrument, reagent and calibrator. Actually, the calculation of current CVs was performed according to all in-control results judged by laboratories' own IQC control rules in February of each year. But to calculate the cumulative CVs, laboratories need to collect all in-control results from the first day on which the same lot of control material was used to the last day until February.⁹ Due to the different collecting time among all laboratories, this study only analyze current CVs to ensure the comparability of data.

The percentages of laboratories meeting different performance specifications of precision for all analytes were calculated according to five imprecision evaluation criteria as follows (Table 1). When current CVs were less than an evaluation criterion, the precision of this laboratory was considered to be acceptable in the performance request of this evaluation criterion. In each group divided by year and manufacturers of measurement systems (referring to closed systems whose reagent and calibrator matched with the instrument deriving from the same manufacture), etc., the pass rates were defined as the ratio of "the number of laboratories with acceptable precision performance" to "the total number of laboratories in each group".

2.3 | Analytical performance specifications for precision

To evaluate precision performance of clinical laboratories, the current CVs should be compared with five objective allowable imprecision specifications (Table 1). The former three respectively were minimum, desirable, and optimum levels of allowable imprecision specifications based on biological variation, which should be calculated according to the formula defined as $0.75CV_i$, $0.50CV_i$, and $0.25CV_i$. CV_i , the within-subject biological variation, was provided by Ricos and updated by Westgard on the website (<http://www.westgard.com/biodatabase1.htm>). The remaining two, 1/3 TEa and 1/4 TEa, were allowable imprecision specifications recommended by CLIA'88.

2.4 | Statistical analysis

The distributions of current CVs were presented as several percentiles including 5th, 25th, 50th (median), 75th, and 95th percentiles. Statistical analyses were conducted through Microsoft Excel 2007 software (Microsoft Inc, Redmond, Washington, DC, USA) and IBM SPSS Statistics Version 19.0 (SPSS Inc, Chicago, IL, USA). For each analyte, comparison of the current CVs between or among different years, as well as different groups categorized by measurement systems in the same year, was performed by Kruskal-Wallis test or Mann-Whitney test when data did not present normal distribution or homogeneous variance. Otherwise, Independent-sample *t* test should be applied. Chi-square (χ^2) test was used to compare pass rates and constituent ratios between or among different years or groups. The differences, between or among comparators, were considered to be statistically significant when the *P*-value is ≤ 0.05 .

3 | RESULTS

In February 2011, 530, 428, 513, 417, and 516 clinical laboratories from the whole China respectively submitted their IQC data for FT3, TT3, FT4, TT4, and TSH assays. Among them, there were only 187 (35.3%), 155 (36.2%), 184 (35.9%), 152 (36.5%), and 188 (36.4%) clinical laboratories continuously participated in IQC survey and submitted their IQC data in the following 5 years.

Unfortunately, laboratories only measuring single concentration level of control materials accounted for 52.14%, 50.43%, 51.09%, 50.55%, and 51.31% on average in 6 years for FT3, TT3, FT4, TT4, and TSH, respectively (Figure 1). Besides, laboratories applying three levels of control materials for all five analytes were around 10.00%. For each analyte, the remaining participant laboratories, less or just a little more than 40.00%, used two levels of control materials.

3.1 | Imprecision analysis

For five analytes, the distribution of current CVs of level 1 was shown in Figure 2. For FT3, the current CVs among 6 years were

different significantly ($P=0.027$). Further comparisons of the current CVs between any 2 years were conducted by Mann-Whitney test, demonstrating there were significant differences between following 2 years: 2011 and 2015 ($P=0.007$), 2011 and 2016 ($P=0.011$), 2012 and 2015 ($P=0.018$), 2012 and 2016 ($P=0.027$). The current CVs for FT3 in 2015 and 2016 were significantly lower than those in 2011 and 2012. However, the current CVs of other four analytes among 6 years all had no significant difference ($P=0.466$ for TT3, $P=0.508$ for FT4, $P=0.253$ for TT4, and $P=0.335$ for TSH).

Percentages of laboratories meeting five different allowable imprecision specifications (also called 'pass rates') according to current CVs of level 1 in 6 years for five analytes were depicted in Figure 3. Only FT3 had increasing pass rates in general. FT3, TT3, FT4, and TT4 analytes got highest pass rates (79.7%-90.4% for FT3, 76.8%-83.2% for TT3, 85.3%-91.8% for FT4, and 59.9%-73.0% for TT4) when the most loosest 1/3TEa specification was used as evaluation criterion, but got pass rates below 80.0% comparing with other four tighter specifications. For TSH, highest pass rates present when the minimum specification was applied. 1/4TEa and optimum specifications were too strict to TSH with pass rates lower than 80.0%. In comparison with optimum specification, TSH possessed pass rates significantly higher than other analytes (3.2%-5.9% for FT3, 0%-3.9% for TT3, 0.5%-3.8% for FT4, 0.7%-3.3% for TT4, and 50.5%-58.5% for TSH), suggesting the precision performance of TSH was better than other analytes.

3.2 | Imprecision analysis by measurement systems

Immunoassays of five thyroid hormones were conducted by four mainstream measurement systems Abbott, Beckman, Roche and Siemens in China. The constituent ratios of these four mainstream measurement systems had not significantly changed with time (all $P>0.05$). For five analytes, the constituent ratios of Roche in 6 years were the highest (about 40%), followed by Siemens (about 25%), Beckman (about 16%-18%), and Abbott (about 10%) in sequence. The proportions of other measurement systems were the lowest (about 7%-8%) and also changed without significant difference ($P>0.05$).

The current CVs of four mainstream measurement systems were further analyzed and shown in Table S1 (in supplementary materials). In general, except for FT3, the current CVs of Roche both were below Beckman and Siemens in 6 years. For Abbott, the current CVs of five analytes decreased in 6 years. In 2016, Beckman and Siemens both had the current CVs significantly higher than Abbott and Roche in TT3, FT4, TT4, and TSH (all $P<0.05$), except that FT3 had the current CVs of Abbott, Roche, and Siemens lower than Beckman with significant difference (all $P<0.05$).

23.5% (44/187), 22.6% (35/155), 22.8% (42/184), 23.7% (36/152), and 24.5% (46/188) laboratories changed their measurement systems from 2011 to 2016 for FT3, TT3, FT4, TT4, and TSH, respectively. The current CVs between laboratories changing and not changing measurement systems revealed no significant difference in both 2011 and 2016 for five analytes (all $P>0.05$), which indicated laboratories did not improve their precision performance due to the change of measurement systems.

FIGURE 1 Percentages of laboratories using different concentration levels of control materials for five items from 2011 to 2016

3.3 | The changes of IQC practice from 2011 to 2016

In the answers of the question what control rules do you choose in the IQC for five analytes, we found that in 2011 more than 40.0%

laboratories categorized in Group 6 (Figure 4) only provided ambiguous answers such as 'L-J', 'Levey-Jennings', 'Westgard rules', or 'Westgard multi-rules', but in 2016 this constituent ratio was reduced to less than 7.0%. Apparently, more and more laboratories grasped the concepts of Westgard rules applied in IQC. In the part of laboratories

FIGURE 2 Current CVs (%) of level 1 of five analytes from 2011 to 2016. (CV% referred to the current CVs. For each analyte, the current CVs of different year were shown in separate boxes. The upper and lower ends of the box stand for 75th and 25th percentile respectively, while the line inside the box is the median. The top and bottom short horizontal line outside box are the 95th and 5th percentile of the data, respectively. Outliers were indicated as small black circles out of boxes.)

offering specific answers, the proportions of laboratories using single IQC rule 1_{25} and 1_{35} (Group 1 and 2), were totally around 20.0% in 2011 but decreased to slightly exceeding 10.0% in 2016. On the contrary, laboratories choosing one or more control rules of 1_{25} , 1_{35} , 2_{25} , R_{45} , 4_{15} and $10_{\bar{x}}$ (Group 3 and 4), were increasing from a little more

than 30.0% in 2011 to more than 80.0% in 2016. In addition, less and less laboratories used other rules (Group 5) such as 1_{45} , 1_{55} , $6_{\bar{x}}$, $8_{\bar{x}}$, $12_{\bar{x}}$, and $7_{\bar{t}}$ with proportions closing to 0 in 2016.

For each analyte, the calculation of control frequency between two IQC measurements was based on the reported number of

FIGURE 3 Percentages of laboratories with acceptable precision performance against five different allowable imprecision specifications for current CVs of level 1 in five items from 2011 to 2016

IQC observation per month. As shown in Figure 5, laboratories were sorted in five groups according to the control frequency. Laboratories that performed IQC at frequency with time intervals

more than 1 day but less than 2 days (Group 2) occupied the largest percentages in 6 years (69.5%-81.8% for FT3, 71.0%-85.2% for TT3, 71.2%-83.7% for FT4, 71.1%-85.5% for TT4,

FIGURE 4 Percentages of laboratories in different groups of control rules for five items from 2011 to 2016. (Group 1 - 12S; Group 2 - 13S; Group 3 - 13S, 22S; Group 4 - multiple, one or more from the control rules of 12S, 13S, 22S, R4S, 41S and 10; Group 5 - others, other control rules such as 6, 7, 8, 12 and 7T; Group 6 - unclear, including laboratories did not submit control rules or not provide a clear description about IQC or Westgard rules.)

FIGURE 5 Percentages of laboratories in different groups of control frequency for five items from 2011 to 2016. (Group 1, one QC per day ($0 < T \leq 1$); Group 2, one QC every 1~2 days ($1 < T \leq 2$); Group 3, one QC every 2~3 days ($2 < T \leq 3$); Group 4, one QC every 3~4 days ($3 < T \leq 4$); Group 5, more than 4 days between each QC ($T > 4$); T, time intervals (days) between two IQC measurements.)

and 71.3%-81.4% for TSH). Except for 2011, percentages of laboratories conducting IQC once or several times a day (Group 1) were gradually increasing from 2012 (1.0%-2.0%) to 2016

(5.0%-7.0%). By contrast, laboratories in Group 3, 4, and 5 that performed IQC with lower frequency accounted for less and less proportion.

4 | DISCUSSION

In China, nationwide monthly surveys of current CVs and IQC practice have officially started since 2011. This study is the first long-term IQC survey to evaluate the precision performance of FT3, TT3, FT4, TT4, and TSH immunoassays in China. In 2010, there was once a preliminary current study about imprecision analysis and IQC practice for these five analytes.¹² To determine whether the precision performance and IQC practice had improved during the following years, we conducted this long-term study including data from 2011 to 2016. Laboratories constantly attending in IQC survey and submitting their IQC data were chosen as the research subjects to ensure the comparability among different years. However, for five analytes, on average, approximately 500 laboratories participated in the 2011 IQC survey and uploaded their IQC information. Only about 36.0% of them continuously submitted IQC data until 2016, indicating the positivity of clinical laboratories toward IQC survey was not high enough.

By monitoring monthly or long-time CVs of IQC data and calculating the pass rates with the comparison between current CVs and allowable imprecision specifications, clinical laboratories can realize their own precision performance. Here, because the number of current CVs in level 1 conformed to the number of participant laboratories and reflected more complete precision performance in each analyte, we only chose level 1 to analyze. Due to the significant decreasing of current CVs and growing pass rates in general, precision performance of FT3 has improved in 6 years. The calculation of pass rates showed the imprecision specification with looser quality request had higher pass rates than those with tighter quality request.

During 6 years, the highest constituent ratios of Roche measurement system reflected it was more popular than other measurement systems. Because smaller current CVs meant better precision performance, Roche and Abbott systems had better precision level than those of Beckman and Siemens in 2016 for TT3, FT4, TT4, and TSH. But for FT3, precision performance of Beckman system was worse than Abbott, Roche, and Siemens systems in 2016. Among these four main measurement systems, the precision level of Abbott system had got great progress with large fluctuation that needed further observation. If laboratories want to select measurement systems with more stable and higher precision level, recommendation would be Roche and Siemens systems for FT3 but only Roche system for other four analytes.

Additionally, the change of IQC practice can also affect the measurement quality. Reasonable IQC practice can effectively increase the rate of error detection and reduce the rate of false rejection. In this paper, several important impact factors were explored such as control rules and control frequency. IQC rules should be correctly selected to ensure the acceptability of testing results. As the data shown, it was quite good that less and less laboratories misunderstood concepts of IQC rules or used single control rule. Conversely, more and more laboratories applied multiple control rules with growing percentages of approximately 50.0% for all analytes, suggesting knowledge deficiency and incorrect use of IQC rules has been improved. Data on control

frequency can partly reflect the plan of IQC practice. Satisfyingly, results demonstrated that the percentages of laboratories conducting IQC with time interval less than 2 days were gradually increasing in 6 years and up to about 90.0% in 2016 for all analytes. Obviously, the situation of control frequency became more and more better during 6 years. No matter the situation of control rules or control frequency, all showed CNCCL and laboratory directors have made more efforts to IQC improvement.

Now, there has been some electronic IQC systems specifically aiming at selection of IQC material and IQC rules, as well as design of analytical run length for quantitative tests.^{13,14} Clinical laboratories can design individual IQC practice based on analytical and management levels with the help of these systems,

However, this study also demonstrated some deficiencies. First, management documents demanded that medical laboratories need to run at least two concentration levels of control material for each analyte.^{15,16} But it was disappointing that all analytes had about 50.0% laboratories still measured single level of control material until 2016. Half of the participant laboratories did not realize the importance of using multiple levels of control materials. Besides, the measurement range of each level was also not specified, so there may be an overlap between two levels.⁸ Second, most of research subjects (about 80.0%) were in tertiary hospitals which can be considered as representatives of laboratories with better analytical performance and focus more on their quality management than those which did not participant. So, results of this study cannot reflect the overall precision level of these analytes in China. Third, this study only gave a data report without advice on how to improve unsatisfied precision performance. It is high time that we should conduct further precision study including different hospital types and provide guidance to participant laboratories how to improve precision performance according to CVs and IQC practice. In addition, comparing with IQC studies in western countries,^{10,11} CNCCL should enrich Chinese IQC surveys to not only estimate within-laboratory imprecision but also assess between-laboratory and within-method imprecision. Meanwhile, the trend of standardization and harmonization for thyroid hormones immunoassays in abroad^{17,18} is also the direction of development for Chinese EQA schemes.

It is meaningful that CNCCL collect long-term IQC data for statistical analysis. With the change of IQC practice in 6 years, only FT3 had the precision performance improved with decreasing current CVs and increasing pass rates. However, in general, unsatisfying current CVs and pass rates illustrated more efforts should be made to improve precision performance of thyroid hormones immunoassays.

ACKNOWLEDGMENTS

We appreciate those participant laboratories and institutions that attended the EQA schemes for this survey. We also thank Clinet website (www.clinet.com.cn) who gave computer technology support to establish the network platform for survey and relevant services.

DECLARATION OF INTEREST

The submission and publication of this manuscript are both approved by all authors without any conflict of interest. On behalf of all authors, I declare that this work was an original research and has not been under review or consideration for publication and published previously in other journals.

REFERENCES

1. Singh RJ, Kaur P. Thyroid hormone testing in the 21st century [J]. *Clin Biochem*. 2016;49:843–845.
2. Weiping T, Xiaoping X, Nanwei T, et al. An epidemiology survey of thyroid diseases of ten cities in China [D]. Proceedings of 9th Annual Meeting of Chinese Society of Endocrinology, PL-07. (in Chinese)
3. Jiawei S, Xiaojun X, Qiumao C, et al. Epidemiological study on thyroid cancer in China [J]. *China Cancer*. 2013;22:690–693. (in Chinese).
4. Ross DS, Burch HB, Cooper DS, et al. 2016 American Thyroid Association Guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis [J]. *Thyroid*. 2016;26:1343–1421.
5. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for adult patients with thyroid nodules and differentiated thyroid cancer [J]. *Thyroid*. 2016;26:1–133.
6. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem*. 1974;20:825–833.
7. Krouwer JS. The problem with total error model in establishing performance specifications and a simple remedy [J]. *Clin Chem Lab Med*. 2016;54:1299–1301.
8. Rong Z, Wei W, Haijian Z, et al. National survey on internal quality control for HbA(1c) Analytical instruments in 331 Hospital laboratories of China [J]. *Clin Lab*. 2015;61:453–460.
9. Fei Y, Wei W, Falin H, et al. Imprecision investigation and analysis of routine chemistry in China [J]. *J Clin Lab Anal*. 2016;30:444–450.
10. Giovannini S, Zucchelli GC, Iervasi G, et al. Multicentre comparison of free thyroid hormones immunoassays: the Immunocheck study [J]. *Clin Chem Lab Med*. 2011;49:1669–1676.
11. Prontera C, Zaninotto M, Giovannini S, et al. Proficiency testing project for brain natriuretic peptide (BNP) and the N-terminal part of the propeptide of BNP (NT-proBNP) immunoassays: the CardioOrmocheck study [J]. *Clin Chem Lab Med*. 2009;47:762–768.
12. Litao H, Zhiguo W. Internal quality control practice of thyroid disease related tests and imprecision analysis in China [J]. *Clin Lab*. 2014;60:301–308.
13. Litao H. *The Computer Simulation Study and Application of Analytical Run Length for Quantitative Tests in Clinical Laboratory* [D]. Beijing: Peking Union Medical College, Chinese Academy of Medical Sciences; 2012: 7–45.
14. Kinns H, Pitkin S, Housley D, et al. Internal quality control: best practice [J]. *J Clin Pathol*. 2013;66:1027–1032.
15. CLSI C24-A3. *Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions; Approved Guideline - Third Edition* [S]. Wayne, PA: Clinical Laboratory Standards Institute; 2006.
16. Westgard JO. *Basic QC practices*, 3rd edn [M]. Madison, WI: Westgard QC, Inc; 2010.
17. Clerico A, Ripoli A, Zucchelli GC, et al. Harmonization protocols for thyroid stimulating hormone (TSH) immunoassays: different approaches based on the consensus mean value [J]. *Clin Chem Lab Med*. 2015;53:377–382.
18. Stockl D, Van Uytfaange K, Van Aelst S, et al. A statistical basis for harmonization of thyroid stimulating hormone immunoassays using a robust factor analysis model [J]. *Clin Chem Lab Med*. 2014;52:965–972.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Zhang S, Wang W, Zhao H, et al. Status of internal quality control for thyroid hormones immunoassays from 2011 to 2016 in China. *J Clin Lab Anal*. 2018;32:e22154. <https://doi.org/10.1002/jcla.22154>