Clinica Chimica Acta 492 (2019) 57–61

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

Practical application of the sigma-metric run size nomogram for multistage bracketed statistical quality control analysis of eight enzymes

Yuping Zeng[a,b], He He[b], Ken Qin[c], Mei Zhang[b], Zhenmei An[d], Hengjian Huang[b,*]

[a] West China School of Medicine, West China Hospital, Sichuan University, Chengdu, China
[b] Department of Laboratory Medicine, West China Hospital of Sichuan University, Chengdu, China
[c] Health Management Center, West China Hospital of Sichuan University, Chengdu, China
[d] Department of Endocrinology and Metabolism, West China Hospital of Sichuan University, Chengdu, China

ARTICLE INFO

Keywords:
Risk management
Nomogram
Multistage quality control
Quality improvement

ABSTRACT

Background: A sigma-metric run size nomogram is used to recommend quality control (QC) strategies to reduce patient risks. Herein, we aimed to evaluate the sigma performance of 8 enzymes and apply multistage bracketed statistical QC (SQC).

Methods: Sigma performance of alanine transaminase (ALT), aspartate transaminase (AST), gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), creatine kinase (CK), amylase (AMY), and lipase (LIP) were determined. Daily workload of each test was estimated and expected reporting QC intervals were designed. Per the nomogram, "start-up" and "monitor" QC rules were determined from sigma performance. SQC was finally applied, followed by quality improvement.

Results: Sigma metrics were as follows: 5.26 (ALT), 4.80(AST), 5.25(GGT), 3.36(ALP), 4.71(LDH), 15.45(CK), 10.77(AMY), and 3.70 (LIP). "Start-up" rules were MR N2, MR N4, MR N2, MR N4, MR N4, 1:2.5 s N1, 1:3 s N1, and MR N4, and "monitor" QC rules were 1:2.5 s N1, 1:3 s N2, 1:2.5 s N1, MR N4, 1:3 s N2, 1:3 s N1, 1:3 s N1, MR N2 for 8 enzymes, respectively.

Conclusion: Multistage bracketed SQC is determined by sigma performance. Risk monitoring is significant during assaying to reduce patient risks and improve quality.

1. Introduction

An appropriate clinical decision cannot be dissociated from accurate and reliable test results, which sets high expectations from laboratories. To meet the clinical demand, laboratory staff are committed to daily quality control (QC) and unceasing improvement for providing the credibility of patient outcomes [1]. Thus far, considerable progress and changes have been made in QC models in medical laboratories.

Currently, QC strategies (rules and numbers) are usually established on the basis of sigma performance of tests [2], an index positively related to the detection performance and degree of stabilization of analytical process, to ensure an at least 90% error detection rate (Ped) when out of control and as a maximally low false rejection rate (Pfr) while in control. However, QC plan (QCP) do not involve the QC frequency, i.e., the time interval of QC events, whereby system errors of the analytical process can be detected in time. In the era of continuous production analyzers, more emphasis is placed on the risks of erroneous outcomes reported and their hazards to patients owing to the lack of timely detection of unmanageable conditions [3]. Therefore, the frequency of QC warrants focus by laboratory staff. In 2016, QC frequency was proposed to be optimized on the basis of MaxE (Nuf) in CLSI C24-Ed4 [4]; however, no specific methods and tools were mentioned. MaxE (Nuf), defined as the maximum expected increase in the number of unacceptable patient results reported before an unmanageable condition is detected, is a risk management indicator proposed by Parvin [5].

Recently, Yago [6] and Bayat [7] have postulated a specific tool, the nomogram, combined with Ped, a parameter of statistical quality control (SQC) with MaxE (Nuf), an index of risk management, to design QCP. Westgard [8] has developed a sigma-metric run size nomogram, which can visually recommend QC rules, numbers, and frequency, and introduced multistage bracketed SQC to monitor patient risks in real time. The MaxE (Nuf) is maintained below 1 and the risk of erroneous

Abbreviations: SQC, statistical quality control; ALT, alanine transaminase; AST, aspartate transaminase; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; LDH, lactate dehydrogenase; CK, creatine kinase; AMY, amylase; LIP, lipase; Pfr, false rejection rate; Ped, error detection rate; QCP, quality control plan; EFLM, European Federation of Clinical Chemistry and Laboratory Medicine; EQA, external quality assessment

* Corresponding author at: No. 37, Guoxue Lane, Wuhou District, Chengdu 610041, Sichuan, China.
E-mail address: huanghenjian@sina.com (H. Huang).

https://doi.org/10.1016/j.cca.2019.02.006
Received 16 September 2018; Received in revised form 6 February 2019; Accepted 6 February 2019
Available online 08 February 2019
0009-8981/

patient outcomes reportedly reduced to the lowest theoretical value. No data regarding the actual application of this method in clinical laboratories have been reported yet. In this study, we designed a multistage bracketed SQC for 8 enzymes based on a nomogram, plotted QC charts to monitor the stability of the detection methods, and performed quality improvement in our laboratory.

2. Materials and methods

2.1. Materials

Eight enzymes were included in our study: alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), creatine kinase (CK), amylase (AMY), and lipase (LIP). They were all assessed using the Roche Cobas 8000 c702 modular analyzer (Germany, Switzerland) with reagents for calibrations (Roche Diagnostics GmbH made in Germany). Liquid QC materials were derived from Bio-Rad Laboratories Inc. (made in USA), with the batch of low level 45,751 and high level 45,753.

2.2. Methods

The detection methods for the 8 enzymes were as follows: ALT, AST, and CK via the rate method, GGT via the rate method involving the gamma-glutamyl-3-carboxyl-4-notroaniline method, ALP via the rate method involving AMP buffer, LDH via the rate method involving lactic acid and pyruvic acid (L-P), AMY via the rate method involving maltoheptaoside method, and LIP via a colorimetric method.

2.2.1. Determination of sigma metrics

Sigma metrics were calculated using the formula [9]:

$$\text{Sigma} = (\text{TEa} - \text{Bias})/\text{CV} \quad (1)$$

TEa was obtained from the updated quality specifications of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) biological variation data published in 2017 [10]. TEa was calculated by the Frasers equation as TEa = Bias + 1.65CV [11].

CV was the coefficient of variation for the internal QC (IQC) data for consecutive 6 months, determined using the following formula:

$$CV_{average}\% = (CV_{high}\% + CV_{low}\%)/2 \quad (2)$$

wherein, CV_{high}% was the CV of batch number 45,753; CV_{low}% was the CV of batch number 45,751.

Bias was calculated by the average of 15 results of the external quality assessment (EQA) of Clinical laboratory center of Ministry of Health during 2017.

2.2.2. Practical application of multistage bracketed SQC

Multistage bracketed SQC were designed on the basis of the "sigma-metric SQC run size nomogram" presented in the Fig. 1 of reference [8]. First, the daily workload of tests and the expected reporting interval of patient samples were estimated. Thereafter, "start-up" and "monitor" QCPs were devised with the principle proposed by Westgard. The "start-up" QC before sample detection was designed on the basis of whether the sample size was greater than or equal to the specified daily workload. QC rules should have a high Ped and the system error can be detected in time. "Monitor" QC of the analytical process was designed on the basis of whether the sample size was greater than or equal to the desired reporting interval. QC rules should have a lower Pfr, ensuring satisfactory quality in the analytical process. The rules, numbers, and analytical batch length (frequency) for QCP were ultimately designed. Designed QCPs were actually applied and QC charts of 8 enzymes were developed in our laboratory.

3. Results

3.1. Sigma of 8 enzymes

The sigma metrics of ALT, AST, GGT, ALP, LDH, CK, AMY, and LIP were 5.26, 4.80, 5.25, 3.36, 4.71, 15.45, 10.77, and 3.70, respectively (Table 1). Based on the sigma value, tests were divisible into different categories, with different analytical performance and QCP set-up.

3.2. Development of multistage bracketed SQC

Multistage bracketed SQC of 8 enzymes were designed on the basis of the Fig. 1 of reference [8]. For CK (sigma > 6.0, workload: 500, QC interval: 100), 1:2.5 s N1 QC rule was recommended in the "start-up" and 1:3 s N1 in the "monitor" process. For AMY (sigma > 6.0, workload: 50, QC interval: 25), 1:3 s N1 was recommended in the whole QC schedule. For ALT and GGT (sigma~5.3, workload: 1000, QC interval: 200), MR N2 in the "start-up" stage and 1:2.5 s N1 in the "monitor" process were used. For AST and LDH with sigma value approaching 5, MR N4 in the "start-up" and 1:3 s N2 in the "monitor" process were suggested. For LIP (sigma < 4.0, workload: 50, QC interval: 25), MR N4 in the "start-up" and MR N2 in the "monitor" process were designed. For ALP(sigma < 4.0, workload: 1000, QC interval: 200), MR N4 was recommended in the whole QC schedule (Table 2).

3.3. Practical applications of this method in our laboratory

The designed QCP strategies for the 8 enzymes were actually applied in our laboratory for one week and Levey- Jennings QC charts were plotted (Fig. 1). According to the designed bracketed QC rules, ALP with the less sigma metrics and most QC materials had a poor analytical performance against 2:2 s and 4:1 s. Additionally, the 11_{th} QC event of ALT was markedly out of control against 1:3 s. While other tests were in control with a stable performance.

4. Discussion

Ped, a characteristic of QC, and sigma, a measure of performance, were both evaluation indicators of traditional SQC based on different context, which can help characterize the analysis performance of laboratory instruments [12]. A certain QC strategy has higher Ped for higher sigma. The higher the Ped or sigma, the better the performance, and vice versa. Clearly, the SQC focal point on the instruments themselves and a lack of interest in the exactitude of reported detection results lead to patient risks [13]. Laboratory staff cannot guarantee the complete absence of a system bias during analysis, thereby yielding inaccurate patient outcomes and resulting in incorrect clinical decisions, thus posing health hazards to patients. Parvin [14] et al. proposed a surrogate index, MaxE (Nuf), which is more closely associated with patient risks and is determined by TEa, QC rules, QC frequency, size, and type of unmanageable conditions. Among them, QC frequency is a relatively new determinant. The higher the frequency, the easier the detection of system errors and prevention of health hazards to patients. In our laboratory, a QC interval of 200 specimens of tests with a daily workload 1000, 100 specimens of workload 500, and 25 specimens of workload 50 were designed, consistent with the examples reported by Westgard. They postulated three parameters, an interval of 200 with a workload of 1000, an interval of 125 with a workload of 500, and an interval of 50 with 200 samples. Based on the results obtained for ALP (Fig. 1), an appropriate QC interval was vital for timely detection of errors for those tests with erratic analytical performance.

Besides run size, bracketed QC rules and numbers are an innovation. Multi-rules based on sigma performance are commonly used nowadays [15]. Nevertheless, the model is that the same QC rules for different tests and different periods are used, which is costly and inefficient. Multistage QC apparently has a personalized design based on different

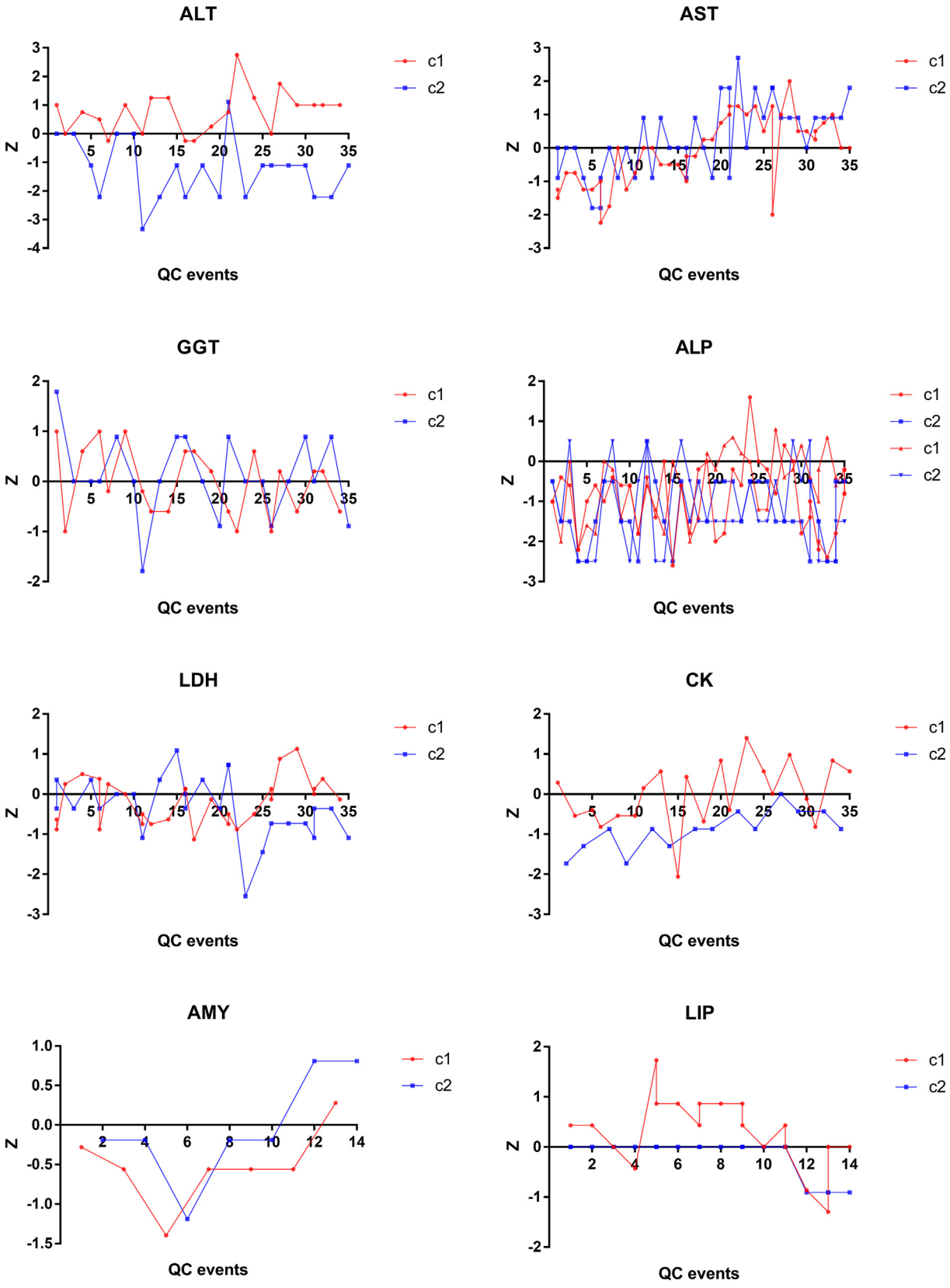

Fig. 1. Quality control (QC) charts of enzymes actually applied in our laboratory. Note: The X-axis represents QC events for one week and the Y-axis displays the *Z*-scores of QC materials. The red line denotes c1 with high-level QC material, and the blue line denoted c2 of low-level QC material.

test performances. Based on sigma performance, test items were divided into different categories with different QCP set-up. For tests with sigma < 4, MR QC rules were recommended throughout the QC schedule and greater QC frequencies were demanded. For ALP (workload: 1000, QC interval: 200), 20 QC materials were needed to monitor the analysis process for one day. And for LIP (workload: 50, QC

Table 1
Sigma metrics of the eight enzymes used in this study.

Tests	TEa (%)	$Bias_{average}$ (%)	IQC			Sigma
			CV_{high} (%)	CV_{low} (%)	$CV_{average}$ (%)	
ALT	14.4	1.75	1.23	3.58	2.41	5.26
AST	13.4	2.76	1.36	3.07	2.22	4.80
GGT	15.7	3.38	1.62	3.07	2.35	5.25
ALP	10.7	1.86	1.97	3.29	2.63	3.36
LDH	7.7	0.85	0.99	1.92	1.46	4.71
CK	20.4	1.32	1.00	1.47	1.24	15.45
AMY	13.4	1.18	0.90	1.37	1.14	10.77
LIP	12.6	2.18	1.70	3.93	2.82	3.70

IQC: internal quality control; ALT: alanine transaminase; AST: aspartate transaminase; GGT: gamma-glutamyl transferase; ALP: alkaline phosphatase; LDH: lactate dehydrogenase; CK: creatine kinase; AMY: amylase; LIP: lipase.

Table 2
Multistage bracketed statistical quality control analysis of eight enzymes.

Tests	Daily workload	Expected QC interval	Start-up QCP	Monitor QCP	Numbers of QC events	Numbers of QC materials
ALT	1000	200	MR* N2	1:2.5 s N1	5	6
AST	1000	200	MR N4	1:3 s N2	5	12
GGT	1000	200	MR N2	1:2.5 s N1	5	6
ALP	1000	200	MR N4	MR N4	5	20
LDH	500	100	MR N4	1:3 s N2	5	12
CK	500	100	1:2.5 s N1	1:3 s N1	5	5
AMY	50	25	1:3 s N1	1:3 s N1	2	2
LIP	50	25	MR N4	MR N2	2	6

Note:*MR means multi-rules, 1:3 s/2:2 s/R:4 s/4:1 s. ALT: alanine transaminase; AST: aspartate transaminase; GGT: gamma-glutamyl transferase; ALP: alkaline phosphatase; LDH: lactate dehydrogenase; CK: creatine kinase; AMY: amylase; LIP: lipase; QC: quality control; QCP: QC plan.

interval: 25) with a sigma of 3.70, MR N4 in the "start-up" and MR N2 in the "monitor" process were designed. Such QC strategies were very cost prohibitive, and it would be better to replace the ALP and LIP method with a better one. For sigma closed values of ~5, QC rules were selected on the basis of run size. For tests with sigma > 6, 1:2.5 s N1 were suggested for run sizes > 1000 down to 400, and 1:3 s N1 for smaller run sizes, and fewer monitoring points can be designed. Beyond doubt, this will save costs and reduce health risks among patients. For CK with a sigma of 15.45, only 5 QC materials were needed for one day. Moreover, AMY (workload: 50, QC interval: 25) with a 6 sigma were controlled with a start-up and monitor QC of 1:3 s N1. Since with a 6 sigma method 1:3 s N1 can be applied with a run size of 350, now that it was applied with a run size of 14 times < 350 (350/25), Max E(Nuf) was reduced from 1 to 0.07 (1/14), which were in much more patient safety for AMY.

QC charts plotted in our laboratory showed that sigma metrics were of importance in QCP. ALP with sigma < 4 was unmanageable and markedly against QC rules, indicating a poor performance and the need for improvements. Therefore, the accuracy of sigma evaluated is also significant. Hence, the TEa of the formula was from the updated quality specification of the EFLM biological variation data published in 2017, which was more accurate and effective [16,17]. Of note, methods with discernably low performance should be rejected and improvements are indispensable instead of focusing on designing QCP to control the stability of the analytical process because this can result in a marked system bias with poor performance, as in the case of ALP (Fig. 1), and irrespective of QC frequency, will also appear unmanageable, accompanied by a complete loss of monitoring stability of the analytical process through QC events. Therefore, more attention should be paid to how to improve the performance of the ALP method. By analyzing the causes of unmanageable conditions, certain improvements were made. Such as, reagent preservation and operational capabilities of the clinical staff received greater attention and specialized training.

The QCP designed in the present study still has some shortcomings. First, our design essentially reduced health risks among patients through an increase in QC frequency, however, this was time-consuming upon detection of unmanageable conditions. Corrective actions must be undertaken to re-assess patient samples and the outcomes could not be reported until the QC results are satisfactory. Secondly, the MaxE (Nuf) is theoretically maintained below 1 by the sigma-metric SQC run size nomogram; moreover, it is cumbersome to calculate the MaxE (Nuf) in the clinical setting and specific risks cannot be known. Parvin has argued that the use of the nomogram is limited without the assessment of injury to patients. They proposed a new calculation index, Risk Management Index (RMI), which can more comprehensively assess health risks among patients [18]. Finally, economic burdens associated with a high QC frequency and increased labor volume due to personalized quality control strategies might not be deemed acceptable by the laboratory staff. Hence, the clinical applications are yet farfetched. Furthermore, more risk management tools should be investigated to improve quality, reduce health risks among patients, and enhance the skills of clinicians.

5. Conclusion

In conclusion, for reducing health risks among patients, the application of multistage bracketed SQC, determined primarily on the basis of sigma performance, is essential and the best method is to improve performance in tests with a low sigma performance.

Acknowledgments

This research was funded by Department of Science and Technology, Sichuan Province with project number 2016ZR0071 and 2017ZR0063.

Conflict of interest

Declarations of interest: none.

References

[1] J.O. Westgard, S.A. Westgard, Quality control review: implementing a scientifically based quality control system, Ann. Clin. Biochem. 53 (Pt 1) (2016) 32–50.
[2] J.O. Westgard, S.A. Westgard, Six sigma quality management system and design of risk-based statistical quality control, Clin. Lab. Med. 37 (1) (2017) 85–96.
[3] C. Parvin, Planning statistical quality control to minimize patient risk: It's about time, Clin. Chem. 64 (2) (2018) 249–250.
[4] C. C24-Ed4, Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions Clinical and Laboratory Standards Institute, 4th Ed, (2016).
[5] C. Parvin, Assessing the impact of the frequency of quality control testing on the quality of reported patient results, Clin. Chem. 54 (12) (2008) 2049–2054.
[6] M. Yago, S. Alcover, Selecting statistical procedures for quality control planning based on risk management, Clin. Chem. 62 (7) (2016) 959–965.
[7] H. Bayat, Selecting multi-rule quality control procedures based on patient risk, Clin. Chem. Lab. Med. 55 (11) (2017) 1702–1708.
[8] J. Westgard, H. Bayat, S. Westgard, Planning risk-based SQC schedules for bracketed operation of continuous production analyzers, Clin. Chem. 64 (2) (2018) 289–296.
[9] J.O. Westgard, S.A. Westgard, The quality of laboratory testing today: an assessment of sigma metrics for analytic quality using performance data from proficiency testing surveys and the CLIA criteria for acceptable performance, Am. J. Clin. Pathol. 125 (3) (2006) 343–354.
[10] A. Carobene, T. Røraas, U. Sølvik, M. Sylte, S. Sandberg, E. Guerra, I. Marino, N. Jonker, G. Barla, W. Bartlett, P. Fernandez-Calle, J. Díaz-Garzón, F. Tosato, M. Plebani, A. Coşkun, M. Serteser, I. Unsal, F. Ceriotti, Biological variation estimates obtained from 91 healthy study participants for 9 enzymes in serum, Clin. Chem. 63 (6) (2017) 1141–1150.
[11] C.G. Fraser, P.H. Petersen, Quality goals in external quality assessment are best based on biology, Scandinavian journal of clinical and laboratory investigation, Supplementum 212 (1993) 8–9.
[12] J. Jairaman, Z. Sakiman, L. Li, Sunway medical laboratory quality control plans

based on six sigma, risk management and uncertainty, Clin. Lab. Med. 37 (1) (2017) 163–176.
[13] J.H. Nichols, Laboratory quality control based on risk management, Ann. Saudi Med. 31 (3) (2011) 223–228.
[14] J. Yundt-Pacheco, C.A. Parvin, Validating the performance of QC procedures, Clin. Lab. Med. 33 (1) (2013) 75–88.
[15] J. Westgard, A total quality-control plan with right-sized statistical quality-control, Clin. Lab. Med. 37 (1) (2017) 137–150.
[16] W. Bartlett, F. Braga, A. Carobene, A. Coşkun, R. Prusa, P. Fernandez-Calle, T. Røraas, N. Jonker, S. Sandberg, A checklist for critical appraisal of studies of biological variation, Clin. Chem. Lab. Med. 53 (6) (2015) 879–885.
[17] C. Perich, J. Minchinela, C. Ricós, P. Fernández-Calle, V. Alvarez, M. Doménech, M. Simón, C. Biosca, B. Boned, J. García-Lario, F. Cava, P. Fernández-Fernández, C. Fraser, Biological variation database: structure and criteria used for generation and update, Clin. Chem. Lab. Med. 53 (2) (2015) 299–305.
[18] C. Parvin, N. Baumann, Assessing quality control strategies for HbA1c measurements from a patient risk perspective, J. Diabetes Sci. Technol. 12 (4) (2018) 786–791.