The application of sigma metrics in the laboratory to assess quality control processes

A research report submitted to the Faculty of Health Sciences, University of Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Medicine

Johannesburg

Submitted by Dr Marli van Heerden

December 2020

Declaration

I ……Marli van Heerden……. (name) declare that:

a) This Research Report is my own, unaided work.
b) It is being submitted for the Degree of Master of Medicine in the branch of Chemical Pathology at University of the Witwatersrand, Johannesburg.
c) It has not been submitted before for any degree or examination at any other university.
d) This research report does not contain other people's data, pictures, graphs or information, unless specifically acknowledged, and sources are listed in the referenced section.
e) This research report is currently under review by the African Journal of Laboratory Medicine.

…….

(Signature of candidate)

---9th--day of -----December----------2020

Acknowledgements

- I would like to express my gratitude to my supervisors Prof Jaya George, Dr Siyabonga Khoza and Dr HE van Deventer for their guidance, support, and encouragement.
- We thank Mrs. Lindi Ntlebi for her assistance in retrieving IQC data for the study period.

Table of contents

List of Tables

List of Figures

Supplementary material

Abstract

Introduction: Laboratories monitor and evaluate analytical performance using quality control processes in terms of precision and bias. Sigma metrics can provide an objective manner to assess and compare quality using an additional parameter of total allowable error. The aim of this study was to calculate the sigma metric of analytes when using different sources of total allowable error and suggest Westgard multirules accordingly.

Materials and methods: A retrospective analysis was performed on 19 general chemistry analytes at a large academic hospital in Johannesburg. Sigma metrics were calculated on two identical analyzers using internal quality control data and the total allowable error from the Ricos biological variation database as well as three alternative sources (Royal College Pathologists of Australia (RCPA), Clinical Laboratory Improvements Amendment (CLIA), and European Federation Laboratory Medicine (EFLM)).

Results: Both analyzers had similar sigma performance. The sigma varied based on the source of total allowable error used with the CLIA guidelines resulting in the best sigma metrics (46% and 53% of analytes achieving acceptable sigma on each analyzer, respectively) and the RCPA guidelines being the most stringent (21% and 23% respectively). Sodium and chloride performed poorly across all guidelines (sigma <3). Month-to-month variation was also noted that may result in acceptable sigma despite poor performance during certain months.

Conclusion: The sigma varies greatly depending on the source of the total allowable error. In a busy laboratory, it could prove to be a valuable tool to save time and decrease costs. There is a clear need for standardization of sigma metrics protocols.

Introduction

Medical laboratories strive to produce accurate reproducible results as physicians rely on these for diagnosis, monitoring, and prognostication of patients (1). To achieve this, medical laboratories monitor and evaluate analytical processes using a number of different quality control (QC) processes, which aim to produce results with no errors. In practice, there are no processes with zero defects (2).

In the analytical process, the examination procedure performances are typically evaluated in terms of precision and accuracy (bias). This is determined by the use of quality control procedures and is performed at intervals determined by laboratory policy. A high standard deviation indicates poor precision, instability, and high random error (3). Most South African laboratories use Levey Jennings control charts and Westgard QC rules to determine whether or not a QC run is acceptable based on an algorithm with specified limits. This approach might not be ideal as one set of rules cannot be applied to all tests due to varying precision and goals (4).

The number of QC levels and the frequency of running QC varies greatly between laboratories (5). The National Accreditation Board for Testing and Calibration

Laboratories (NABL) guidelines report that two level controls should be run at a peak hour and subsequently one level every 8 hours for laboratories that run continuously (6). Several rules may be applied to determine if the QC values are acceptable or not. Most laboratories use 1_2S as a warning rule. This implies that a single control measurement exceeding two standard deviations from the mean (in any direction) may indicate a problem (5). However, when this rule is used as a control rule, it can cause a false rejection rate of up to 14% (5). Internal quality control policies regard the 1_{3s}, R_{4s} and 2_{2s} rule as criteria for rejection, while 10 consecutive observations on one side of the mean (10x rule) require further investigation. Combinations of rules (multirules) are sometimes employed to reduce the rate of false rejections and to save time and effort by incorporating rules that are sensitive to both random and systematic errors.

Six sigma elaborates on this further by individualizing control rules based on analytical performance of the test (7). The lower the sigma metric, the greater the number of control rules required to monitor the analyte effectively. Analytes achieving sigma $\geq$6 require a single 1_3S control rule with 2 control measurements performed once daily. In contrast, when the sigma is between 3 and 4, the maximum number of control rules are required and the number of control levels or the frequency of runs needs to be increased. When sigma <3 there is no acceptable QC strategy available.

The sigma scale provides an objective manner to assess and compare the quality that is required by incorporating both the imprecision and bias observed in a laboratory's performance (8). Sigma metric is based on 3 parameters: total allowable

error (TEa), bias, and imprecision. The TEa from various sources is associated with significantly different sigma metrics for the same assay (9,10).

Six Sigma can be used to decide on the best Westgard rule by judging the performance of a process against a reference method and to assess quality or identify processes needing improvement. As demonstrated by Litten (7), implementation of a Six Sigma-designed QC program can result in fewer controls per run, fewer false rejections, simpler Westgard rules, and a 45% saving on reagents and supplies in the laboratory.

Our laboratory runs two lines of analyzers in parallel and currently does not use Sigma metrics to manage QC. The aim of this study was to calculate the sigma metrics of selected analytes, to determine how sigma changed with different sources of TEa in our setting and to see if sigma differed across lines and suggest Westgard rules accordingly.

Materials and method

The study was conducted in a National Health Laboratory Service (NHLS) laboratory at Charlotte Maxeke Johannesburg Academic Tertiary Hospital, Johannesburg, which is ISO 15189 accredited. There are two identical Cobas® 8000 (Roche Diagnostics, Mannheim, Germany) systems (referred to as Line 1 and Line 2) running in parallel; with some tests run on both lines.

A retrospective analysis of 19 analytes was performed using internal quality control data obtained from the Roche Cobas® 8000 chemistry analyser IT middleware system over a period of 12 months (January-December 2017). We used Roche QC materials: one in the normal range and one in the abnormal range (Precicontrol ClinChem Multi 1 and 2, Precicontrol Tumour marker, and Precicontrol Universal). A single lot number of QC material was used for the 12 months except for PSA and TSH, for which different lot numbers were used. Ethics was obtained from the Human Research Ethics Committee of the University of the Witwatersrand (number W-CBP-180216-01).

Tests included in our study: alanine aminotransferase (ALT), alkaline phosphatase (ALP), aspartate aminotransferase (AST), calcium (Ca), chloride (Cl), cholesterol (CHOL), creatine kinase (CK), creatinine (CREA), direct bilirubin (DBIL), glucose (Glc), high density lipoprotein-cholesterol (HDL), potassium (K), sodium (Na), thyroid-stimulating hormone (TSH), total bilirubin (TBIL), total protein (TP), prostate-specific antigen (PSA), urea and uric acid (UA).

The reagent, QC, and calibrator materials were all provided by Roche Diagnostics.

Statistical analysis

Using internal QC means and standard deviations, we calculated the bias and coefficient of variation (CV). The bias was determined by subtracting the package insert target value from the observed quality control mean. The CV was determined by the following calculation: CV=100 x (SD/mean)

For our study, the biological variation database from Ricos *et al*. was used to determine the desirable test-specific quality requirements. This database was last updated by the Spanish Society of Laboratory Medicine in 2014 (11). This was compared to the total allowable error (TEa) guidelines from Clinical Laboratory Improvement Amendments (CLIA) (12), the Royal College of Pathologists of Australia (RCPA) (13), and the European Federation of Laboratory Medicine (EFLM) (14).

TEa values given as a percentage were converted to units with the following calculation:

TEa (units) = (TEa%/100)*Target value.

Sigma metrics were calculated for each analyte, on two levels, as follows:

Sigma= [(TEa – (Bias)/SD].

Thereafter the average annual sigma metric was calculated by combining the values obtained for each month and dividing it by the number of months.

The Quality Goal index (QGI) indicates the possible source of error and represents the relative degree to which bias and precision meet their quality goals (15).

The QGI was determined as follows:

QGI= Bias/1.5CV.

A QGI score of <0.8 indicates imprecision, QGI score 0.8-1.2 indicates both imprecision and inaccuracy, and a score of >1.2 indicates inaccuracy (15).

Metanalysis data for within (CV_i) and between (CV_g) subject variation were obtained from the European Federation of Clinical Chemistry Laboratory Medicine Biological Variation database (14).

The desirable specifications for imprecision, bias, and total allowable error were calculated as follows:

$CV\% = 0.5(CV_i)$,

$Bias\% = 0.25(CV_i^2 + CV_g^2)^{0.5}$

$TEa\% = (1.65 \times CV\ \%) + Bias\%$

The allowable limits of performance for total and direct bilirubin, calcium, and uric acid were obtained from the Westgard website based on EFLM data (16).

The Multirules are based on the Westgard Sigma Rules™ (17). The data was captured, and statistical analysis was performed using Windows 10, Microsoft Excel.

Results

Figure 1 demonstrates the average annual sigma metric obtained for each analyte when combining the performance of the two QC levels. Eight analytes (ALT, AST, TBIL, CK, Urea, DBIL, PSA and TSH) achieved sigma > 3 on both lines, while one analyte (UA) had a sigma of >3 on line 1 and < 3 on line 2 (Figure 1a and 1b).

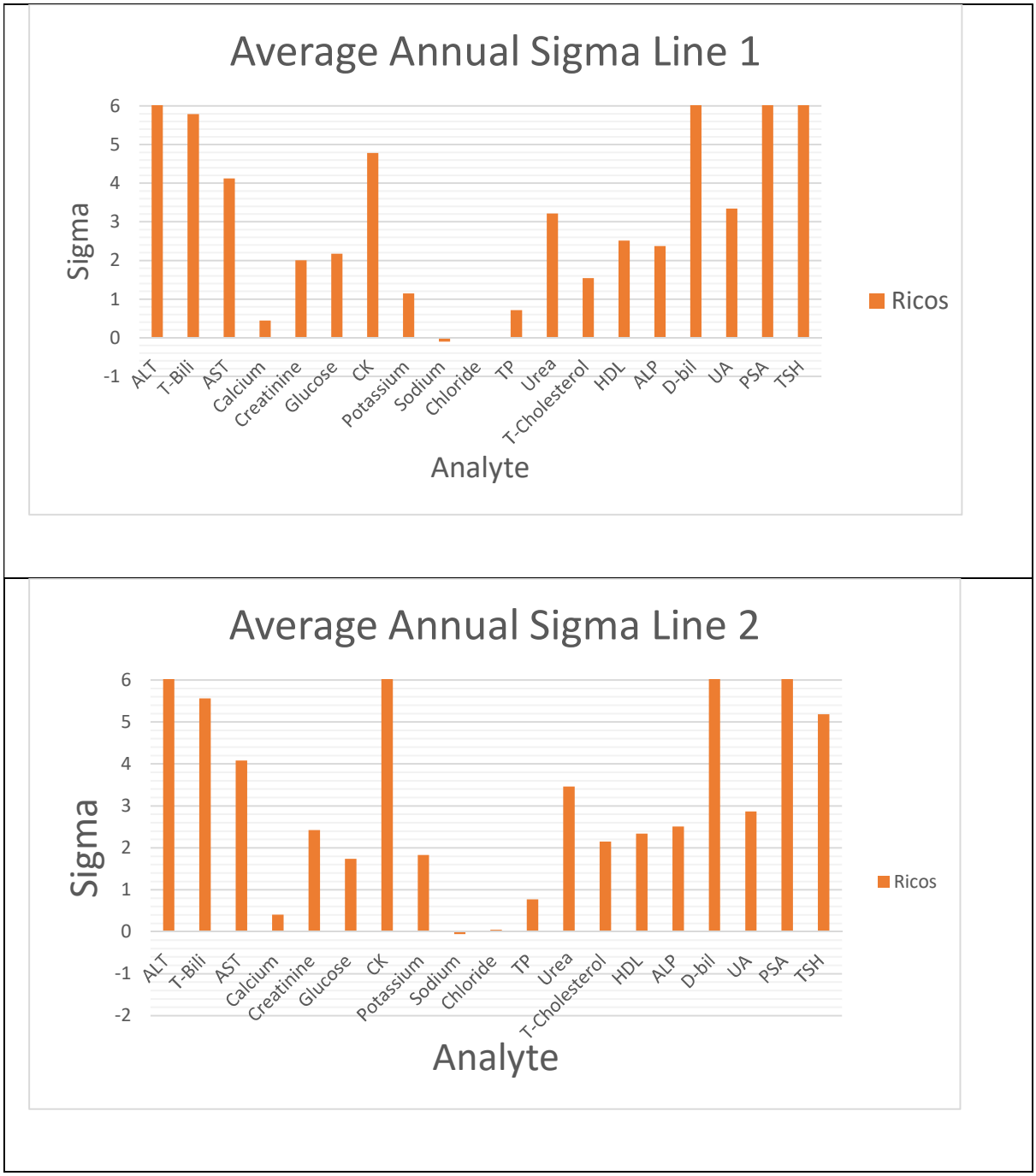


Figure 1a and b: Average annual sigma for Line 1 and Line 2

Table 1 demonstrates the sigma values for each analyte grouped according to performance for both lines based on different sources of TEa and concentration of control. The combined sigma across levels per TEa goal is demonstrated in Table 2, together with the QGI.

The suggested Westgard multirules are based on the sigma metric obtained (Table 3).

Table 1: Analyte sigma metric classification by different TEa guidelines.

	Ricos		CLIA		RCPA		EFLM	
	Line 1	Line 2	Line 1	Line 2	Line 1	Line 2	Line 1	Line 2
Sigma <3	ALP Ca CHOL Cl CREA Glc 1 HDL 2 K Na TP	ALP 1 Ca CHOL Cl CREA Glc1 HDL K Na TP	Ca 2 CHOL Cl Glc 1 K 2 Na TP 1 TSH Urea	Ca 2 CHOL 1 Cl Glc 1 Na TP 1 Urea	ALT 1 ALP 2 Ca CHOL Cl CK CREA Glc 1 HDL 2 K Na PSA TBIL 2 TP Urea UA 2	ALT AST 1 Ca Cl CHOL2 CREA HDL 1 Glc1 Na K PSA TBIL 2 TP Urea 1	ALP Ca CHOL Cl CREA Glc HDL 2 Na K TP	AST 1 ALP Ca Cl CHOL CREA Glc 1 HDL Na K TP UA 1
Sigma 3-6	ALT 1 AST CK HDL 1 Glc 2 TBIL Urea UA	ALT 1 ALP 2 AST Glc 2 TBIL TSA Urea UA PSA 2	ALT AST Ca 1 CREA 2 CK K ISE 1 TBIL 2 TP2 UA	ALT AST Ca 1 CHOL 2 CREA 2 Glc 2 K TBILI 2 TP L2 TSH UA	ALP 1 ALT 2 AST DBIL HDL 1 TBIL 1 UA 1	ALP AST 2 CK CHOL1 DBIL Glc 2 HDL 2 TBIL 1 TSH Urea 2	ALT AST CK HDL1 PSA TBIL UA Urea	ALT AST 2 CK 2 Glc 2 PSA TBIL TSH Urea UA 2
Sigma >6	ALT l2 DBIL PSA TSH	ALT 2 CK DBIL PSA 1	ALP CREAT 1 HDL TBIL level 1 Glc 2	ALP CREA 1 CK HDL TBIL 1	TSH Glc 2		DBIL TSH	CK 1 DBIL

ALT-alanine aminotransferase, ALP-alkaline phosphatase, AST-aspartate aminotransferase, Ca-calcium, Cl-chloride, CHOL-cholesterol, CK-creatine kinase, CREA-creatinine, DBIL-direct bilirubin, Glc-glucose, HDL-high density lipoprotein, K-potassium, Na-sodium, TSH-thyroid stimulating hormone, TBIL-total bilirubin, TP-total protein, PSA-prostate specific antigen, UA-uric acid.

Table 2a: Line 1 average sigma across QC levels by different TEa guidelines					
Analyte	LINE 1 sigma metrics				QGI
	Ricos	EFLM	CLIA	RCPA	
ALT	7	3	5	2	0.36
TBIL	6	6	6	3	0.17
AST	4	3	5	3	0.15
Ca	0	0	3	1	0.14
CREA	2	2	6	2	0.12
Glc	2	3	4	3	0.18
CK	5	3	5	2	015
Na	1	2	2	1	0.17
K	0	0	1	1	0.17
Cl	0	0	1	1	0.24
TP	1	1	3	1	0.15
Urea	3	4	2	2	0.17
CHOL	2	2	2	2	0.20
HDL	3	2	7	3	0.29
ALP	2	2	7	3	0.22
DBIL	10	8	x	5	0.18
UA	3	3	5	3	0.18
PSA	11	5	x	1	0.56
TSH	8	7	2	6	0.28

Table 2b: Line 2 average sigma across QC levels by different TEa guidelines					
Analyte	LINE 2 sigma metrics				QGI
	Ricos	EFLM	CLIA	RCPA	
ALT	7	3	4	2	0.44
TBIL	6	5	6	3	0.21
AST	4	3	5	3	0.11
Ca	0	0	2	1	0.15
CREA	2	2	7	2	0.20
Glc	2	2	4	3	0.19
CK	13	6	13	5	0.26
K	2	2	3	2	0.13
Na	0	0	1	1	0.19
Cl	0	0	1	1	0.16
TP	1	1	3	1	0.15
Urea	3	4	2	2	0.29
CHOL	2	2	2	2	0.24
HDL	2	2	8	2	0.50
ALP	3	2	7	3	0.20
DBIL	13	9	X	5	0.23
UA	3	2	4	2	0.18
PSA	7	3	X	1	0.44
TSH	5	5	3	4	0.13

Table 3: Suggested Westgard Multirules based on sigma classification	
Sigma	Suggested Westgard Multirules
<3	No acceptable QC strategy available
3-6	Sigma 3-4: $1_3S/2_2S/R_4S/4_1S/8x$ 2 levels 4 times/day OR 4 levels 2/day Sigma 4-5: $1_3S/2_2S/R_4S/4_1S$ 4 levels once/day OR 2 levels twice/day Sigma 5-6: $1_3S/2_2S/R_4S$ 2 levels once/day
>6	$1._3S$ with two levels of QC per day

This data is also presented in a normalized method decision chart available in the supplementary material (Figure 1). Performance was very similar across both lines for each guideline. For example, using the Ricos TEa goals, the same ten analytes had sigma < 3 across both lines and the same 4 analytes had a sigma of > 6 on both lines. Only AST achieved sigma > 3 with all guidelines and at both analyte concentrations, but this was seen only on line 1. No single analyte achieved sigma≥6 across all guidelines and for two control levels.

When comparing the sigma performance based on the different sources of TEa, the CLIA guidelines resulted in the best sigma metrics with 46% of analytes and 53% of analytes achieving sigma ≥3 on Line 1 and 2, respectively. Sixteen percent of analytes

achieved sigma>6 on Line 1 and 20% on Line 2 using the Ricos TEa goals. Using the Ricos TEa golas, 43% had sigma ≥ 3 on line 1, and 36% on line 2. This was followed by the EFLM with 39% and 32% of analytes having sigma > 3 for lines 1 and 2, respectively. The worst sigma metrics were obtained using RCPA guidelines, 80% with sigma < 3 for line 1, and 77% for line 2 (Figures 2a and 2b). No analyte on line 2 achieved sigma ≥ 6 using RCPA guidelines.

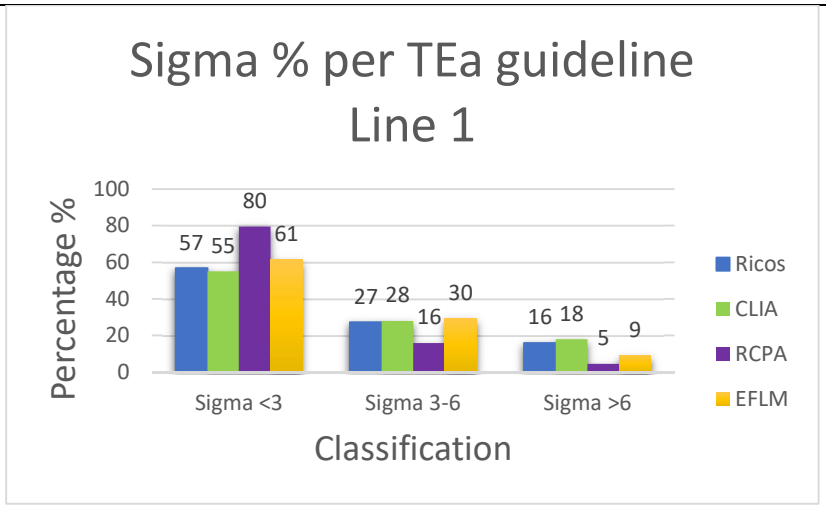


Figure 2a

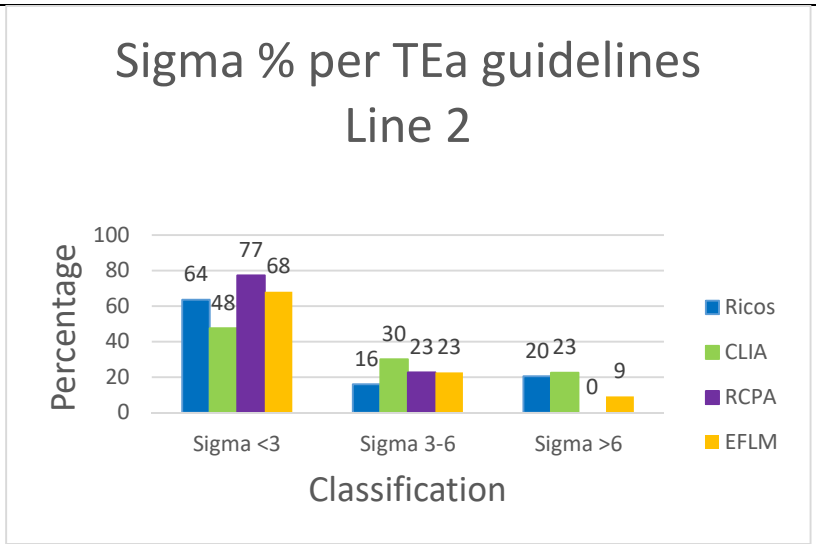
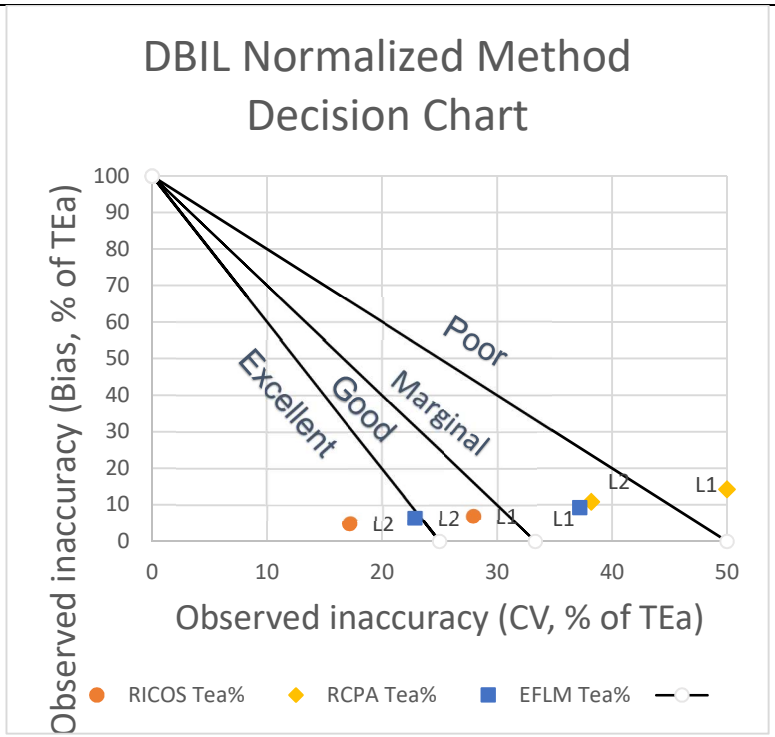


Figure 2b

Figure 2a and b: Line 1 and 2 Sigma Performance when different TEa guidelines are utilized

When looking at the method decision charts, specific analytes were chosen to demonstrate the general findings of this study. More analytes were classified as excellent performers when using the Ricos TEa goals compared to that of the RCPA. (Figures 3a and 3b). For TSH, we noted a similar pattern with the EFLM guidelines also resulting in good performance (Figure 3c). There was also variation in performance between QC levels with level 2 generally performing better (Figures 3 a-c). As demonstrated in Figure 4a, the Ricos and CLIA TEA goals appear to be the most lenient, resulting in better performance for CK on Line 2. The EFLM guidelines resulted in good performance, while the RCPA guidelines are stricter, resulting in poor performance. HDL generally performed poorly, except when the CLIA guidelines were applied (Figure 4b).


DBIL Normalized Method Decision Chart
Observed inaccuracy (Bias, % of TEa)
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
Observed inaccuracy (CV, % of TEa)
Poor
Marginal
Good
Excellent
L2
L1
L2
L2
L1
L1
L2
L1
RICOS Tea%
RCPA Tea%
EFLM Tea%

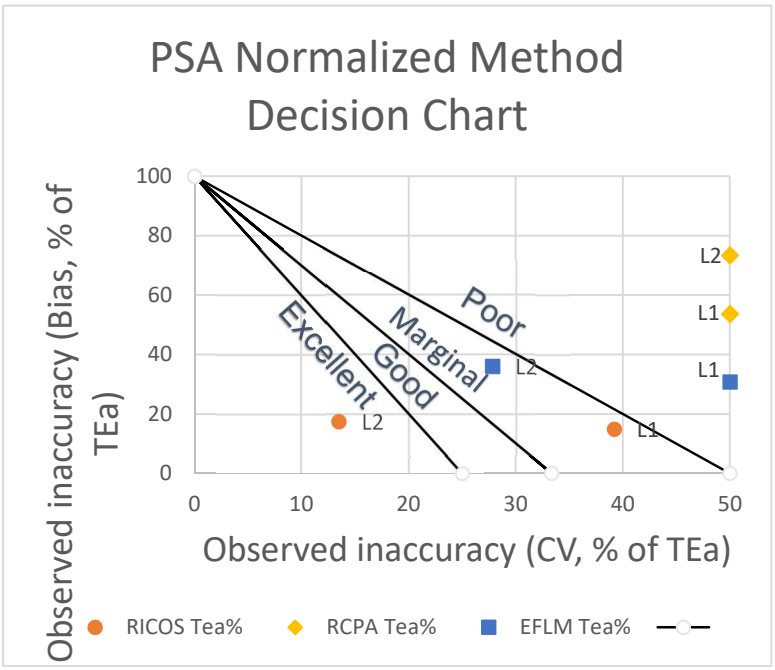

PSA Normalized Method Decision Chart
Observed inaccuracy (Bias, % of TEa)
100
80
60
40
20
0
0
10
20
30
40
50
Observed inaccuracy (CV, % of TEa)
Poor
Marginal
Good
Excellent
L2
L1
L2
L2
L1
L1
RICOS Tea%
RCPA Tea%
EFLM Tea%

Figure 3b

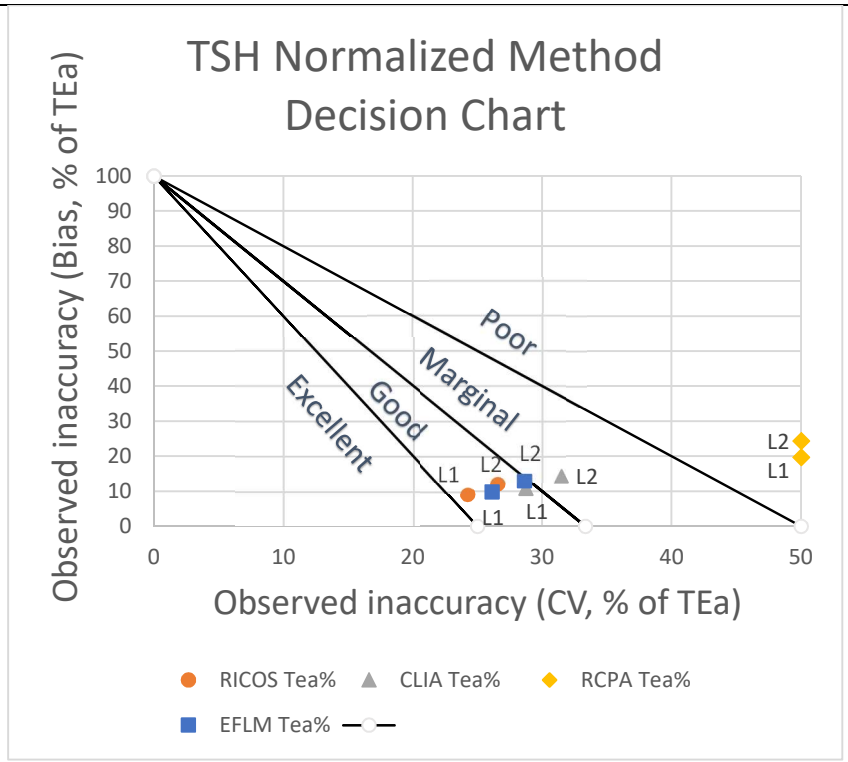


Figure 3c

L1-Level 1, L2- Level 2

Figure 3a, b and c: Normalized Method Decision Chart for DBIL, PSA and TSH on Line 1.

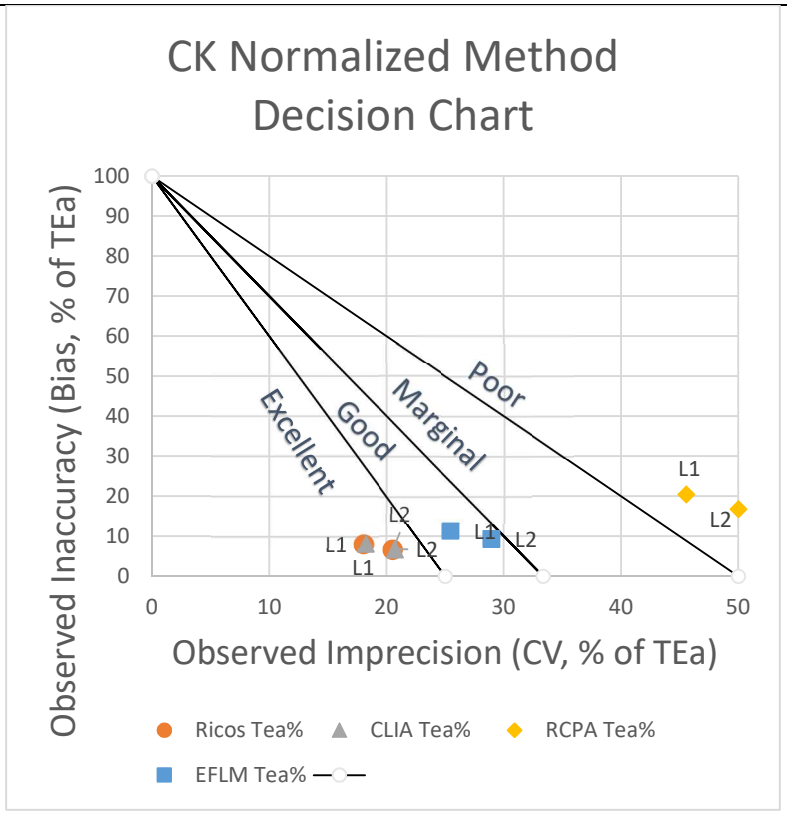


Figure 4a

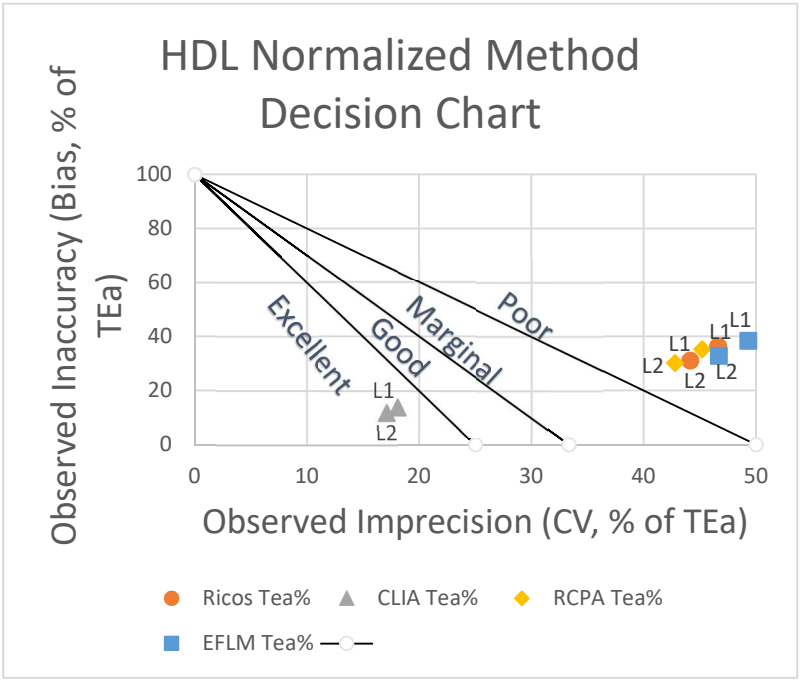


Figure 4b

L1- Level 1, L2- Level 2

Figure 4a and b: Normalized Method Decision Chart for CK and HDL on Line 2.

A number of analytes had sigma <3 for both QC levels across both lines when using the Ricos TEa performance specifications. These included Na, Cl, Glc, and CHOL. AST and urea achieved sigma metrics between 3 and 6 on both lines respectively. Three analytes achieved sigma >6 on both control levels, namely DBIL, PSA, and TSH on line1, as well as CK, DBIL, and PSA on line 2 using the Ricos TEa goal. (See Table 1).

We observed large variations in performance from month to month (Figure 5a and b). Many analytes achieved an annual average acceptable sigma despite poor or marginal performance during certain months. Examples include TSH (Supplementary figures 7a and 7b), AST, TBIL and ALT (Supplementary figures 2 and 5). On the other hand, CREA (Supplementary figures 2 and 5), CK level 1 (Supplementary figures 2a and 5a) and CHOL level 2 (Supplementary figures 4b and 7b) failed to achieve acceptable sigma despite sigma >3 during certain months. ALT, TBIL (Supplementary figure 2), and AST level 1 (Supplementary figure 2a) on Line 1 (module 702) displayed similar patterns with improvements during the months of April, May, July, September, and November.

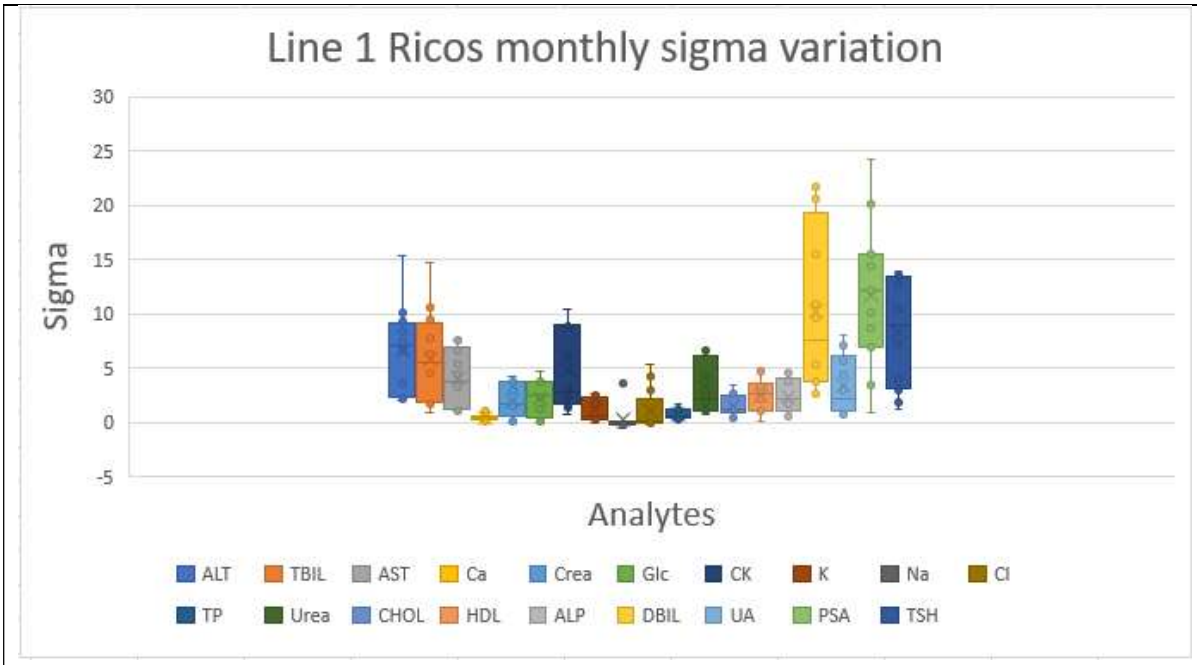


Figure 5a

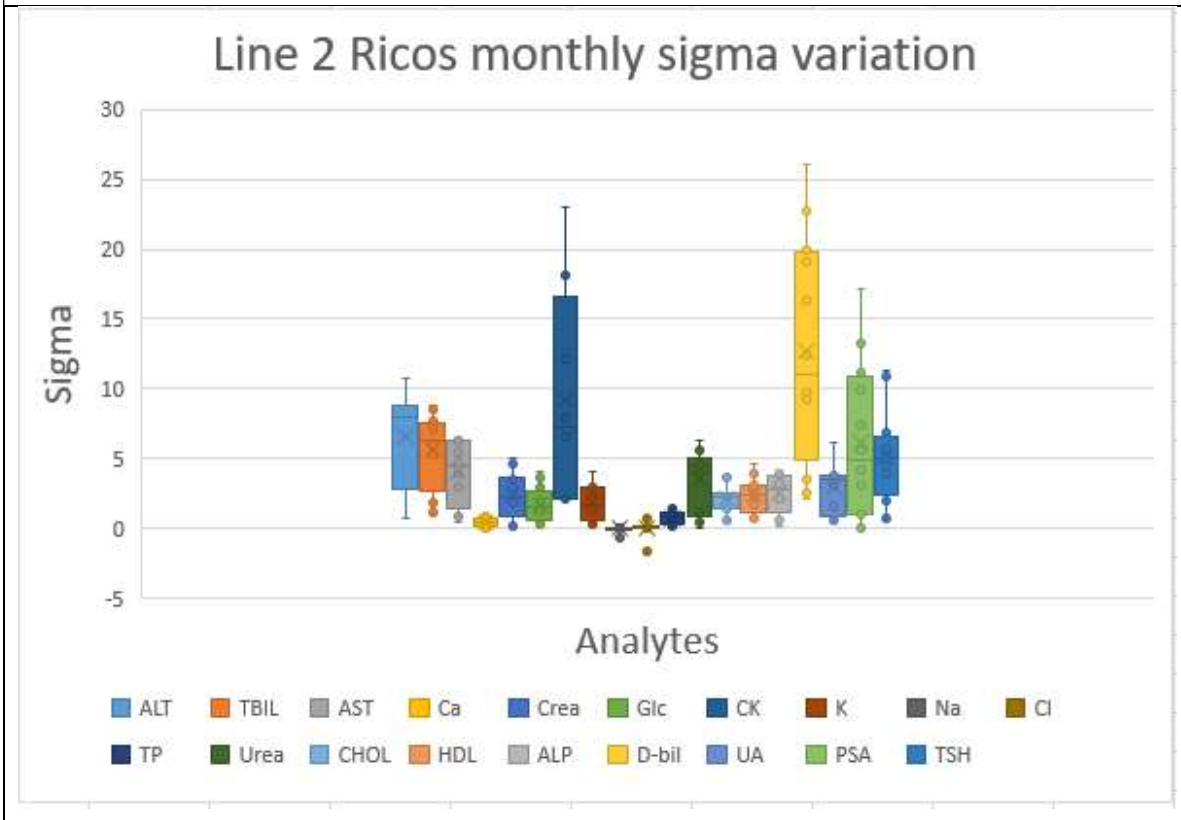


Figure 5b

Figure 5a and 5b: Box and whisker plots demonstrating the variability in analyte performance during the 12-month period.

Discussion

In our study, we analysed quality control (QC) data for 19 chemistry analytes over a 12-month period and calculated sigma for each analyte to objectively evaluate the performance and to determine if we can use the sigma metrics to identify appropriate control rules. We, like others, have shown that six sigma values vary depending on the source of TEa used. We also showed variation in sigma with concentration of analyte and variation from month to month.

There are no universally accepted TEa goals and published data show that sigma metrics vary with TEa goals used (8,10,19). There are a number of ways to address this. For example, Sharkawy *et al*. (19) created a harmonization protocol for sigma calculation for comparison of sigma metrics across laboratories by using similar TEa goals. We noted fewer differences across both lines using the same TEa goals compared to sigma variation with different TEa goals. Another proposed approach is to assess the effect of TEa on patient outcomes. Researchers in China (20) assessed the performance of "severity of harm" from TEa being exceeded in 36 analytes. A risk priority number (PRN) was assigned by multiplying the sigma metric by the score of the intended use. The authors suggested that TEa should be defined by the highest possible hierarchical model and recommend that tests that show negligible risk to patient should be allowed to reach lower sigma metrics (20).

In our study, the CLIA TEa guidelines resulted in the best sigma metrics. This may change in the future as CLIA proposed new limits that appear to be less lenient (21).

The RCPA guidelines appeared to be the most stringent, which is in keeping with the results of a study by Liu et al. (22). Recently, EFLM has established the Working Group on Biological Variation (WG-BV) and the Task and Finish Group for the Biological Variation Database (TFG-BVD) to assess the quality of existing BV data and to compile global estimates in an attempt to harmonize analytical performance specifications world-wide (23). Another option would be to use an algorithm as proposed by Varela and Pacheco (24) which standardizes the selection procedure of the most appropriate TEa for the test analytical performance. In addition to TEa, analyte concentration is crucial when determining sigma metrics (8). In a study that investigated the performance of verified versus non-verified reagents, Cao et al. (25) observed that sigma varied with analyte concentration and suggested that different rules should be used for different analyte concentrations. The observed changes in sigma value with analyte concentration may be due to changes in precision and/or bias with analyte concentration (8).

In contrast to our study, Gulbahar et al (32) found unacceptable sigma for urea and sigma >6 for potassium, AST, uric acid and total protein on the same platform. When these authors compared the Cobas platform to that of other analytical systems, results varied except sodium performing poorly throughout and ALP achieving six sigma on all platforms(32). Possible causes for these differences include the facility where studies are conducted, different environments and the level of staff training. Instruments may also be maintained, calibrated and operated differently (31). Difference in analyser performance from the same manufacturer can also be explained by differences in equipment and models and the use of kits with different lot numbers (32).

There are different approaches to bias and CV calculation, which may influence the final sigma calculation. For example, bias may be determined from External Quality Assessment reports instead of from IQC data as we did. Guo *et al*. (9) showed that both methods can be used for sigma metric determination and suggest that laboratories evaluate sigma metrics multiple times to optimize quality control schedules. The month-to-month variation noted with sigma was due to changes in CV and bias over these months. The CV, which is a measure of imprecision, was based on results over a 12-month period and therefore would be influenced by changes in reagents, calibrators or personnel. It would, therefore, be expected to be wider than CV determined over shorter periods with consequently lower sigma metrics (26).

To address the variability of sigma performance between different QC levels, months and lines, we could use the average sigma or lowest sigma metric for determining QC procedures (27). The suggested Westgard multirules will be easy to implement for analytes that are performing well such as PSA and DBIL. It will however be complicated for analytes with a low sigma metric such as the electrolytes; as well as those performing differently on the different levels of QC. The goal is to achieve 90% error detection and a 5% false rejection rate, while using the lowest possible number of control rules (7). The lower the error detection the more likely that more than one QC run will be required to detect a critical shift in performance (28). By selecting the appropriate control rule(s) and number of control samples, sigma metrics allows for the selection of QC procedures that prevent the reporting of erroneous patient results.

This is an important concept of risk management whereby QC strategies can be developed to prevent failures and lower the probability of patient harm. (30)

The level of performance that is based on the sigma metric has other practical implications for laboratories such as the design of QC programs. Laboratories that achieve high sigma metrics (requiring less control levels or run numbers) can reduce control material usage, recalibration and consumption of reagent and materials. In addition, it can decrease unnecessary troubleshooting and staff labour (18)

Moving average quality control has been suggested for high-volume analytes with low sigma metrics to decrease the risk of reporting erroneous results between scheduled quality control. Moving average QC uses a mathematical algorithm that can be incorporated into the middleware and calculate the average assay results. This may provide information regarding method stability and continuous (real-time QC) (29). Sigma metrics may also be an important tool when sourcing new products or methods e.g. the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) working group recommends that laboratories use sigma metrics to assess the quality of a method when selecting a new HbA1C method. (18) The result is more effective instrument operation, optimized laboratory workflow, and more reliable test results, ultimately assisting the clinician in the treatment of the patient (31).

Our study demonstrated the need for increased IQC frequency and calibration for some analytes such as the poor performance of electrolytes regardless of the TEa source used. The poor sigma performance of electrolytes is however not unique to our

laboratory. Potassium, sodium, and chloride have low biological variation and tight quality specifications are expected to give low sigma results (27,33). We used the QGI to investigate the reason for sigma < 3 and showed that the main problem was imprecision. It has been suggested that “a search for new and improved calibration methods may improve the precision and subsequently the sigma” (34). It has been noted that when imprecision is poor relative to analytical goals, good error detection is hard to attain regardless of the QC rules used (28).

This study had a few strengths and limitations. One of the strengths of this study is the 12-month time period which gave a good reflection of the data as estimation of accuracy and imprecision is expected to improve with more data points (28). We also looked at sigma performance of multiple analytes on two levels of QC. There was no lot change and performance related change can therefore be attributed to other factors. Comparing two identical analysers allowed us to assess the performance of analysers performing under the same environmental conditions and also improves the consistency of our findings. We also used different and widely available TEa guidelines to calculate sigma metrics. A limitation of this study is that no third party QC was used. The recommendation is that third party QC should be used whenever possible (28). Another consideration is looking at the performance at critical medical decision levels due to the implications on patient safety.

Conclusion

Laboratory results are crucial in the diagnosis, monitoring and prognostication of patients and often further action relies on the value of one test result. We should

therefore aim to minimize our errors that can affect patient outcomes. This is where sigma metrics has the potential of being a valuable quality management tool by monitoring analytical performance and comparing it to world class performance. However, it is important to set up standardized protocols for the determination of sigma metrics including the source of the bias and TEa.

References:

1. Jordan B, Mitchell C, Anderson A, Farkas N, Batrla R. The Clinical and Health Economic Value of Clinical Laboratory Diagnostics. EJIFCC. 2015;26(1):47–62.
2. Charuruks N. Sigma Metrics Across the Total Testing Process. Clin Lab Med. 2017;37(1):97–117.
3. Cooper G. Basic Lessons in Laboratory Quality Control. Bio-Rad Laboratories; 2008.
4. Westgard JO. Internal quality control: Planning and implementation strategies. Ann Clin Biochem. 2003;40(6):593–611.
5. Shah S, Saini R, Singh SB, Aggarwal O GA. Six Sigma Metrics and Quality Control in Clinical Laboratory. Int J Med Res Rev. 2014;2(2):140–9.
6. Specific Criteria for Accreditation of Medical Laboratories (NABL 112; Issue No 04, Issue date 11/02/2019). Haryana; 2019.
7. Litten J. Applying Sigma Metrics to Reduce Outliers. Clin Lab Med. 2017;37(1):177–86.

8. 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(7):973–80.

9. Guo X, Zhang T, Gao X, Li P, You T, Wu Q, et al. Sigma metrics for assessing the analytical quality of clinical chemistry assays: A comparison of two approaches. Biochem Medica. 2018;28(2 Special Issue).

10. Xia J, Chen S, Xu F, Zhou Y. Quality specifications of routine clinical chemistry methods based on sigma metrics in performance evaluation. J Clin Lab Anal. 2018;32(3):e22284.

11. Ricós C, Garcia-Lario J, Alvarez V, Cava F, Domenech M, Hernández A, et al. Biological Variation database, and quality specifications for imprecision, bias and total error. The 2014 update. Available from: https://www.westgard.com/biodatabase1.htm. Accessed January 8th, 2020

12. U.S Department of Health and Human Services. Medicare, medicaid, and CLIA Programs; Regulations Implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA). Federal Register. 1992;57(40):7002–186.

13. Royal College of Pathologists of Australasia. Allowable Limits of Performance for Biochemistry. Available from: http://www.rcpaqap.com.au/docs/2014/chempath/ALP.pdf. Accessed May 5th, 2019

14. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, et al. The EFLM Biological Variation Database. Available from: https://biologicalvariation.eu/meta_calculations. Accessed April 30th, 2020

15. Verma M, Dahiya K, Ghalaut V, Dhupper V. Assessment of quality control system by sigma metrics and quality goal index ratio: A roadmap towards preparation for NABL. World J Methodol. 2018;8:44–50.

16. Westgard SA. Consolidated Comparison of Chemistry Performance Specifications. Available from: https://www.westgard.com/consolidated-goals-chemistry.htm. Accessed May 13th, 2020

17. Westgard JO, Westgard SA. Westgard Sigma Rules. Available from: https://www.westgard.com/westgard-sigma-rules.htm. Accessed May 13th, 2020

18. Westgard S, Bayat H, Westgard JO. Analytical Sigma metrics: A review of Six Sigma implementation tools for medical laboratories. Biochem Medica. 2018;28(2):1–12.

19. El Sharkawy R, Westgard S, Awad AM, Ahmed AOI, Iman EH, Gaballah A, et al. Comparison between sigma metrics in four accredited Egyptian medical laboratories in some biochemical tests: An initiative towards sigma calculation harmonization. Biochem Medica. 2018;28(2):233–45.

20. Xia Y, Xue H, Yan C, Li B, Zhang S, Li M, et al. Risk analysis and assessment based on Sigma metrics and intended use. Biochem Medica. 2018;28(2):195–203.

21. U.S Department of Health and Human Services. Centers for Medicare and Medicaid Services. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Proficiency Testing Regulations Related to Analytes and Acceptable Performance - Proposed Changes. Federal R. Vol. 84. 2019.

22. Liu Q, Fu M, Yang F, Liang W, Yang C, Zhu W, et al. Application of Six Sigma for

evaluating the analytical quality of tumor marker assays. J Clin Lab Anal. 2019;33(2):e22682.

23. Aarsand AK, Røraas T, Bartlett WA, Coşkun A, Carobene A, Fernandez-Calle P, et al. Harmonization initiatives in the generation, reporting and application of biological variation data. Clin Chem Lab Med. 2018;56(10):1629–36.

24. Varela B, Pacheco G. Comprehensive evaluation of the internal and external quality control to redefine analytical quality goals. Biochem Medica. 2018;28(2 Special Issue):20710.

25. Cao S, Qin X. Application of sigma metrics in assessing the clinical performance of verified versus non-verified reagents for routine biochemical analytes. Biochem Medica. 2018;28(2 Special Issue):20709.

26. Li R, Wang T, Gong L, Peng P, Yang S, Zhao H, et al. Comparative analysis of calculating sigma metrics by a trueness verification proficiency testing-based approach and an internal quality control data inter-laboratory comparison-based approach. J Clin Lab Anal. 2019;33(9):1–9.

27. Tran MTC, Hoang K, Greaves RF. Practical application of biological variation and Sigma metrics quality models to evaluate 20 chemistry analytes on the Beckman Coulter AU680. Clin Biochem. 2016;49(16):1259–66.

28. Jones G, Calleja J, Chesher D, Parvin C, Yundt-Pacheco J, Mackay M, et al. Collective Opinion Paper on a 2013 AACB Workshop of Experts seeking Harmonisation of Approaches to Setting a Laboratory Quality Control Policy. Clin Biochem Rev. 2015;36(3):87–95.

29. Van Rossum HH. Moving average quality control: Principles, practical application

and future perspectives. Clin Chem Lab Med. 2019;57(6):773–82.

30. Yago M, Alcover S. Selecting statistical procedures for quality control planning based on risk management. Clin Chem. 2016;62(7):959–65.

31. Westgard S, Petrides V, Schneider S, Berman M, Herzogenrath J, Orzechowski A. Assessing precision, bias and sigma-metrics of 53 measurands of the Alinity ci system. Clin Biochem. 2017;50(18):1216–21.

32. Gülbahar Ö, Kocabiyik M, Çiraci MZ, Demirtaş C, Uçar F, Bayraktar N, et al. The use of six sigma methodology to evaluate the analytical performances of clinical chemistry analyzers. Turkish J Biochem. 2018;43(1):1–8.

33. Gami B, Patel D, Chauhan K, Shah H, Haridas N. Sigma Metrics as a quality marker for analyzing electrolytes in the laboratory. Int J Adv Res. 2013;1(7):197–201.

34. Moya-Salazar J, Pio-Dávila L. Evaluation of inter-batch variability in the establishing and quality control of glucose. Med Univ. 2016;18(71):85–90.

Supplementary material

Normalized Method decision charts

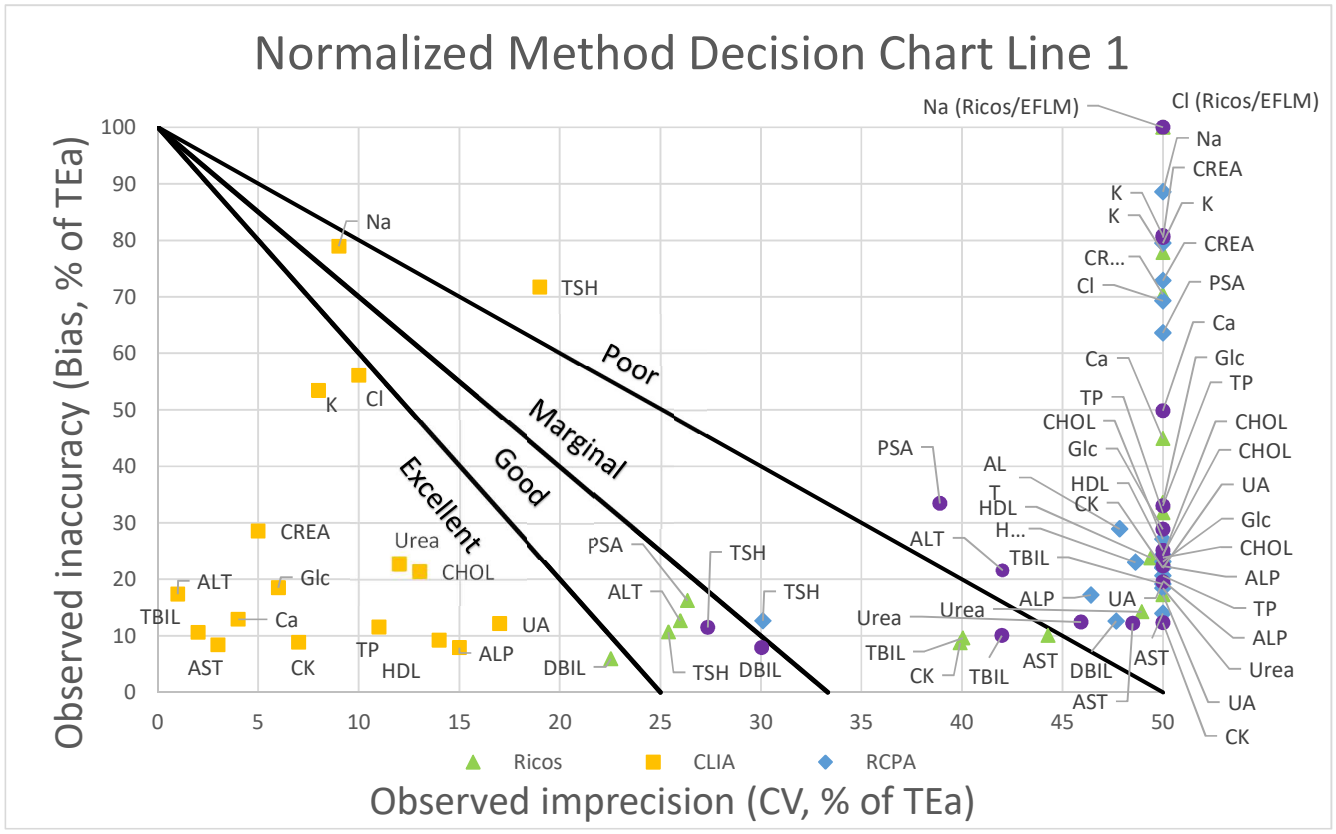


Figure 1a

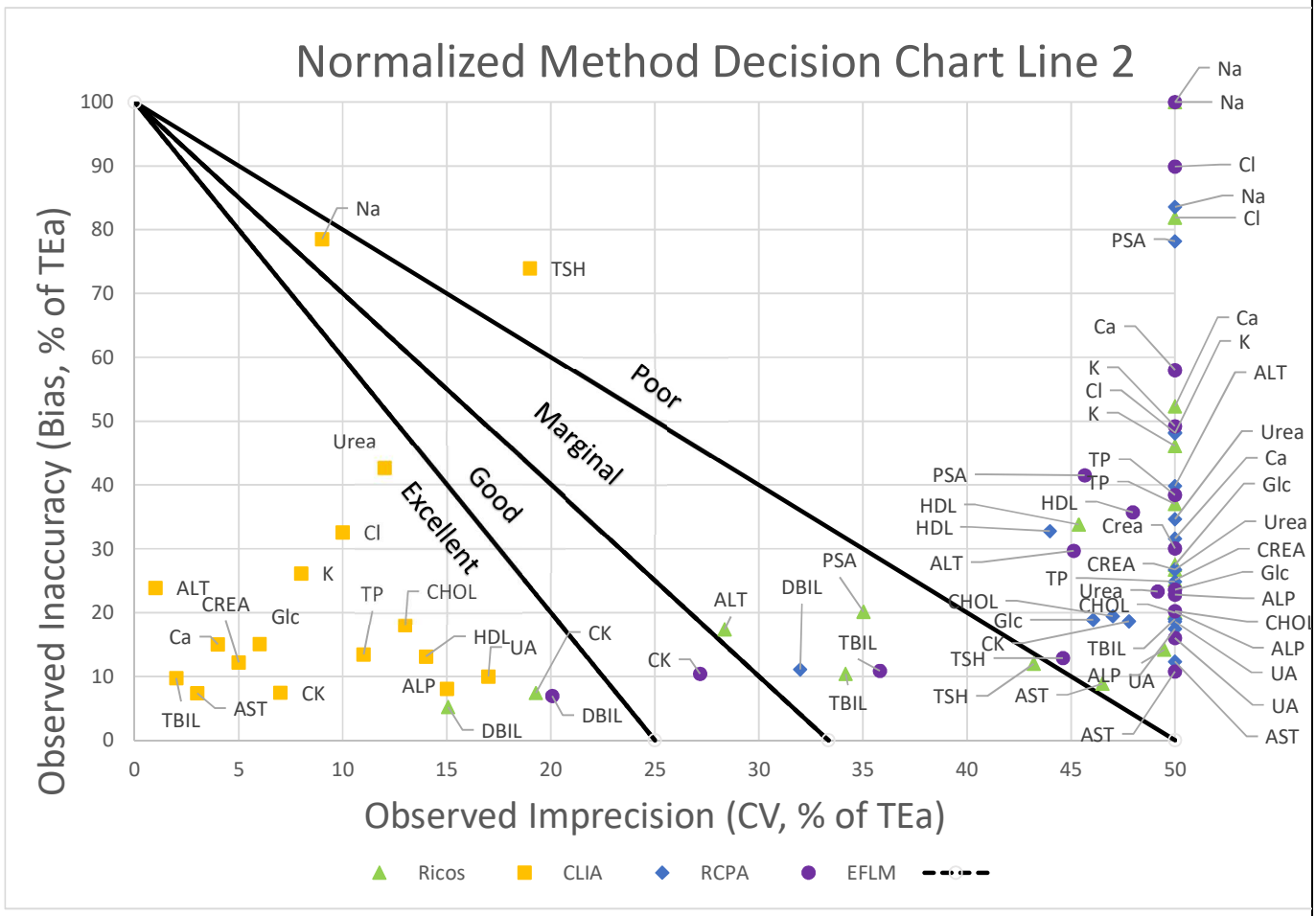


Figure 1b

Figure 1a and b: Normalized method decision charts demonstrating the performance of analytes based on different TEa guidelines on both analyzers

Line 1 monthly sigma

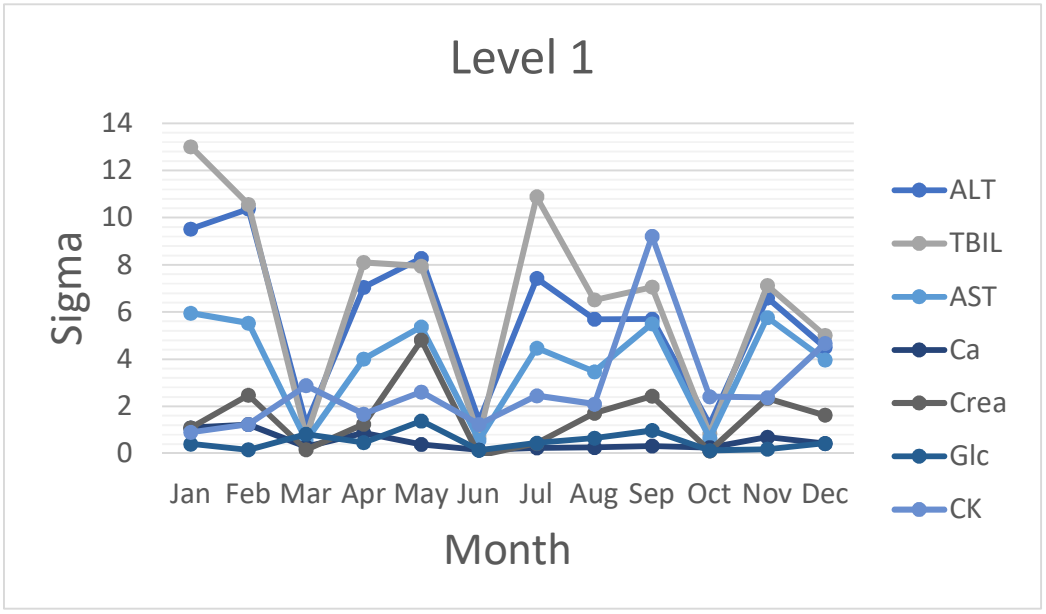


Figure 2a

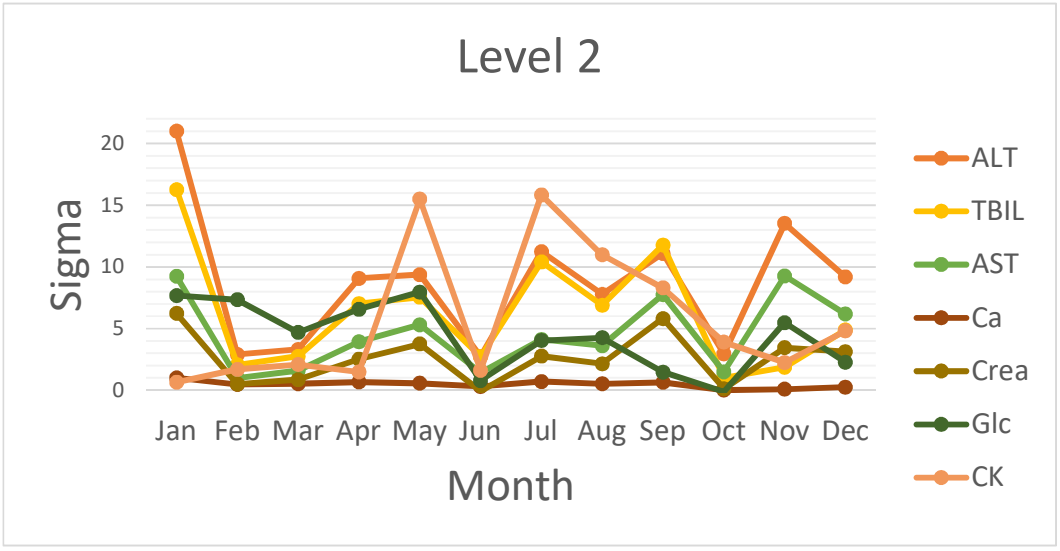


Figure 2b

Figure 2a and b: Monthly sigma performance of ALT, TBIL, AST, Ca, CREA, Glc and CK

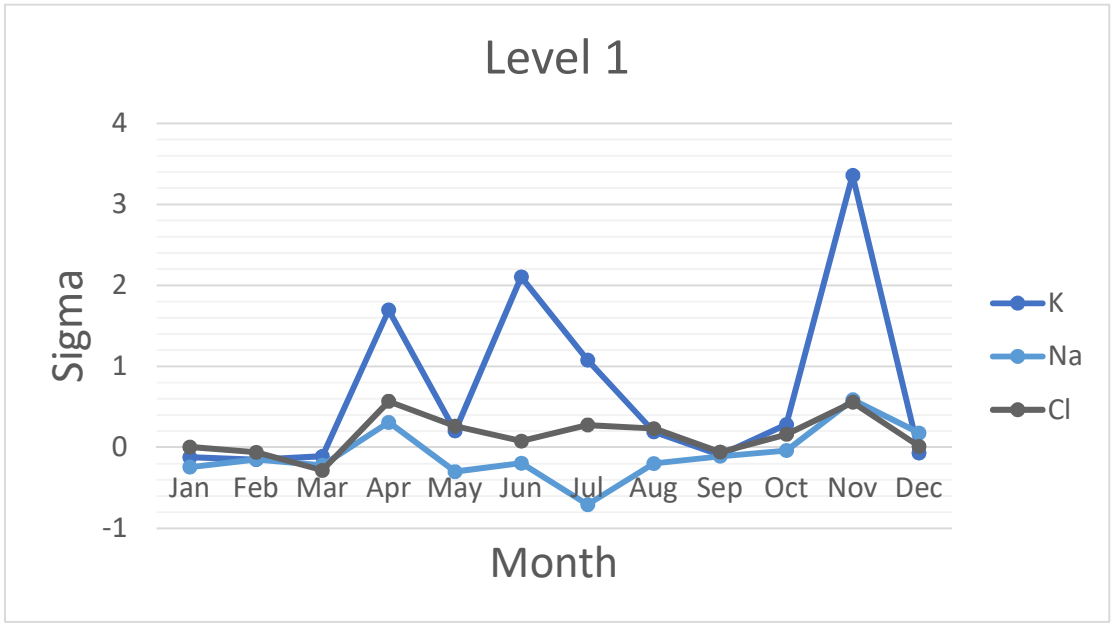


Figure 3a

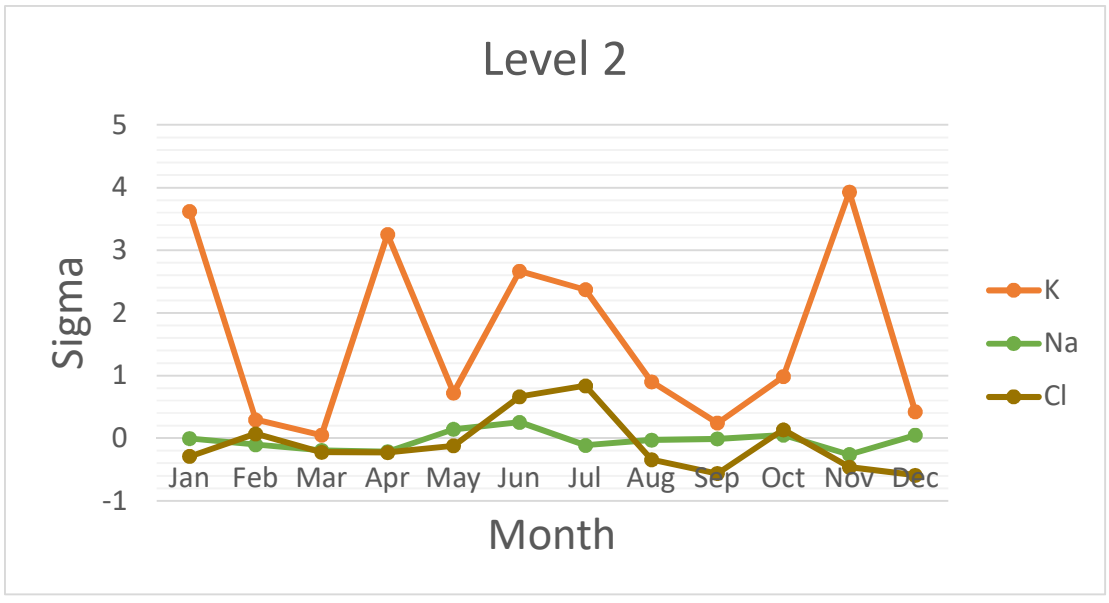


Figure 3b

Figure 3a and b: Monthly sigma performance of Na, K and Cl

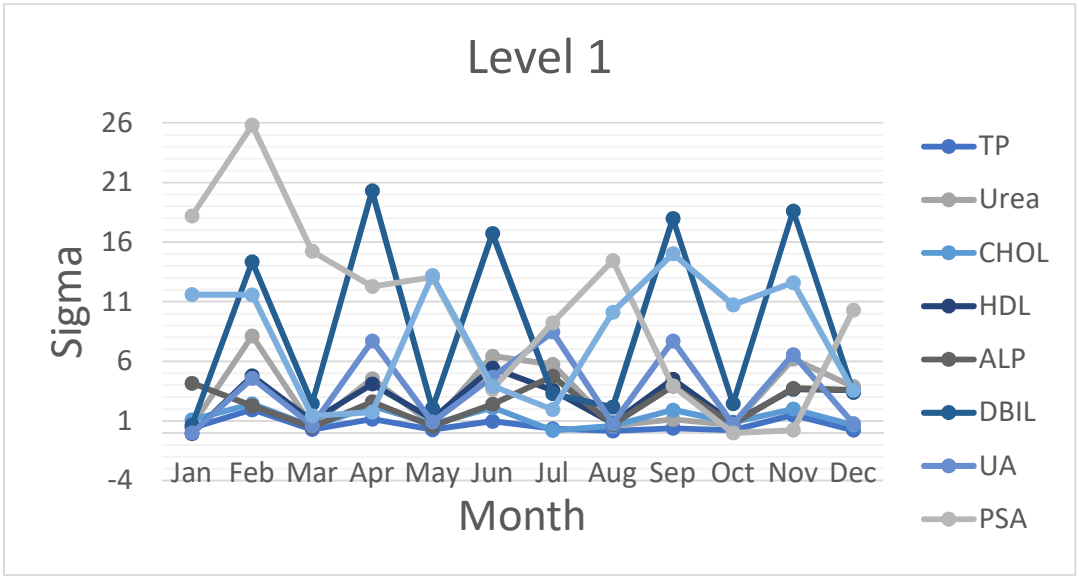


Figure 4a

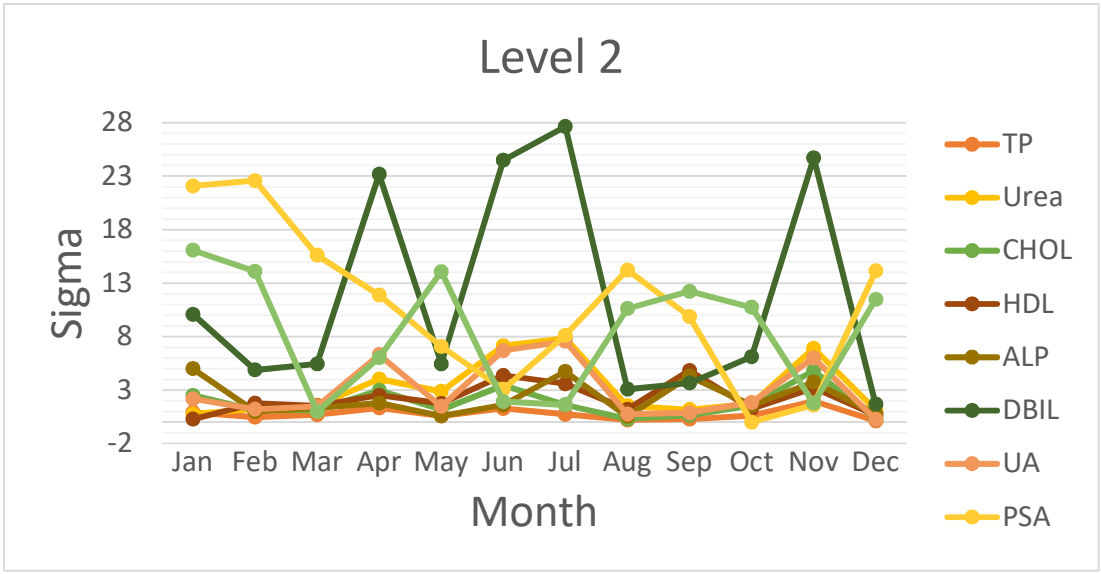


Figure 4b

Figure 4a and b: Monthly sigma performance of TP, urea, CHOL, HDL, ALP, UA and PSA

Line 2 Monthly sigma

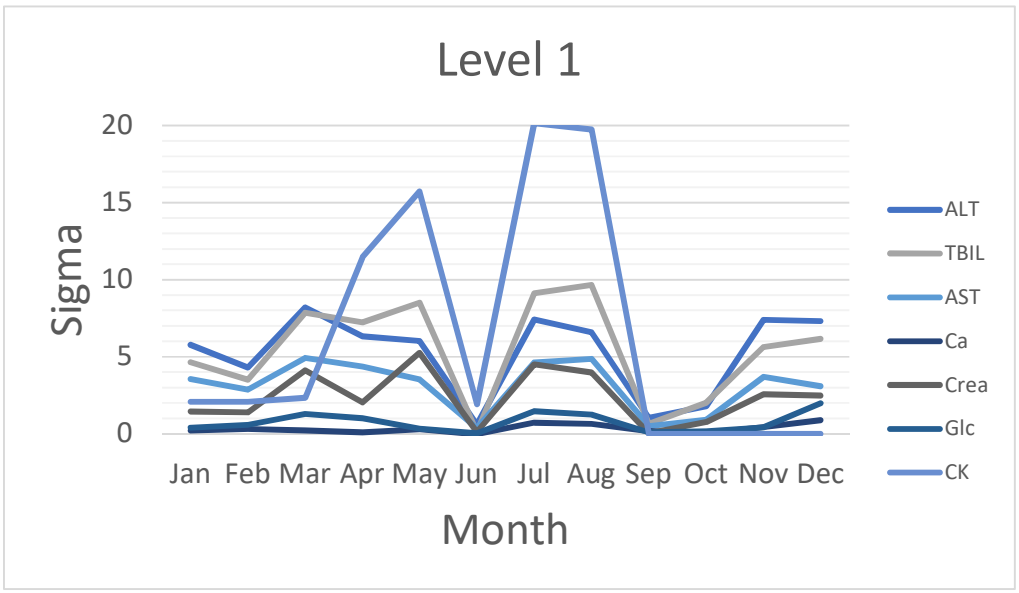


Figure 5a

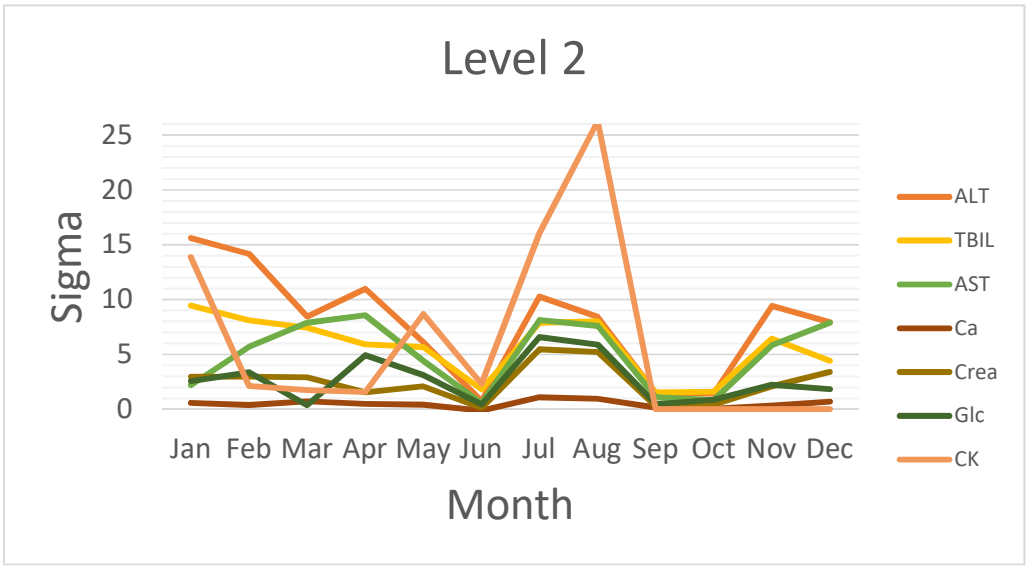


Figure 5b

Figure 5a and b: Monthly sigma performance of ALT, TBIL, AST, Ca, CREA, Glc and CK

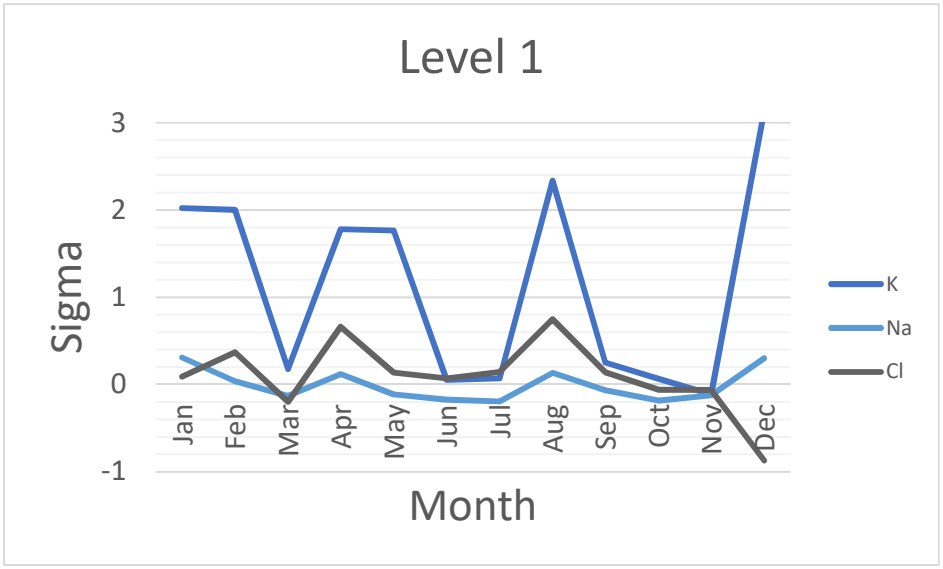


Figure 6a

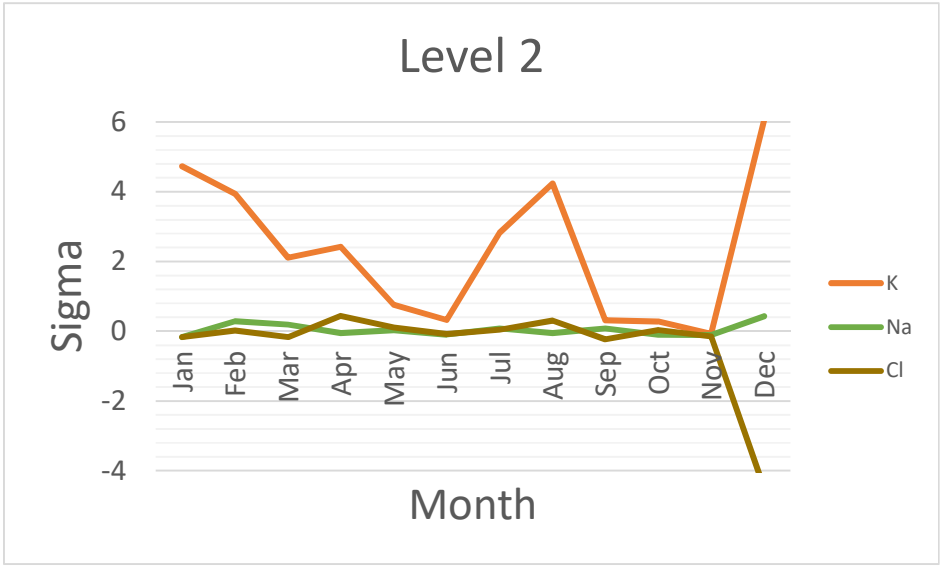


Figure 6b

Figure 6a and b: Monthly sigma performance of K, Na and Cl

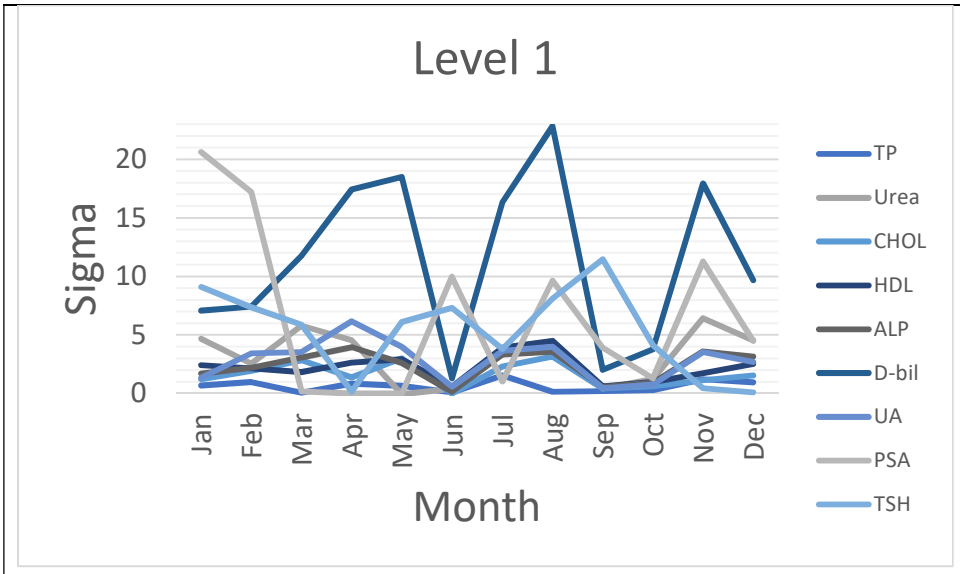


Figure 7a

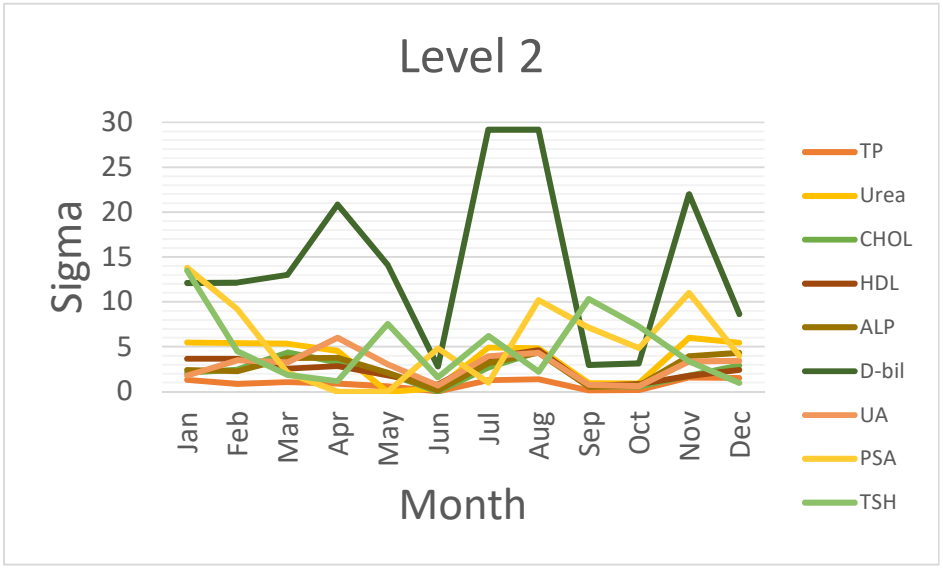


Figure 7b

Figure 7a and b: Monthly sigma performance of TP, urea, CHOL, HDL, ALP, DBIL, UA, PSA and TSH

Research Protocol
The application of sigma metrics in the laboratory to assess quality control processes

Dr Marli van Heerden
Student number: 0702096W

Introduction:

It is vital that medical laboratories produce accurate reproducible results as physicians rely on these for diagnosis, monitoring and prognostication of patients. To achieve this medical laboratories monitor and evaluate analytical processes using Quality Control (QC) processes which allow the laboratory to validate the performance of analyzers and reagents as specified by the manufacturer. Whilst we aim for laboratory results with no errors, in practice there are no processes with zero defects. [1; 2]. Errors may occur at any time; it is the laboratorians' duty to do their best to control and monitor these errors.

The sigma scale provides a convenient way to assess and compare the quality that is required using the precision and bias observed in a laboratory's performance. This scale measures the degree to which any process deviates from the target. The sigma value indicates how often defects are likely to occur; the higher the sigma value, the less likely the process will generate defects. Average services or products, regardless of their complexity, have a quality performance value of about 4-sigma. The best, or world-class, performance has a level of six-sigma.

Sigma metrics is a concept initially introduced by Motorola to reduce cost, eliminate errors, decrease variability and to prevent wasting of resources and effort. In business, manufacturing, service delivery, management and other industries it is used to measures the degree to which

a process deviates from its goal. Similarly, sigma metrics can be applied in the laboratory setting to monitor quality processes and alert us to change that needs investigation in order to achieve the desired level of quality control and confidence in our results.

The Chemical Pathology Laboratories of the National Health Laboratories currently do not use Sigma metrics to manage QC and the aim of this study is to calculate the sigma metrics of selected analytes and then suggest appropriate QC rules based on the Sigma metrics of the analytes selected.

Background and literature:

In analytical process, the examination procedure performances are typically evaluated in terms of precision and accuracy (bias). This is determined by the use of quality control procedures and is performed at intervals determined by laboratory policy. Interpretation of the performance of quality control is commonly done using Levey Jennings charts which are based on the mean and standard deviation (SD) of the control material. The mean provides the best estimate of the analyte's true value for a given QC level [4]. The SD determines how close the QC values are in relation to each other and refers to the total analytical standard deviation of the test method. A high standard deviation indicates poor precision, instability and high random error.

The Levey-Jennings chart is created by calculating the mean and SD for each level of QC and then plotting the daily quality control values. When the analytical process in functioning optimally, 68% of the values should fall within 1SD, 95.5% within 2SD and 99.7% within 3SD [5]. The Levey Jennings chart can indicate bias or imprecision as demonstrated in Figure 1a (bias) and Figure 1b (imprecision).

The probability of false rejections (due to random/inherent imprecision) and probability of analytical error detection should be kept in mind. The ideal QC system is stable, reliable, and easy to interpret and allows for continuous monitoring or correction [6].

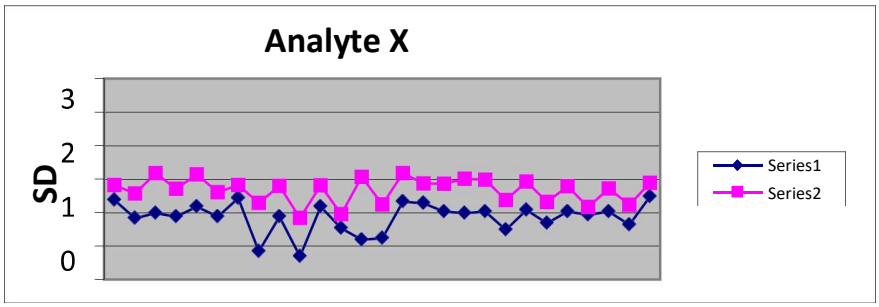


Figure 1a: Example of a Levey-Jennings chart (bias)

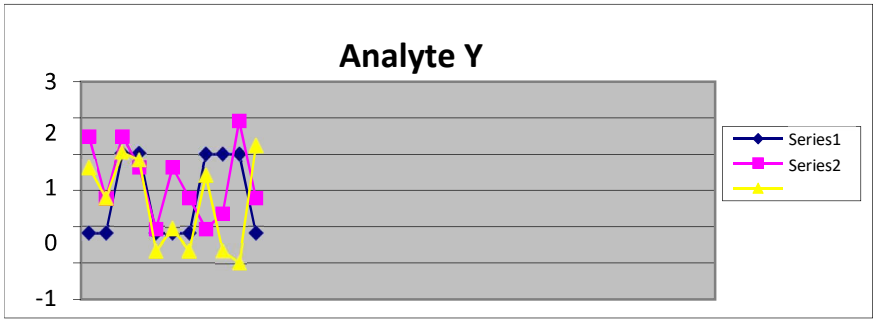


Figure 1b: Example of a Levey-Jennings chart (imprecision)

A number of rules may be applied to determine if the QC values are acceptable or not and some of these are illustrated in table 1 below

Table 1

13S The run is rejected when a single measurement exceed 3SD in either direction due to unacceptable random error	R4S: The run is rejected when the overall distance from the mean exceeds 4SD. This violation indicates that a random error is present.
12S: This rule is often used as a warning rule and the subsequent data should be evaluated carefully as a	10x: The run is rejected when 10 consecutive results are on the same side of the mean due to a small

random or systematic error might be present. If no error can be identified it is an acceptable random error.	systematic error or bias that can usually be resolved by maintenance or calibration
22s: The rule is violated when two consecutive QC results are greater than 2SD on the same side of the mean due to systematic error.	

[5; 7]

When multiple analytes are affected, recent changes to the system may be implicated such as failure to calibrate after a reagent change.

Two types of error are described:

Systematic errors often appear as a positive or negative shift and may be due to a number of causes including reagent/calibrator issues, may follow changes to the system after a major maintenance, problems with the sampling or dispensing system as well as temperature or light source deterioration. [5]

Random error implies a sudden, unexpected change from expected result. Causes include sampling errors from bubbles or clots, an electric supply problem, fluctuating temperature or an operator not following proper procedure [5].

If during the investigation no cause can be identified, troubleshooting guidelines and manufacturer's recommendations should be followed.

In a busy laboratory it is very time consuming to apply all the rules to every analyte and so the 2SD rule is applied as a warning and the 3SD as rejection rule. The 2SD control rule is generally regarded as a warning rule and if used as a control rule can cause a false rejection rate of up to 9% [8] Combinations of rules (multirules) are sometimes employed to reduce the rate of false rejections and to save time and effort.

The Westgard Multirules procedure, which uses a combination of criteria to assess an analytical run, increases sensitivity but can be too complex for routine use. The Westgard Multirules procedure generally uses 5 different control rules with 2 or 4 control measurements per run. The goal is to achieve 90% error detection and 5% false rejection rate, while using the lowest possible number of control rules [9].

Graphs provide a visual representation of processes and can assist in identifying problem areas requiring further investigation. Graphs include Method Evaluation Decision Charts, Sigma QC selection Graphs (sigma value added to a power function graph) and Normalized Operating Specifications (OPSpecs) charts. (See Appendix B). Normalized charts allows for different tests to be represented on the same chart [10]. This provides a practical way to choose the right QC method. When a method decision or OPSpecs chart is used, the CV and bias percentage needs to be calculated.

The Six Sigma scale is another method for quality management which is easily interpreted by laboratory staff. It runs from 0 to 6 and a value of 3 is regarded as the minimal acceptable performance [11]. Bill Smith, a Motorola engineer, is seen as the father of Six Sigma and he decided to quantify errors as defects per million (DPM) where 1 sigma corresponds to 698 000 defects per million and 6 sigma corresponds to 3, 4 errors per million [12].

Six Sigma can be used to decide on the best Westgard rule by judging the performance of a process against a reference method and to assess quality or identify processes needing improvement.

David E. Nevalainen was one of the first authors to apply sigma metrics in the laboratory setting [13]. In 2000 he published a study where quality indicators and national data were

converted to parts-per-million defects. He recognized the opportunity to improve laboratory performance, quality and cost by applying a strategy that is common practice in in other industries.

The number of QC levels and frequency of running QC varies greatly between laboratories [5]. According to the National Accreditation Board for Testing and Calibration Laboratories(NABL) guidelines, when a laboratory runs more than 75 samples per day at least two level QC's should be run at least twice a day [14]. Six sigma elaborates on this further by individualizing control rules based on analytical performance of the test

As the sigma increases, the consistency, reliability, robustness and performance of the test improves. This can provide guidance for designing QC strategies [5]:

A. High sigma values

This implies a good quality testing process and therefore simple QC rules are effective. The rationale is that the greater the error, the greater the probability for error detection (in addition to inherent random error) and the greater number of measurements and control rules are required. The false rejection rate should be kept as low as possible.

B. Low sigma values

More QC samples and more stringent rules are required. This simple parameter allows laboratories to easily determine the performance of current methods. For a six sigma of <3 the method performance needs to be re-evaluated as it is unstable and unacceptable. [11; 15]

According to Westgard "sigma metrics analysis provides not only an objective assessment of analytical methods and instrumentation, but it also provides the critical design information

needed for operational implementation". The guidelines proposed by Westgard are given in table 2 [16]:

Table 2

Sigma value	Number of QC samples per run	QC rule applied
Greater than or equal to 6	2	1:3S single control rule
5	2	Multirule: 13S/22s/R4S
4	4	Multirule: 13S/22S/R4S/41S
<4	6 (3 levels in duplicate)	Multirule including the 8X rule

The "Westgard Sigma Rules" graphic tool (Figure 2) was developed in an attempt to simplify control rule selection by choosing a sigma value and following a flow chart to guide the number of runs and number of controls per run. The Sigma value is read at the bottom of the chart and the control rules to the left of it indicates the ideal number of control levels (N) and runs (R). It should be kept in mind that the frequency of QC may also be affected by other events such as changes in reagent lots, calibrator lots, maintenance, replacement of parts etc.

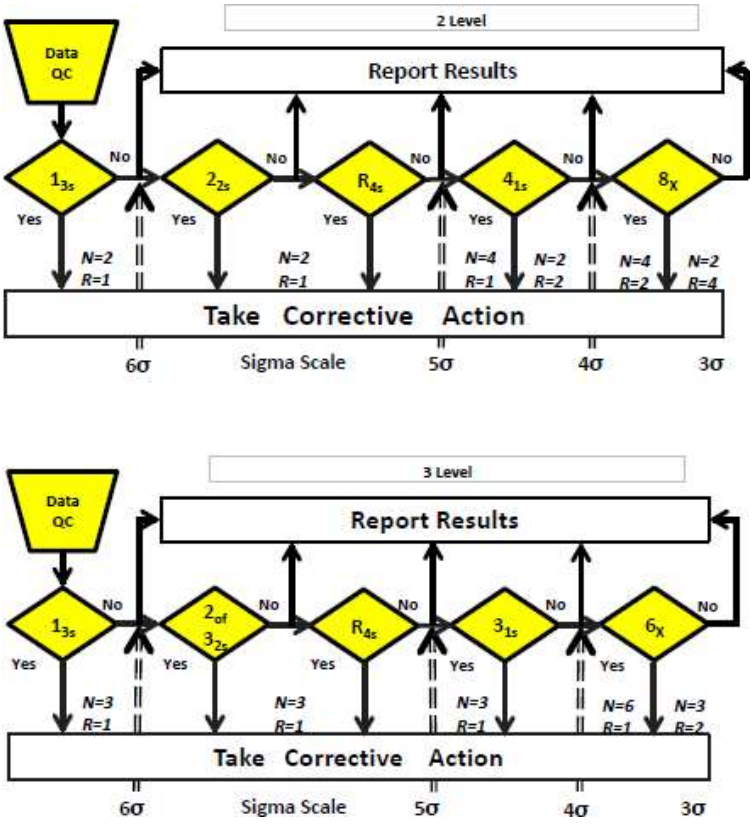


Figure 2

[IQCP-Startup-Kit4 -3-2015 download available from

https://www.westgard.com/downloads/worksheets-downloads/66-iqcp-startup-kit.html]

Month to month performance of analytes can be assessed using the critical systematic error.

The critical systematic error indicates how far the mean can shift before crossing the total allowable error limits. When the critical systematic error is zero, the chance of producing unreliable patient results is more than 5%. This also indicates that the total error is more than the total allowable error and corrective action should be taken.

There is limited information on the use of Six Sigma in clinical laboratories but results from a pilot study done in Ghana during 2014 [17] revealed unacceptable sigma values and high imprecision for all analytes. This indicates an unstable laboratory analytical system requiring further investigation to ensure reliability of patient results. In this study analytes representing endocrine, lipids, enzymes, electrolytes and the total protein levels were assessed over a four-month period. The decision was made to perform detailed assessments of the analytical procedures to determine the cause of the low sigma results.

Results of a pilot study done in India during 2015 [18] showed that the sigma values for sodium and potassium were unacceptable. Subsequent studies showed unacceptable sigma values for a number of other commonly run analytes [19] The authors came to the conclusion that sigma metrics can be used as a self-assessment tool to guide QC strategies and that upgraded analyzers, better methodologies and strict ion selective electrode maintenance in their laboratory may help achieve better sigma values.

Iqbal and Mustansar [20] performed sigma metric analysis in a tertiary hospital laboratory setting and found only five out of ten analytes investigated had acceptable sigma on both levels. Chloride had the lowest sigma (1.1) and creatinine the highest (10.1). This study allowed them to focus on strategies for problematic analytes.

In our laboratory the Levey Jennings control charts (Figure 1) and Westgard QC rules are applied to determine whether or not a QC run is acceptable based on an algorithm with specified limits. This approach might not be ideal as one set of rules cannot be applied to all tests due to varying precision and goals.

The aim of my study is therefor to apply sigma metrics in our laboratory to assess the performance of high-volume analytes. This may guide us in identifying our own problematic analytes and developing new QC procedures based on scientifically proven data.

Objectives:

- Determine the mean, SD and target value of selected analytes from data available on the Roche analyzer
- Determine the total error of selected analytes and compare these to the literature [3]
- Calculate the bias, total error and sigma-metrics of these analytes
- Assess the critical systematic error for each analyte monthly for the six-month period
- Suggest quality control rules based on Sigma metrics for the various analytes

Study design:

- This is a retrospective study for a six-month period at the Charlotte Maxeke Johannesburg Academic Hospital National Health Laboratory Service autolab during 2017
- We will determine the sigma matrix of the following 19 analytes:

General chemistry N= 672 per analyte	Tumour markers N= 672 per analyte	Endocrinology N= 672 per analyte	Electrolytes N= 1344 per analyte
ALT	PSA	Glucose	Sodium
Total and direct bilirubin		TSH	Potassium
AST		Total cholesterol	Chloride
Calcium		HDL	
Creatinine			
Total protein			
Urea			
ALP			
Uric acid			

These analytes were chosen to include high volume tests, endocrine components and analytes used in other studies as a way to compare sigma values.

Methods:

We will retrieve internal QC data for these analytes from the Cobas 8000 data manager. The QC for each analyte is routinely run twice a day unless otherwise indicated. There are two levels of QC per run to cover normal and abnormal ranges.

For all the analytes, except sodium, potassium and chloride, there will be four data points per day. There are two ion-selective electrode (ISE) modules on each line and there will be eight data points per day for sodium, potassium and chloride.

Most studies looked at a period between one and six months and we chose a six-month period to broaden our data range and determine the average values.

Calculations

The sigma metrics for each analyte will be determined using the following calculation

Sigma= Tea-Bias/SD

For a full description of each term, please see Appendix C. The bias is determined by subtracting the true value from the mean. The SD and mean is available on the instrument and is determined by precision studies. The total error is calculated using the bias and SD:

TE= Bias + Z*SD

If the Z (chance of exceeding the quality requirement) is set to 1.645 (1.65) then 95% of measurements errors should fall within the TE limits or 5% error may occur before rejecting the run. If the Z is set to 2.33 (2) then 99% of measurements should fall within the TE limits [9].

The critical systematic error will be calculated by subtracting 1.65 from the calculated sigma value.

All data will be captured on an Excel spreadsheet. Please see Appendix A.

Variables that may affect the QC results and sigma calculations include:

- QC material reconstitution, handling and storage
- Lot number changes (to allow month to month comparison)
- Random and systematic error
- Maintenance: Temperature/calibration etc.

Total error will be calculated for each analyte and compared to published guidelines.

The proposed Westgard rules for the sigma values obtained will be determined and compared to what is currently applied in our laboratory. Low sigma values will require investigations such as more QC to be added or more runs per cycle. High sigma indicates satisfactory performance which requires less stringent rules and less time spent on the analyte.

Critical systematic error will be determined for each analyte for the six-month period.

Excepted outputs:

1. To obtain my MMED degree
2. Publication
3. Provide QC recommendations for the NHLS laboratory at Charlotte Maxeke Johannesburg Academic Hospital e.g. more stringent QC rules to be applied when the sigma metrics values are low.

Budget: Not applicable

Ethics: To be obtained

Work plan:

Protocol presentation	22/01/2018
Protocol submission	31/01/2018
Ethics submission	05/02/2018
Assessor meeting	21/02/2018
Data collection	01/03/2018-30/04/2018
Data analysis	01/05/2018-31/07/2018
Write-up	01/08/2018-31/12/2018

References

1. Westgard J.O., Seehafer J.J., and Barry P.L.: Allowable imprecision for laboratory tests based on clinical and analytical test outcome criteria. Clin Chem 1994; 40(10): 1909-1914
2. Chinchilli V.M., and Miller W.G.: Evaluating test methods by estimating total error. Clin Chem 1994; 40(3): 464-471
3. Cooper WG. Basic Lessons in Quality Control. Bio-Rad Laboratories, Inc. Clinical Diagnostics Group. 1998. P15.
4. Ricos C, Alvarez V, Cava V Et al. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999; 59(7): 491-500. [2014 updated version available from https://www.westgard.com/biodatabase1.htm]
5. Shah S, Saint R, Singh SB et al. Six Sigma Metric and Quality Control in Clinical Laboratory. Indian J. Med Res. 2014: 2(2): 140-149.
6. Badrick T. The Quality Control System. Clin Biochem Rev. 2008; 29(Suppl 1): 67-70.
7. Westgard JO. Westgard rules and Multirules. [homepage on the internet] Available from: https://www.westgard.com/mltirule.htm
8. Westgard JO, Barry PL, Hunt MR and Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem, 1981; 27(3): 493-501.
9. Curtis A, Parvin A, Yundt-Pacheco J et al. Sigma Metrics, Total Error Budgets and QC. Bio-Rad laboratories. 2016 December 06. Available from: http://www.qcnet.com/rabid/8002/default.aspx
10. Westgard JO, Westgard SA. Total Analytic Error. Clinical Laboratory News. 2013 Sep 1. Available from: https://www.aacc.org/publications/cin/articles/2013/September/total-analytical-error
11. Westgard S. Six Sigma Metric Analysis for Analytical Testing Processes. Available from website: https://corelaboratory.abbott/sal/whitePaper/SixSigma_WP_MAATP_ADD-00058830
12. Sawalakhe PV, Deshmukh, Lakhe RR. Evaluating Performance of Testing Laboratory using Six Sigma. Int J Innov Eng Res Techn. 2016; 1(1): 13-20
13. Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L, Morgan T. Evaluating laboratory performance on quality indicators with the six sigma scale. Arch Pathol Lab Med 2000; 124(4):516-519.

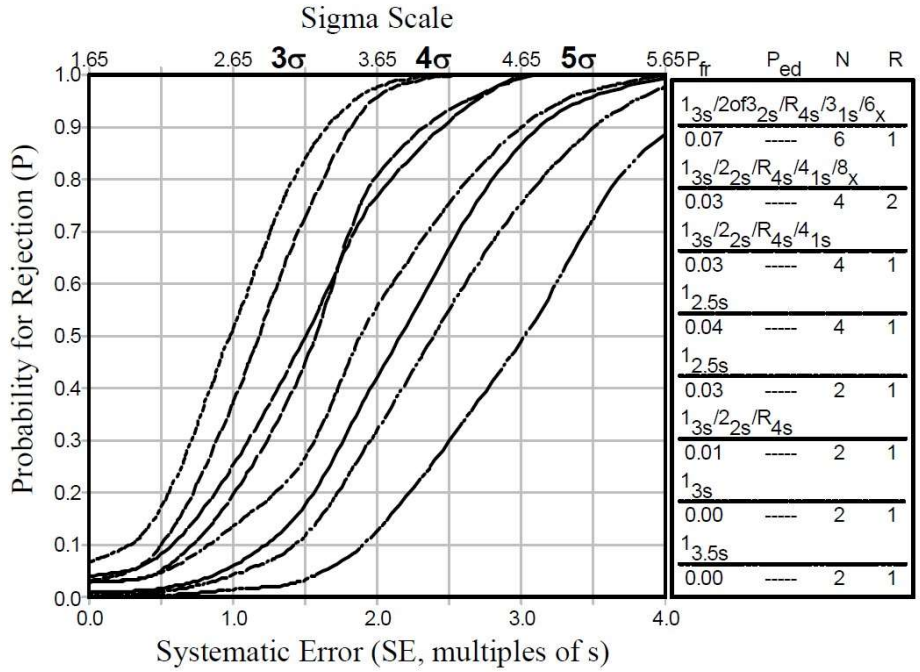
14. NABL 112: Specific criteria for Accreditation of Medical Laboratories: 2008 1 Feb; 03:28-45. Amend 2012 16 Oct; 03:28-45.
15. Nanda AK, Ray L. Quantitative Application of Sigma Metrics in Medical Biochemistry. J Clin Diagn Res. 2013; 7(12): 2689-2691.
16. Westgard JO, Westgard SA. Introducing Westgard Sigma Rules. Available from: https://www.westgard.com/westgard-sigma-rules.htm. September 2014.
17. Afrifa J, Gyekye SA, Owiredu WKBA et al. Application of sigma metrics for the assessment of quality control in clinical chemistry laboratory in Ghana: A pilot study. Niger Med J. 2015; 56(1): 54-58.
18. Adiga A, Preethika A. Sigma Metric of Electrolytes- A Pilot Study. Int J Res Stud Biosci. 2015; 3(11): 33-37.
19. Adiga US, Preethika A. Sigma metrics in clinical chemistry laboratory- A guide to quality control. Al Ameen J Med Sci. 2015; 8(4): 281-287.
20. Iqbal S, Mustansar T. Application of Sigma Metrics Analysis for the Assessment and Modification of Quality Control Program in the Clinical Chemistry Laboratory of a Tertiary Care Hospital. Indian J Clin Biochem. 2017; 32(1): 106-109.

Appendix A:

Analyte	Bias	TE	TEA	SEc	Sigma		Mean	True value	SD	%TEA
	‘=(H2-I2)	‘=(B2+(1.65*J2))	‘=((K1/100%)*I2)	‘=((D2-B2)/J2)-1.65	‘=E2+1.65					
ALT										
T-Bili										
AST										
Calcium										
Creatinine										
Glucose										
CK										
Potassium										
Sodium										
Chloride										
TP										
Urea										
T-Cholesterol										
HDL										
ALP										
D-bil										
UA										
PSA										
TSH										

Appendix B:

Sigma-Metrics QC Selection Tool for 2 Levels Control



P_{fr}	P_{ed}	N	R
$1_{3s}/_{2of3}2_{s}/_{R4s}/_{3}1_{s}/_{6X}$			
0.07	-----	6	1
$1_{3s}/_{2}2_{s}/_{R4s}/_{4}1_{s}/_{8X}$			
0.03	-----	4	2
$1_{3s}/_{2}2_{s}/_{R4s}/_{4}1_{s}$			
0.03	-----	4	1
$1_{2.5s}$			
0.04	-----	4	1
$1_{2.5s}$			
0.03	-----	2	1
$1_{3s}/_{2}2_{s}/_{R4s}$			
0.01	-----	2	1
1_{3s}			
0.00	-----	2	1
$1_{3.5s}$			
0.00	-----	2	1

Test _____

Allowable Total Error (TEa%)
or (TEa units) _____

Clinical Decision Concentration _____

Observed Imprecision (CV%)
or (SD units) _____

Observed Inaccuracy (Bias %)
or (Bias units) _____

Sigma [(TEa%-Bias%)/CV%)]
or [(TEa_{Units}-$Bias_{Units}$)/SD_{Units}] _____

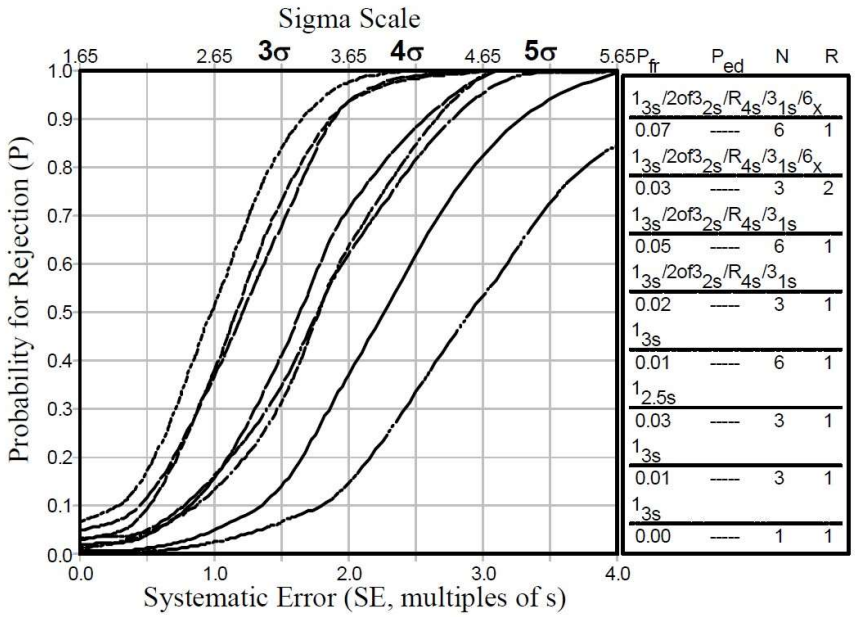
Critical Systematic Error
[Sigma – 1.65] _____

Selected QC Procedure _____

Analyst _____

Date _____

Sigma-Metrics QC Selection Tool for 3 Levels Control



Test _____

Allowable Total Error (TEa%)
or (TEa units) _____

Clinical Decision Concentration _____

Observed Imprecision (CV%)
or (SD units) _____

Observed Inaccuracy (Bias %)
or (Bias units) _____

Sigma [(TEa%-Bias%)/CV%)]
or [(TEa_{Units}-$Bias_{Units}$)/SD_{Units}] _____

Critical Systematic Error
[Sigma – 1.65] _____

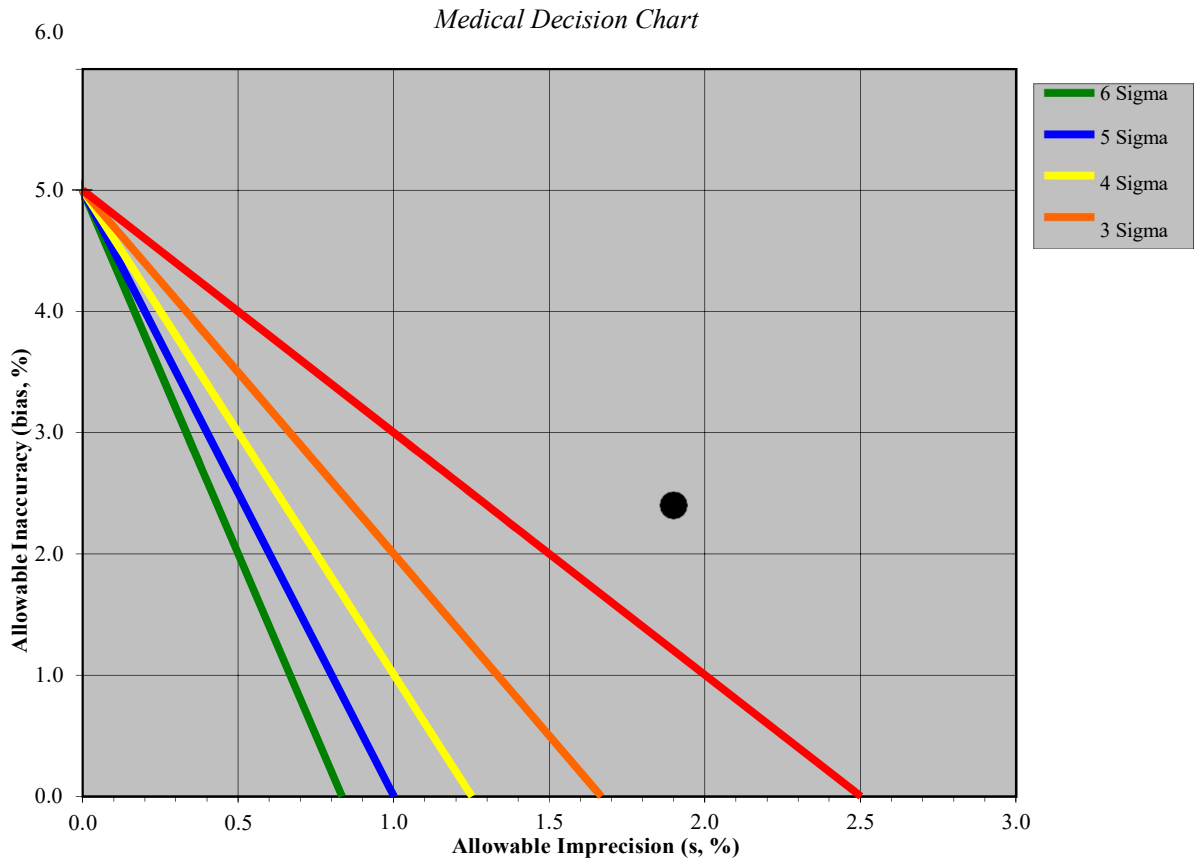
Selected QC Procedure _____

Analyst _____

Date _____

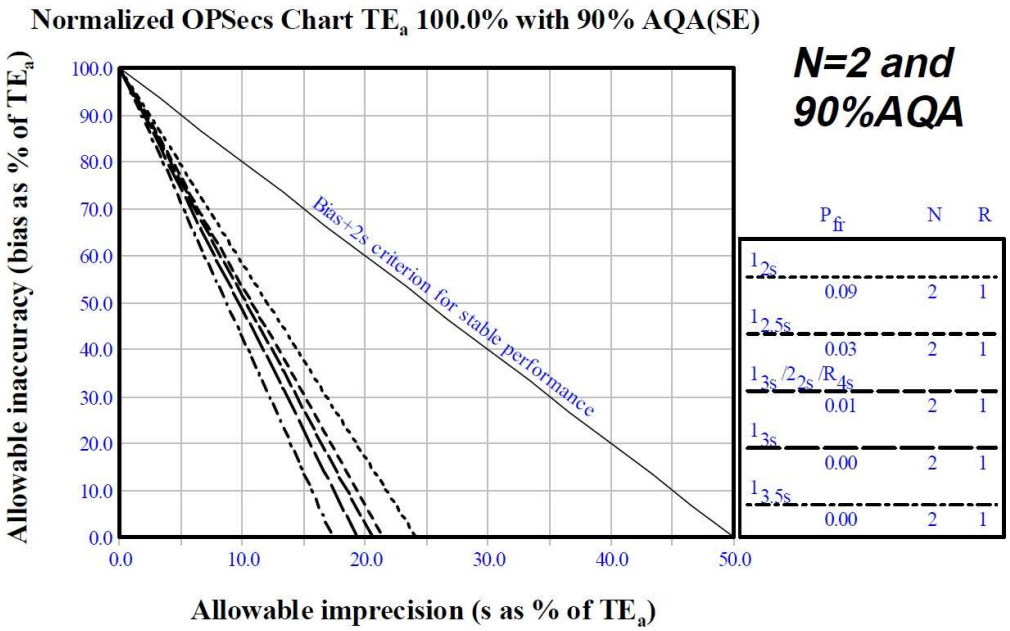
Sigma Metrics Tool Handout available from:

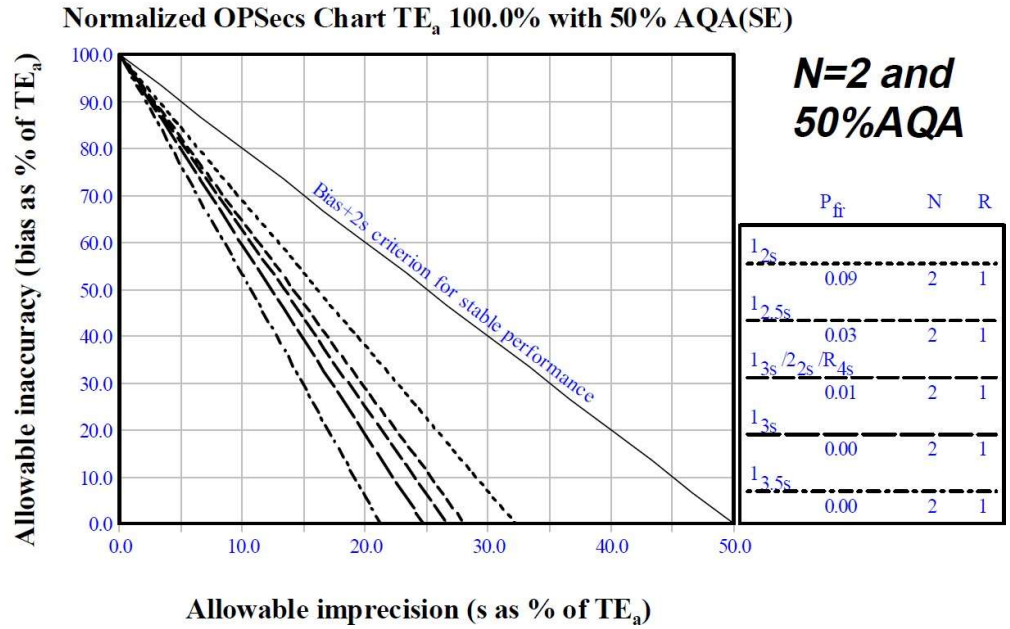
https://www.westgard.com/downloads/worksheets-downloads/40-sigma-metrics-tool-handout/file.html



Method Performance Evaluation chart available from:

https://www.westgard.com/downloads/worksheets-downloads/21-six-sigma-medx-chart.html





Normalized

OPSpecs Chart available from: https://www.westgard.com/downloads/worksheets-downloads/22-nopspecs-charts/file.html

Appendix C

Terms:

C*V* (coefficient of variation): Calculated by dividing the standard deviation by the mean. The CV allows comparison of precision especially when two different methods are compared [18]. Possible sources of acceptable CV include the instrument manual/ package insert, inter-laboratory comparison programs, proficiency schemes, journals or regulatory bodies' proficiency limits.

Bias: The difference between the result obtained by a laboratory method and the result that would be obtained from an accepted reference method [18]. Bias can have a big impact on analytical quality and should be investigated where identified [9].

Critical systematic error: Indicates how far the mean can shift before quality performance requirements are jeopardized. It is calculated by subtracting 1.65 from the sigma.

Total error budget (TEB): Includes the test system bias and imprecision. The TEB should not exceed 50% [9] as it is associated with poor sigma values.

Total allowable error: refers to the analytical requirements that need to be maintained to ensure clinically valid results [10]. It is based on a test system's analytical imprecision (random error) and bias (systematic error). If the total allowable error is exceeded the result is unreliable [15]. In other words, the total allowable error is the combined imprecision and inaccuracy that can be tolerated without negatively impacting patient care based on that result [11]. This allows the laboratory to assess method performance, calibration status, reagent lot number variability, analytical system changes and appropriate QC rules. Method validation studies, proficiency testing or external quality assessment surveys as well as routine QC data

can be used to make estimates. Regulatory bodies such also provide acceptable performance criteria.

Sources of Total allowable error according to the Stockholm hierarchy of desirability:

1. Medical decision levels from peer-reviewed data and publications. This is specifically relevant to glucose and cholesterol.
2. Biological variation within and between individuals is taken into account
3. Proficiency testing schemes. The peer group's median CV% can be multiplied by 3.

   This however allows for large variation and may not cover the preferred reference ranges.
4. Tonks' rule: [(Reference range span/4)/mean of range x 100%].
5. Current SD x 3

When the TAE provided is in units the upper and lower limits is calculated by adding and subtracting the value from the target value. When both units and a percentage are given the greater value of the two should be applied.

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research & Post Graduate Affairs)

TO: Dr M van Heerden
School: Pathology
Department: Chemical Pathology
NHLS

E-mail: mvanheerden88@gmail.com

CC: Supervisor: Professor J George <Jaya.George@nhls.ac.za>
and <HREC-Medical.ResearchOffice@wits.ac.za>

FROM: Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 16/02/2018

REF: *R14/49*

PROTOCOL NO: W-CBP-180216-01 *(This is your ethics application study reference number. Please quote this reference number in all correspondence relating to this study)*

PROJECT TITLE: *The application of sigma metrics in the laboratory to improve quality control processes*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to the Government funding of the University.

MSWorks2000/Iain0007/ClearScanWaiver.wps