RESEARCH ARTICLE

WILEY

Utility of process capability indices in assessment of quality control processes at a clinical laboratory chain

Ping Dong[1] | Yong-Bo Wang[2] | De-Zhi Peng[1] | Jia-Jia Wang[1] | Ya-Ting Cheng[3,4] | Xiao-Yan Deng[4] | Biao Zheng[4] | Ran Tao[3,4] 🆔

[1]Laboratory Diagnosis Department, Jinan KingMed Center for Clinical Laboratory, Jinan, China

[2]Laboratory Diagnosis Department, Qingdao KingMed Center for Clinical Laboratory, Qingdao, China

[3]Laboratory Diagnosis Department, Guangzhou KingMed Center for Clinical Laboratory, Guangzhou, China

[4]KingMed School of Laboratory Medicine, Guangzhou Medical University, Guangzhou, China

Correspondence
Ran Tao, Laboratory Diagnosis Department, Guangzhou KingMed Center for Clinical Laboratory, No.10 Luoxuan 3rd road, International Biological Island, Guangzhou 510005, China.
Email: labtr@kingmed.com.cn

Abstract

Background: To evaluate the utility of the process capability indices C_p and C_{pk} for assessing the quality control processes at chain laboratory facilities.

Methods: In April 2020, the minimum C_p and C_{pk} values for 33 assays of a laboratory chain with 19 facilities were collected for further analysis and a total of 627 datasets (C_p and C_{pk}) were compared. In addition, standard values for C_p and C_{pk}, defined as the lowest of the top 20%, were obtained for comparison and the indices were used to determine whether precision or trueness improvements were required for the corresponding assay.

Results: A total of 627 datasets of 33 assays from 19 laboratory facilities were collected for further analysis. Based on the C_p results, 329 (52.5%), 211 (33.7%), 65 (10.3%), and 22 (3.5%) were rated as excellent, good, marginal, and poor, respectively. While the corresponding results for C_{pk} were 300 (47.8%), 216 (34.4%), 79 (12.6%), and 32 (5.1%). In addition, it was noteworthy that eight (C_p criteria) and six assays (C_{pk} criteria) were rated as excellent or good at all 19 facilities. Comparison of the process capability indices at the Jinan KingMed Center with the standard values revealed that total protein, albumin, and urea showed trueness individual improvement, precision individual improvement, and precision common improvement, respectively, while the results of other assays were stable.

Conclusion: Process capability indices are useful for evaluating the quality control procedures in laboratory facilities and can help improve the precision and trueness of laboratory tests.

KEYWORDS
improvement, precision, process capability index, quality control, trueness

1 | INTRODUCTION

In laboratory medicine, internal quality control (IQC) and external quality assessment (EQA) are two important processes for quality management.[1,2] Several tools have been developed for internal quality management. For example, the coefficient of variation (CV) and bias are used to evaluate the precision and trueness of assays, respectively. Sigma metrics, (TEa-|bias|)/CV, are employed to evaluate the analytical performance of a test[3,4]; for example, a Sigma value >3

indicates that testing procedure is able to meet clinical need and can be controlled using a selection of standard Westgard rules.[5] However, when the Sigma value is ≤3, the quality goal index (QGI) needs to be calculated, which can help the laboratory to analyze the cause of bad performance and propose improvement measures. QGI < 0.8 indicates that the main reason for the bad performance of the method is that the precision exceeds the allowable range, and the priority is to improve the precision. When QGI > 1.2, the trueness of the proposed method is poor, and the trueness is improved first. QGI of 0.8–1.2 indicates that both trueness and precision need to be improved.[6]

External quality assessment has several following challenges: (1) although bias is used to evaluate the trueness of a test and can be estimated based on the EQA, it is difficult to identify the bias in a timely manner due to the long period required for the assessment, which may have a significant effect on the laboratory test results; (2) there remains a lack of evaluation factors (eg, analyzer or reagent) associated with precision or trueness, and it appears to be impossible to determine whether a problem is limited to one laboratory or exists in all laboratories in a chain. The standard deviation index (SDI) and coefficient of variation index (CVI) are designed to evaluate factors associated with trueness and can help identify the key point necessary for improvement.[7] However, the calculation of SDI and CVI requires a special tool and potentially additional time, which contributes to an extended lag period.

In 1996, the process capability indices C_p and C_{pk} used in the manufacturing industry were introduced to laboratory medicine and their utility for selecting appropriate QC rules were assessed by Burnett et al.[8] The higher the values of the process capability indices are, the greater the number of products that can be produced under the permitted specifications.[9,10] Interestingly, an algebraic relationship was noted between the C_p, C_{pk}, and Sigma values, suggesting that the indices may be useful to improve the trueness or precision of an assay. Hence, the process capability indices can indicate the following: (1) whether the performance of an assay is abnormal; (2) the error is random or systematic; and (3) the abnormal finding only occurs within an individual laboratory or exists in most laboratories. Therefore, in this study, the process capability indices C_p and C_{pk} were assessed for the quality control of a clinical laboratory chain and their utility was also evaluated in a single center (Jinan KingMed Center).

2 | MATERIALS AND METHODS

2.1 | Facilities

Between April 1 and 30, 2020, 19 laboratories from a clinical laboratory chain (KingMed Center for Clinical Laboratory) participated in the investigation.

2.2 | Tests

Thirty-three tests, specifically alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), creatine kinase (CK), creatine kinase isoenzyme MB (CK-MB), hydroxybutyrate dehydrogenase (HBDH), total protein (TP), albumin (Alb), total bilirubin (TBIL), direct bilirubin (DBIL), glucose (Glu), urea (Urea), creatinine (Cr), uric acid (UA), triglycerides (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), apolipoprotein A1 (APOA1), apolipoprotein B (APOB), immunoglobulin A (IgA), immunoglobulin G (IgG), immunoglobulin M (IgM), complement 3 (C3), complement 4 (C4), antistreptolysin O (ASO), rheumatoid factor (RF), C-reactive protein (CRP), Potassium (K), sodium (Na), and Chlorine (Cl) assays, were included in the analysis.

2.3 | C_p and C_{pk}

Process capability is an indicator that describes the relationship between the three parameters in the manufacturing process: allowable specification, centralized tendency, and variance. The better the capability, the greater the number of products that can be produced within the permitted specifications.[8] For laboratory medicine, allowable specification, centralized tendency, and variance are equivalent to the total allowable error (TEa), the difference between the cumulative mean and the fixed mean of quality control data, and the variation of quality control data. The process capability index includes the index for bilateral specifications (C_p) and the index of existing unilateral bias (C_{pk}). C_p emphasizes the relationship between total allowable error (TEa) and the imprecision. It does not consider any bias. The calculation formula is as follows: (USL-LSL)/6σ. USL and LSL are the upper- and lower-specification limits. σ is the standard deviation (SD) of the process.[8,11] For laboratory medicine, C_p =2*TEa/(6*CV), that is C_p = TEa/(3*CV).

In addition to considering TEa and imprecision, C_{pk} also increases the factor of bias between the cumulative mean and the fixed mean, which reflects imprecision and untrueness simultaneously. C_{pk} is defined as follows, where USL, LSL, and σ are for C_p, and μ is the mean of the process:

$$Cpk = \min\left[\frac{\mu - LSL}{3\sigma}, \frac{USL - \mu}{3\sigma}\right]$$

If the bias is nonzero and the specification limits are symmetrical with regard to the "true value" ($\overline{x_0}$), then: USL = $\overline{x_0}$ + TEa, LSL = $\overline{x_0}$ − TEa. Therefore, C_{pk} is converted as follows[8,11]:

$$Cpk = \min\left[\frac{TEa + \mu - \overline{x_0}}{3\sigma}, \frac{TEa + \overline{x_0} - \mu}{3\sigma}\right]$$

For laboratory medicine, the difference between μ and $\overline{x_0}$ can be regarded as bias. So C_{pk} = (TEa-|bias|)/(3*CV).

The TEa was derived from the *Analytical quality specifications for routine analytes in clinical biochemistry*,[12] or the standard released by the National Health Commission of the People's Republic of China for EQA.[13]

2.4 | Absolute criteria for C_p or C_{pk}

Analytes with a C_p or $C_{pk} \geq 2$ were considered "excellent" for assays (green). Similarly, analytes with $1.33 \leq C_p$ or $C_{pk} < 2$ were considered to be "good" for assays (light green), those with $1 \leq C_p$ or $C_{pk} < 1.33$ were considered to be "marginal" for assays (light red), and those with C_p or $C_{pk} < 1$ were considered "poor" for assays (red).[8] Table 1 shows the relationships between C_p, C_{pk}, and the corresponding Sigma values. If an analyte had C_p and C_{pk} values of 2.12 and 1.23, respectively, the assay had excellent precision and marginal trueness. Therefore, measures to improve the trueness were necessary.

2.5 | Relative criteria for C_p or C_{pk}: Evaluation at a single center

The utility of C_p and C_{pk} at a single center (Jinan KingMed Center) was evaluated. Firstly, standard values for C_{pk} and C_p, defined as the lowest of the top 20% results, were obtained for comparison and the indices collected from Jinan KingMed Center were used to determine whether the assays required improvement according to the criteria listed in Table 2.

2.6 | System designing

The quality control products used by 19 laboratories include *BIO-RAD* Assayed Chemistry Control, *BIO-RAD* Lipids Control, *BIO-RAD* Cardiac Markers Plus Control, and *CLINIQA* Serum Protein Control etc. Generally, two concentrations are used for each item and tested at least once a day. The equipment and reagents used by each laboratory are different. Take ALT test as an example, Jinan KingMed Center uses Beckman AU5800 Automatic Analyzer and Maccura reagent, while Roche c702 Automatic Analyzer and Roche reagent are used in Guangzhou Kingmed Center. The laboratories involved in the comparison had a variety of testing systems, including different equipment, reagents, and calibrators.

The comparison process of quality control data contains two parts (Figure 1) as follows: internal quality control system: daily calculation the C_p and C_{pk} of each QC data in every laboratory; external comparison system: extraction and comparison of the minimum C_p and C_{pk} every month.

3 | RESULTS

3.1 | Performance assessment

3.1.1 | C_p

Based on the C_p, 329 of the 627 datasets (52.5%) were rated as excellent, 211 (33.7%) were rated as good, 65 (10.3%) were rated as marginal, and 22 (3.5%) were rated as poor. Eight tests, specifically HBDH, UA, LDL-C, IgM, C3, C4, ASO, and RF, were all rated as good or excellent, and the others were partially rated as marginal or poor (Table 3).

3.1.2 | C_{pk}

Based on the C_{pk}, 300 of the 627 datasets (47.8%) were rated as excellent, 216 (34.4%) were rated as good, 79 (12.6%) were rated as marginal, and 32 (5.1%) were rated as poor. Six tests, namely UA, LDL-C, IgM, C3, C4, and ASO, were all rated as good or excellent, and the others had partial results rated as marginal or poor.

### 3.2 | Utility of C_p and C_{pk}: Evaluation at a single center

The data collected from the Jinan KingMed Center (Table 4) were evaluated according to the rules in Table 2. Based on the results, suggestions were provided and measures were implemented to improve performance. According to the C_p and C_{pk} of the analytes, three tests, namely TP, Alb, and Urea, were found to require following improvements: (1) the precision and trueness of TP test were rated as good and marginal respectively and the evidence suggested that the trueness should be improved at the Jinan KingMed Center;

TABLE 1 The relationships between C_p, C_{pk}, and corresponding Sigma values and their potential implications

C_p value	C_p rating	C_{pk} value	C_{pk} rating	Sigma value	Implications
$C_p \geq 2$	Excellent	$C_{pk} \geq 2$	Excellent	Sigma ≥ 6	N/A
$C_p \geq 2$	Excellent	$1.33 \leq C_{pk} < 2$	Good	$4 \leq$ Sigma < 6	N/A
$1.33 \leq C_p < 2$	Good	$1.33 \leq C_{pk} < 2$	Good	$4 \leq$ Sigma < 6	N/A
$C_p \geq 2$	Excellent	$1 \leq C_{pk} < 1.33$	Marginal	$3 \leq$ Sigma < 4	Trueness
$1.33 \leq C_p < 2$	Good	$1 \leq C_{pk} < 1.33$	Marginal	$3 \leq$ Sigma < 4	Trueness
$1 \leq C_p < 1.33$	Marginal	$1 \leq C_{pk} < 1.33$	Marginal	$3 \leq$ Sigma < 4	Precision
$C_p \geq 2$	Excellent	$0 \leq C_{pk} < 1$	Poor	Sigma < 3	Trueness
$1.33 \leq C_p < 2$	Good	$0 \leq C_{pk} < 1$	Poor	Sigma < 3	Trueness
$1 \leq C_p < 1.33$	Marginal	$0 \leq C_{pk} < 1$	Poor	Sigma < 3	Precision
$0 \leq C_p < 1$	Poor	$0 \leq C_{pk} < 1$	Poor	Sigma < 3	Precision

TABLE 2 Suggestions based on comparisons of the standard values with the C_p and C_{pk} collected from a single center

C_p and C_{pk} at a single center	Standard values for C_p and C_{pk}	Suggestions	
Marginal/poor	Excellent/good	Individual improvement	Practice standardization
Marginal/poor	Marginal/poor	Common improvement	Shared problems among laboratories, such as reagent quality, TEa setting
Excellent/good	Excellent/good/marginal/poor	Stable	Keeping
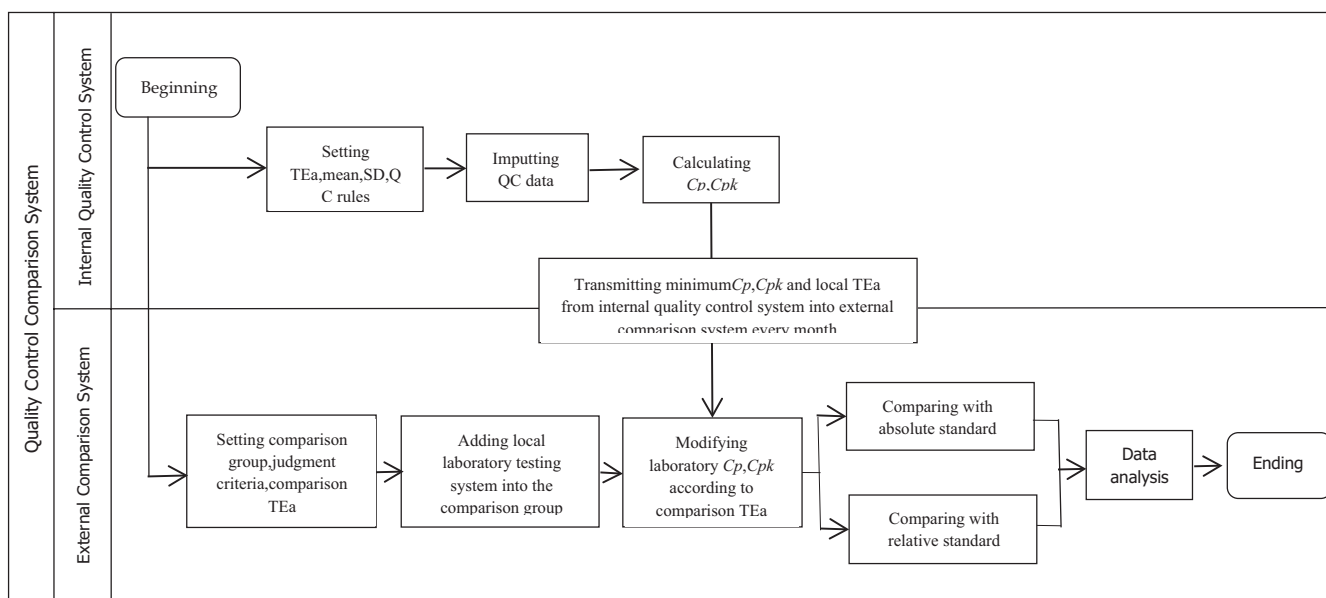
FIGURE 1 The external comparison procedure of IQC data based on C_p and C_{pk}

(2) the Alb test was rated as a marginal and poor assay. In this center, attention should be paid to the precision and trueness, and more effort should be devoted to improving the precision; (3) the Urea tests in most facilities were rated as marginal, which suggested that the precision and trueness had to be improved in most facilities and precision should be given first priority for improvement in the hospital chain. With the exception of these assays, the C_p and C_{pk} in other tests were stable.

4 | DISCUSSION

Precision is usually evaluated through IQC, and bias can be measured by EQA or comparing the IQC data to evaluate the performance of assays. This process can help laboratories improve the performance of the detection system.[1] Several tools have been developed to evaluate the performance of analytes, such as the Sigma value and QGI. When the Sigma value is ≤3, the quality goal index (QGI) needs to be calculated. A QGI score <0.8 indicates imprecision, a score higher than 1.2 indicates untrueness, and a score between 0.8 and 1.2 indicates both imprecision and untrueness. The combined application of Sigma and QGI can identify key point for improvement (precision or trueness) based on IQC datas.[6] But due to the presence of two

variables, bias and CV, it is impossible to further analyze the causes of imprecision or untrueness by means of Sigma external comparison. Another tool is SDI and CVI, which can identify key point for improvement (precision or trueness) and the causes of abnormal based on IQC data.[7] But some disadvantages limit its wider application as follows: (1) influenced by the quality specifications of the laboratories involved in the comparison, the artificial division of "outliers" in statistics and the number of laboratories in the comparison group, false positive, or false negative would occur. When false positive, correctional measures would be made but the effect would be unsatisfactory, and when false negative, abnormal results would be reported due to no correction; (2) commercial software such as Unity Real Time is required,[7,14] and laboratory costs are high; (3) only applicable to the use of specific quality control items, cannot cover all tests; and (4) direct use of parameters bias and CV for calculation cannot exclude the quality difference caused by the QC material, so it can only be applied to the same manufacturer's QC products with the same batch number, and the comparison has major limitations.

In view of the above questions, we have been trying to find another solution method. We learned the calculation thought of C_p and C_{pk} from MultiQC software in 2015.[15] This software introduces C_p and C_{pk} into IQC analysis. Coincidentally, we were developing an laboratory-built quality control software at that

TABLE 3 Assays rated based on the C_p or C_{pk} value among the 19 laboratory facilities

Tests	TEa (%)	C_p				C_{pk}			
		Excellent	Good	Marginal	Poor	Excellent	Good	Marginal	Poor
ALT	±16	9	8	2	0	6	11	2	0
AST	±15	8	9	2	0	7	6	6	0
GGT	±11	8	8	3	0	7	9	3	0
ALP	±18	8	9	2	0	7	9	3	0
LDH	±11	8	7	2	2	8	7	1	3
CK	±15	9	8	2	0	8	8	3	0
HBDH	±30	15	4	0	0	12	6	1	0
CK-MB	±25	7	7	4	1	6	8	3	2
TP	±5	3	9	4	3	2	7	7	3
Alb	±6	0	9	4	6	0	8	6	5
TBil	±15	13	4	2	0	11	4	3	1
DBil	±15	12	6	1	0	11	5	2	1
Glu	±7	4	11	2	2	4	9	4	2
Urea	±8	0	5	9	5	0	4	9	6
Cr	±12	11	5	3	0	10	6	3	0
UA	±12	14	5	0	0	13	6	0	0
TG	±14	14	4	1	0	13	5	1	0
TC	±9	9	8	1	1	7	9	2	1
HDL-C	±30	13	5	1	0	12	5	1	1
LDL-C	±30	15	4	0	0	14	5	0	0
APOA1	±30	12	6	1	0	11	6	1	1
APOB	±30	12	6	1	0	12	6	1	0
IgA	±25	13	5	1	0	13	4	2	0
IgG	±25	11	7	1	0	10	8	0	1
IgM	±25	15	4	0	0	14	5	0	0
C3	±25	14	5	0	0	14	5	0	0
C4	±25	13	6	0	0	12	7	0	0
ASO	±25	13	6	0	0	12	7	0	0
RF	±25	15	4	0	0	14	4	1	0
CRP	±25	12	6	1	0	11	6	1	1
K	±6	8	7	4	0	8	7	3	1
Na	±4	4	8	7	0	4	8	7	0
Cl	±4	7	6	4	2	7	6	3	3

time, so the indexes of C_p and C_{pk} were also put into the software we developed.[8,11,16,17] As CV is used for the calculation of C_p, CV and bias are both included in the calculation formula for C_{pk}, the two indices can intuitively reflect problems associated with technical procedures. Compared with other methods, the following advantages of C_p and C_{pk} seem clear: (1) because C_p and C_{pk} are compared with specific standard values, the key point to be improved (precision or trueness) can be determined, and the data are more authoritative; (2) the further reasons of imprecision or untrueness (common factor or individual factor) can be acquired by comparative analysis with a relatively excellent laboratory as the benchmark; (3) compared with a single laboratory, it can reflect

the actual level of the laboratory better by excluding the influence of different quality specifications of multiple laboratories and the inherent defects of statistical method; (4) the software for C_p and C_{pk} analysis is self-developed and designed. There is no subsequent investment in software and (5) as the transformed IQC data are used for comparison, the influence of different manufacturers and different batch numbers has been excluded, and all items can be evaluated, so the application range is wider.

A major difference between the process capability indices and Sigma metrics is the method used to measure bias.[18] The traceability of a test is thought to be determined by the calibrator, reagent, and analyzer. Two methods are usually employed in practice to evaluate

TABLE 4 Comparison of C_p and C_{pk} between the standard values and data collected from Jinan KingMed Center

Tests	C_p			C_{pk}			Conclusion
	Value	Grade	Standard	Value	Grade	Standard	
ALT	1.75	Good	2.36	1.65	Good	2.30	Stable
AST	2.83	Excellent	1.99	2.3	Excellent	1.8	Stable
GGT	2.12	Excellent	2.44	1.9	Good	2.43	Stable
ALP	1.84	Good	2.47	1.68	Good	2.44	Stable
LDH	1.57	Good	2.46	1.36	Good	2.35	Stable
CK	3.11	Excellent	2.88	2.88	Excellent	2.82	Stable
HBDH	2.11	Excellent	3.4	2.03	Excellent	3.4	Stable
CK-MB	2.26	Excellent	4.57	1.78	Good	2.44	Stable
TP	1.42	Good	1.39	1.21	Marginal	1.36	Trueness individual improvement
Alb	1.01	Marginal	1.45	0.98	Poor	1.4	Precision individual improvement
TBil	2.45	Excellent	3.52	2.37	Excellent	2.47	Stable
DBil	4.73	Excellent	4.35	3.87	Excellent	4.23	Stable
Glu	1.53	Good	1.96	1.46	Good	1.74	Stable
Urea	1.08	Marginal	1.31	1.04	Marginal	1.24	Precision common improvement
Cr	1.94	Good	2.92	1.75	Good	2.84	Stable
UA	3.03	Excellent	3.28	2.477	Excellent	3.27	Stable
TG	3.11	Excellent	3.67	3.08	Excellent	3.62	Stable
TC	1.54	Good	2.55	1.47	Good	2.18	Stable
HDL-C	2.15	Excellent	4.10	1.95	Good	3.98	Stable
LDL-C	2.3	Excellent	5.61	2.26	Excellent	5.36	Stable
APOA1	3.27	Excellent	4.71	2.9	Excellent	4.32	Stable
APOB	2.39	Excellent	5.82	2.07	Excellent	5.05	Stable
IgA	3.79	Excellent	5.84	3.77	Excellent	5.68	Stable
IgG	3.13	Excellent	5.58	3.1	Excellent	5.43	Stable
IgM	5.42	Excellent	7.09	5.32	Excellent	5.94	Stable
C3	4.85	Excellent	6.86	4.1	Excellent	6.44	Stable
C4	4.42	Excellent	10.08	4.13	Excellent	9.21	Stable
ASO	4.09	Excellent	3.61	3.81	Excellent	3.57	Stable
RF	4.17	Excellent	4.53	3.59	Excellent	4.38	Stable
CRP	2.04	Excellent	2.95	2.64	Excellent	2.89	Stable
K	2.20	Excellent	2.81	1.76	Excellent	2.50	Stable
Na	2.12	Excellent	1.61	1.94	Good	1.58	Stable
Cl	1.96	Good	1.99	1.85	Good	1.82	Stable

systematic error as follows: (1) calculation of difference according to EQA results provided by the organizer and (2) calculation of difference between cumulative mean and the fixed mean of IQCs. Although the bias calculated for the EQA results can be used, several following factors should be taken into consideration: (1) the levels of analytes in IQC samples are usually not the same as the levels in EQA samples; (2) the influence of the matrix effect on IQC and EQA samples remains unclear[19-21]; (3) difference in factors (eg, reagent and analyzer) associated with traceability between laboratories makes the situation more complicated, so the obtained bias value does not necessarily reflect the true technical level of the laboratory; (4) not all items have an official EQA plan in China and (5) the detection and evaluation cycle of the EQA is so long. For example, General Chemistry organizes three times a year. That means the EQA results are obtained once every 3 months on average. The organization frequency of other specialties is lower. So, the bias that can reflect the status of the detection system cannot be obtained in real-time.

C_p is equal to Sigma/3 when the bias is 0, suggesting that C_p can reflect the imprecision of analytes objectively. Furthermore, C_{pk}, which is calculated based on Sigma/3 and bias, can reflect the

trueness and precision of analytes. Hence, when C_p and C_{pk} are combined, it would be helpful to know whether the trueness or the precision needs to be improved. When both C_p and C_{pk} are ≥1.33, which means Sigma ≥4, the laboratory can determine its quality control scheme (including quality control rules, levels of QCs, and QC frequency) in accordance with the Westgard Sigma Rules.[22,23] When either C_p or C_{pk} is <1.33, Sigma is <4, suggesting the laboratory should conduct a comprehensive analysis of C_p and C_{pk} to determine which parameter requires improvement. In addition, standard C_p and C_{pk} values were selected from the top 20% value of the clinical laboratory chain for comparison. This is helpful for identifying abnormal results caused by a "common factor" or "individual factor." A standard C_p or C_{pk} value >1.33 suggests the test is excellent or good. However, when C_p or C_{pk} for a single center is <1.33, an individual factor may be responsible for the bad performance of the tests and correctional measures, such as maintenance, standard operation, and new reagents/instruments should be taken within this individual laboratory.

In our study, lipid analysis showed a poor performance in one laboratory. In this laboratory, TC, TG, and APOB had low C_p and C_{pk} values (<1.33), suggesting the precision for these assays required improvement. Based on the C_{pk} value, APOA1 was rated as a poor assay and improvements to the trueness were recommended. Although the C_p and C_{pk} values for the HDL-C and LDL-C assays at this facility indicated good performance, the assays still required improvement compared with those at the other laboratories. According to the data presented in Table 3, an error in the assessment of the lipid analysis was suspected and this was confirmed by subsequent analysis. The Liquichek™ Lipids control purchased from BIO-RAD needs to be stored at −20 to −70℃. However, it was incorrectly stored at 2–8℃. Once new controls were used, the tests improved significantly and the corresponding C_p and C_{pk} values returned to normal.

For the standard values defined at the top 20% of results, a standard C_p or C_{pk} value <1.33 indicates that the assay has a marginal or poor status in most laboratories, which may introduce a slight TEa limitation in the analytical method or the difference between different reagents. In our study, the utility of the process capability indices was assessed in a single center (Jinan Kingmed Center). Analytes such as Urea, TP, and Alb had relatively low standard C_p and C_{pk} values. The results suggested that the poor performance for the Urea assay was a common problem and required improvement in most laboratories. Although the TP/Alb assay showed individual improvement, the standard C_p and C_{pk} values were only a litter higher than 1.33. The bad performance for those three assays may be attributed mostly to the small Tea (8%, 5% and 6%, respectively), which was defined by the National Health Commission of China in 2012 and must be used compulsively.[12] While failing a criteria may be a signal, comparison with levels of the Milan criteria, for example, biological variation and state of the art, can allow alternate assessment of the assay quality. So, we compared the TEa of those three tests used in China to the European Society for External Quality Assessment (ESfEQA), which were 20%, 10%, and 20%, respectively. It showed that the

TEa of China is more stringent and this may explain the unsatisfactory result.[24,25] Further analysis demonstrated that when the C_p and C_{pk} values for Urea, TP, and Alb were calculated using the TEa defined by the ESfEQA,[26] all results were rated as excellent or good. Perhaps, through this study, we can feed back to the standard-setters the excessively strict TEa of some tests, so as to ensure the applicability and conformity of performance specifications. By comparing the C_p and C_{pk} values for analytes between data collected from a single center and the standard values, the cause of abnormal results can be easily identified and the performance of the tests can be efficiently improved.

As mentioned above, besides evaluating the performance of assays, the process capability indices are also used to analyze the cause of abnormal QC data and determine whether an immediate correction is required. If a process capability index is low, more quality control rules and immediate interventions may be required. For example, when the C_p and C_{pk} values are 1.15 and 1.21, respectively, and two consecutive control results are outside the limits in the same direction (Westgard rule 2_{2S}), a systematic error is assumed. However, the two process capability indices demonstrated that both trueness and precision require improvements and precision should receive the first priority. Conversely, if C_p and C_{pk} are both >2, even if the result for the control is outside the limit (Westgard rule 1_{3S}), the analytical process is still acceptable and no additional measure is required.

Compared with traditional methods for quality management in laboratory facilities, the process capability indices C_p and C_{pk} have several advantages. These indices can help laboratories discover issues with assays through comprehensive analysis of both parameters collected from chain laboratories and by comparing C_p and C_{pk} between data collected from a single center and the standard values in the top 20%. The precision and trueness of laboratory tests can be improved significantly using this approach. In addition, based on our experience with 19 facilities, the process capability indices C_p and C_{pk} may be applicable for QC management in all KingMed Center laboratories and can make the QC process more standardized and practical.

ACKNOWLEDGEMENTS

CONFLICT OF INTERESTS
The authors declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

AUTHOR CONTRIBUTIONS
Ran Tao and Ping Dong researched literature and conceived the study. Yong-Bo Wang, De-Zhi Peng, Biao Zheng, and Xiao-Yan Deng were involved in protocol development, gaining ethical approval, patient recruitment, and data analysis. Ping Dong, Jia-Jia Wang, and Ran Tao wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ran Tao https://orcid.org/0000-0003-1126-6050

REFERENCES

1. Westgard JO, Westgard SA. Quality control review: implementing a scientifically based quality control system. *Ann Clin Biochem.* 2016;53:32-50.
2. Jones GRD, Albarede S, Kesseler D, et al. Analytical performance specifications for external quality assessment – definitions and descriptions. *Clin Chem Lab Med.* 2017;55:949-955.
3. Petrides V, Schneider S. Using Sigma metrics to establish analytical product performance requirements and optimize analytical performance of an in vitro diagnostic assay using a theoretical total PSA assay as an example. *Biochem Med.* 2018;28:020903.
4. Shaikh MS, Moiz B. Analytical performance evaluation of a high-volume hematology laboratory utilizing sigma metrics as standard of excellence. *Int J Lab Hematol.* 2016;38:193-197.
5. Wang H, Ma Y, Shan X. Evaluating the analytical quality control of urinary albumin measurements using sigma metrics. *Clin Biochem.* 2019;73:109-111.
6. Sayeed S, Ganji SB, Mopuri R. A short-term assessment of routine chemistry parameters by sigma metrics and quality goal index ratio in a tertiary care hospital laboratory. *J Evolution Med Dent Sci.* 2019;8:2303-2306.
7. Yoon YA, Jang MA, Lee JS, et al. Effect of accreditation on the trueness of diagnostic hematologic tests: standard deviation index analysis. *Ann Lab Med.* 2018;38:67-70.
8. Burnett L, Hegedus G, Chesher D, et al. Application of process capability indices to quality control in a clinical chemistry laboratory. *Clin Chem.* 1996;42:2035-2037.
9. Aslam M, Azam M, Jun CH. A mixed repetitive sampling plan based on process capability index. *Appl Math Model.* 2013;37:10027-10035.
10. Aslam M, Wu C-W, Azam M, et al. Variable sampling inspection for resubmitted lots based on process capability index Cpk for normally distributed items. *Appl Math Model.* 2013;37:667-675.
11. Chesher D, Burnett L. Equivalence of critical error calculations and process capability index Cpk. *Clin Chem.* 1997;43:1100-1101.
12. National Health Commission of China. Analytical quality specifications for routine analytes in clinical biochemistry. 2013. http://www.nhc.gov.cn/wjw/s9492/201301/c8dd48222ab14387a6503667be78bec3.shtml. Accessed April 10, 2020.
13. National Center for Clinical Laboratories of China. The external quality assessment standard released by the National Health Commission of the People's Republic of China. 2019. https://www.nccl.org.cn/showEqaPlanProDetail?id=146. Accessed April 10, 2020.
14. BIO-RAD. Unity™ Interlab reports. https://www.qcnet.com. Accessed April 10, 2020.
15. MultiQC 5.2-user manual. http://www.sble.fr/wordpress/wp-content/uploads/2011/05/MultiQC5.pdf. Accessed April 10, 2020.
16. Niza-Ribeiro J, Noordhuizen JP, Menezes JC. Capability index–a statistical process control tool to aid in udder health control in dairy herds. *J Dairy Sci.* 2004;87:2459-2467.
17. Oliva A, Monzón C, Santoveña A, et al. Development of an ultra high performance liquid chromatography method for determining triamcinolone acetonide in hydrogels using the design of experiments/design space strategy in combination with process capability index. *J Sep Sci.* 2016;39(14):2689-2701.
18. Thelen MHM, Jansen RTP, Weykamp CW, et al. Expressing analytical performance from multi-sample evaluation in laboratory EQA. *Clin Chem Lab Med.* 2017;55:1509-1516.
19. Badrick T, Punyalack W, Graham P. Commutability and traceability in EQA programs. *Clin Biochem.* 2018;56:102-104.
20. Jennings I, Kitchen D, Kitchen S, et al. The importance of commutability in material used for quality control purposes. *Int J Lab Hematol.* 2019;41:39-45.
21. Zhang S, Zeng J, Zhang C, et al. Commutability of possible external quality assessment materials for cardiac troponin measurement. *PLoS One.* 2014;9:1-9.
22. Westgard S, Bayat H, Westgard JO. Analytical Sigma metrics: a review of Six Sigma implementation tools for medical laboratories. *Biochem Med.* 2018;28:020502.
23. Kumar BV, Mohan T. Sigma metrics as a tool for evaluating the performance of internal quality control in a clinical chemistry laboratory. *J Lab Physicians.* 2018;10:194-199.
24. Hens K, Berth M, Armbruster D, Westgard S. Sigma metrics used to assess analytical quality of clinical chemistry assays: importance of the allowable total error (TEa) target. *Clin Chem Lab Med.* 2014;52:973-980.
25. Westgard JO, Westgard SA. Assessing quality on the Sigma scale from proficiency testing and external quality assessment surveys. *Clin Chem Lab Med.* 2015;53:1531-1535.
26. European Society for External Quality Assessment. Permissible ranges of European Society for External Quality Assessment. https://www.esfeqa.eu/fileadmin/media/esfeqa.eu/pdf/Permissible_Ranges_ESfEQA_Version_08.20_conventional_units_04_08_2020.pdf (Version: 08.2020). Accessed April 10, 2020.