Original Research Article

Sigma metrics as quality indicators in guiding and tracking laboratory process improvement

Sohini Sengupta[1]*, Douglas W. Chung[2], Ramdhir Kumar[3], Paritosh Gupta[4], Sonu Upadhyay[4], Ashok Rattan[5]

[1]Laboratory Medicine, Redcliffe Labs, India
[2]Abbott Laboratories, Singapore
[3]Abbott Laboratories, India
[4]National Reference Laboratory, Redcliffe Labs, India
[5]Medical Committee and Quality, Redcliffe Labs, India

*Correspondence:
Dr. Sohini Sengupta,
E-mail: dr.sohini@redcliffelabs.com

ABSTRACT

Background: To demonstrate the utility of sigma metrics towards assessing the quality of processes, and optimization of statistical quality control rules in a high-volume clinical laboratory, in a two-phase quality improvement project.
Methods: In the "pre" period, the sigma score was assessed across 25 routine high-volume assay parameters in our laboratory, comprising of 20 clinical chemistry and 5 immunoassay methods. Measures were taken to improve the analytical quality of low sigma score parameters within a 6-month period. Another sigma metric analysis was then performed in the "post" period to examine any measurable improvement.
Results: The average sigma metric increased from 6.4σ to 9.2σ. Out of 25 analytes, 17 showed a significant improvement, defined as an increase in the sigma metric by greater than 1.0.
Conclusions: The changes in sigma metric had a significant positive impact on the DPMO and reinforced the reliability of our test results. It showed that our quality control processes can be streamlined and simplified further, to optimize the frequency of internal quality control, while still maintaining the same level of error detection and analytical quality assurance. The analysis also provided additional benefits of achieving lesser errors, fewer sample reruns and troubleshooting, and improved turnaround time, for better clinician and patient satisfaction.

Keywords: DPMO, Quality assurance, Sigma metrics, Total allowable error

INTRODUCTION

Six Sigma is a data-driven process improvement approach, first originated in Motorola and GE, which has been widely adopted in many industries since the 1980s. At the center of the Six Sigma approach is the sigma metric, represented by the Greek letter sigma (σ), which is a measurement of the efficiency of the process in staying within certain quality specifications. As the sigma metric increases, process quality improves, less errors are produced and the Defect Per Million Opportunities (DPMO) metric gets logarithmically smaller. Hence, the goal of Six Sigma is to achieve "world-class" performance at 6σ which translates to less than 3.4 DPMO.

In the clinical laboratory, the total testing cycle can also be viewed as a large-scale industrial process, where thousands of results are generated every day and must conform within certain quality specifications. In 2000,

Nevalainen et al. performed the first study that examined the clinical laboratory's performance on a sigma scale, and did so across the pre-analytical, intra-analytical and post-analytical phases.[1] In 2001, Westgard incorporated the sigma approach into statistical quality control processes within the intra-analytical phase.[2]

Since those days, a large body of studies have been published across different applications: assessing analytical quality for many analytes on one analyzer; comparing quality for one analyte across multiple analyzers; comparing performance for multiple analytes between multiple analyzers; tracking performance across multiple sites in a regional network; and pooling worldwide data to derive representative insight on the global performance of specific analyzer platforms such as the Alinity ci-series and the Atellica Solution.[3-8]

While such studies demonstrate that it is possible for a particular methodology or analytical system to reach 6σ performance for a given analyte, we observe that there is significant variation in performance even between different laboratories using the same platform and the same reagents, as discovered by Taher et al, where analytes including chloride, glucose and β HCG had widely variable sigma metrics ranging from 3σ in some sites to over 7σ in others.[7] Thus, it is worth noting that beyond instrument technology and material formulation factors that are dependent on the manufacturers, the local operational factors at the individual laboratory level have a significant impact on analytical quality.

This study aims to investigate and demonstrate the utility of sigma metrics in a two-phase quality improvement project. In the "pre" period, baseline performance was assessed across 25 routine high-volume assay parameters in our laboratory comprising of 20 clinical chemistry and 5 immunoassay methods. For each analyte, imprecision was determined from the cumulative percentage CV in daily quality control and the inaccuracy is determined by the percentage bias versus the peer group in the external quality assurance assessment programme. From this data, we derived the baseline sigma metric and provided guidance on which assays needed improvement. Error reduction tactics were taken to improve the analytical quality of these parameters within a 6-month period, following which another sigma metric analysis was taken in the "post" period to examine whether there was any measurable improvement. A review of the literature suggests that this is the first study examining this application of sigma metrics, in demonstrating the measurable impact of quality improvement via the longitudinal tracking of sigma metrics for the same parameters on the same analyzer over time.

METHODS

The study was performed at the National Reference Laboratory of Redcliffe Labs, India. Daily quality control data was first collected in August 2022 for 20 clinical chemistry analytes and 5 immunoassay analytes on the Alinity ci-series instrument from Abbott Laboratories (Chicago, IL, USA). For each parameter, sigma metrics were calculated in the "pre" and "post" periods using the formula:

$$\text{Sigma} = (\text{TEa} - [\%\text{Bias}])/\%\text{CV}$$

For the determination of imprecision, two levels of quality control were run twice daily, and the statistical data including laboratory mean and percentage CV were gathered at the clinical decision level. Independent third-party quality control materials from Randox Laboratories (Crumlin, UK) were used, which were human assayed multi-sera normal control (lot number 1543UN) and elevated control (lot number 1211UE), Immunoassay Premium Plus Control (lot numbers 2105EC & 2107EC).

For the determination of inaccuracy (or bias), the laboratory participated in external quality assurance programs from Bio-Rad Laboratories (Hercules, CA, USA), EQAS Clinical Chemistry Cycle 23 and EQAS Immunoassay Cycle 19, which provided an estimation of the bias relative to the peer group from other laboratories around the world for the study period.

The Total Allowable Error (TEa) specification for each analyte is taken from publicly available sources and after discussion with our laboratory management. These sources include CLIA, biological variation specifications taken from the European Biological Variation Study (EuBIVAS), the Spanish Society of Clinical Biochemistry and Molecular Pathology, the Ricos database and CAP limits, and were decided depending upon medical decision points as shown in Table 1.[9,10]

Table 1: Total allowable error for each analyte used in calculation of sigma metrics.

Test	Abbreviation	TEa	Source
Albumin	ALB	10%	CLIA
Alkaline phosphatase	ALKP	20%	CLIA
Alanine aminotransferase	ALT	16.10%	EuBIVAS
Aspartate aminotransferase	AST	16.70%	EuBIVAS
Bilirubin, direct	DBIL	44.50%	Ricos Desirable
Bilirubin, total	TBIL	20% or 0.4 mg/dL	CLIA
Calcium	CAL	1 mg/dL	CLIA
Chloride	CL	5%	CLIA

Continued.

Test	Abbreviation	TEa	Source
Cholesterol	CHOL	10%	CLIA
Creatinine	CREA	15% or 0.3 mg/dL	CLIA
Gamma-glutamyl transferase	GGT	22.11%	Ricos desirable
Glucose	GLUC	10% or 6 mg/dL	CLIA
HDL cholesterol	HDL	17.40%	Ricos 2014 Min
Phosphorous	PHOS	10.7% or 0.3 mg/dL	CAP
Potassium	K	0.3 mEq/L	CLIA
Protein, total	TPRO	10%	CLIA
Sodium	NA	4 mEq/L	CLIA
Triglycerides	TRIG	18%	Spanish minimum
Urea nitrogen	UREA	9% or 2.0 mg/dL	CLIA
Uric acid	URIC	17%	CLIA
Thyroid stimulating hormone	TSH	23.7%	Ricos desirable
Total T3	TT3	24%	Spanish minimum
Total T4	TT4	24%	Spanish minimum
Vitamin B12	B12	30%	Ricos desirable
Vitamin D	VITD	25%	NYS

CLIA: Clinical Laboratory Improvement Amendments. NYS: New York State Department of Health Clinical Laboratory Evaluation Program. Spanish Minimum: 2015 Spanish Minimum Consensus Performance Specifications. Ricos: see publication. EuBIVAS: European Biological Variation Study

After the "pre" data was gathered, various quality improvement tactics were proposed through discussion between the laboratory staff, and experts and specialists from the instrument manufacturer. Some of these actions (shown below in Table 2) were implemented during the improvement period and the sigma metrics were re-assessed in the "post" period in January 2023.

Table 2: Some of the quality improvement actions taken in the laboratory.

Actions
Setting the QC target value to reflect the correct cumulative statistical mean
Reporting results to the appropriate number of decimal places for each analyte
Ensuring all material stability and shelf life requirements are being adhered
Following the appropriate calibration frequency as per manufacturer recommendations
Improvement of reagent and QC handling, aliquoting, and running practices in the laboratory
Improvement of training of operators and technical staff to ensure proper maintenance of equipment

A normalized method decision (MEDx) chart is also plotted to graphically represent the data in two dimensions, following the procedure described by Westgard which yields more insight into the method performance of each analyte by mapping the inaccuracy on the y-axis and imprecision on the x-axis.[11] On the MEDx chart, lines are also drawn representing different sigma zones, where the lower left-hand corner with low imprecision and low bias represents the area of 6σ "world class" performance; conversely, increases in either inaccuracy or imprecision corresponds to moving upwards or rightwards on the MEDx chart to the zones of lower sigma metrics.

RESULTS

The results of the sigma metric analysis in the "pre" and "post" periods for each analyte are presented below in Table 3. The average sigma metric is also calculated across all analytes in each time period. This analysis showed that some of the analytes saw a significant improvement (sigma metric increased by more than 1.0), some did not show a significant change, while other analytes saw a decrease in the sigma metric. Overall, the average sigma metric across all 25 parameters did see an improvement from 6.4σ in the "pre" period to 9.2σ in the "post" period. A paired-sample t test was performed, the one-tailed p-value is less than 0.001, indicating a statistically significant improvement has indeed occurred.

In Figure 1, a combined summary showing the sigma metric distribution of the 25 analytes is presented for both the "pre" and the "post" period, graphically showing the shift in the distribution. This chart demonstrates the dramatic shift in the distribution, where 9 analytes were performing at or below 3σ in the "pre" period, only 3 of the analytes remained at or below 3σ performance in the "post" period after the implementation of quality improvement actions.

Lastly, a few analytes with notable observations were selected and plotted on a normalized method decision (MEDx) chart to visualize and evaluate such changes. As shown in Chart 2a to 2f, the change in the performance of a particular analyte is reflected in the change in its position along the horizontal and vertical axes.

Table 3: Sigma metric scores in "pre" and "post" periods.

Test	Sigma (pre)	Sigma (post)	Δ sigma	Direct-ion
ALB	6.6	8.6	+2.0	↑
ALKP	8.6	13.0	+4.3	↑
ALT	5.0	4.8	-0.2	↔
AST	7.1	10.7	+3.6	↑
B12	5.3	3.4	-1.8	↓
CAL	3.7	5.6	+2.0	↑
CHOL	2.8	5.2	+2.3	↑
CL	2.4	4.6	+2.2	↑
CREA	3.8	7.5	+3.8	↑
DBIL	15.7	16.1	+0.4	↔
GGT	5.7	5.7	-0.0	↔
GLUC	3.9	26.3	+22.4	↑
HDL	19.6	17.1	-2.6	↓
K	2.6	14.3	+11.7	↑
NA	1.1	1.7	+0.7	↔
PHOS	5.0	6.9	+1.9	↑
TBIL	7.9	9.8	+1.9	↑
TPRO	6.9	5.8	-1.2	↓
TRIG	13.0	15.6	+2.5	↑
TSH	4.7	9.4	+4.8	↑
TT3	4.5	8.3	+3.8	↑
TT4	5.7	6.0	+0.3	↔
UREA	1.9	3.3	+1.4	↑
URIC	12.3	16.2	+3.8	↑
VITD	2.9	5.5	+2.6	↑
Aver-age	6.4	9.2	+2.9	(P<0.001)

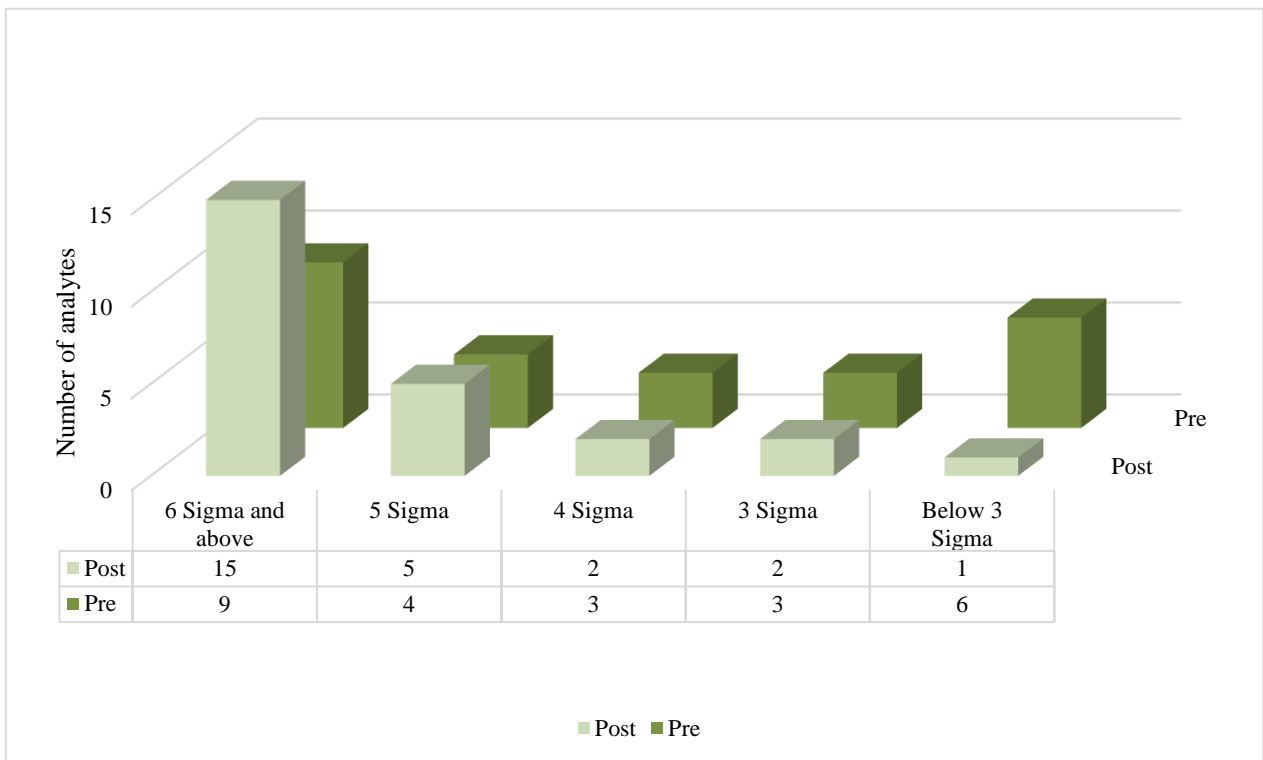|	6 Sigma and above	5 Sigma	4 Sigma	3 Sigma	Below 3 Sigma
Post	15	5	2	2	1
Pre	9	4	3	3	6

Figure 1: Summary of quality improvement across all 25 analytes.

DISCUSSION

The results in Table 3 demonstrate that even before the quality improvement initiatives were taken, our laboratory was already performing at a "world class" level, with an average of 6.4σ across the 25 clinical chemistry and immunoassay methods on the Alinity ci-series system. After the implementation of effective interventions in Table 2, we saw an impressive further increase in the average score to 9.2σ.

Furthermore, a detailed analyte-level examination of the results in Table 3 demonstrates that a few specific analytes exhibited notable changes. MEDx charts (Figure 2 a-f) provide further insight on the impact of imprecision and inaccuracy on overall sigma metric scores. On GLUC, K, and TSH, for example, we saw that the improvement from 2σ and 3σ to the 6σ zone was driven by a combination of both reduced inaccuracy and reduced imprecision (downward movement on the y-axis and leftward movement on the x-axis).
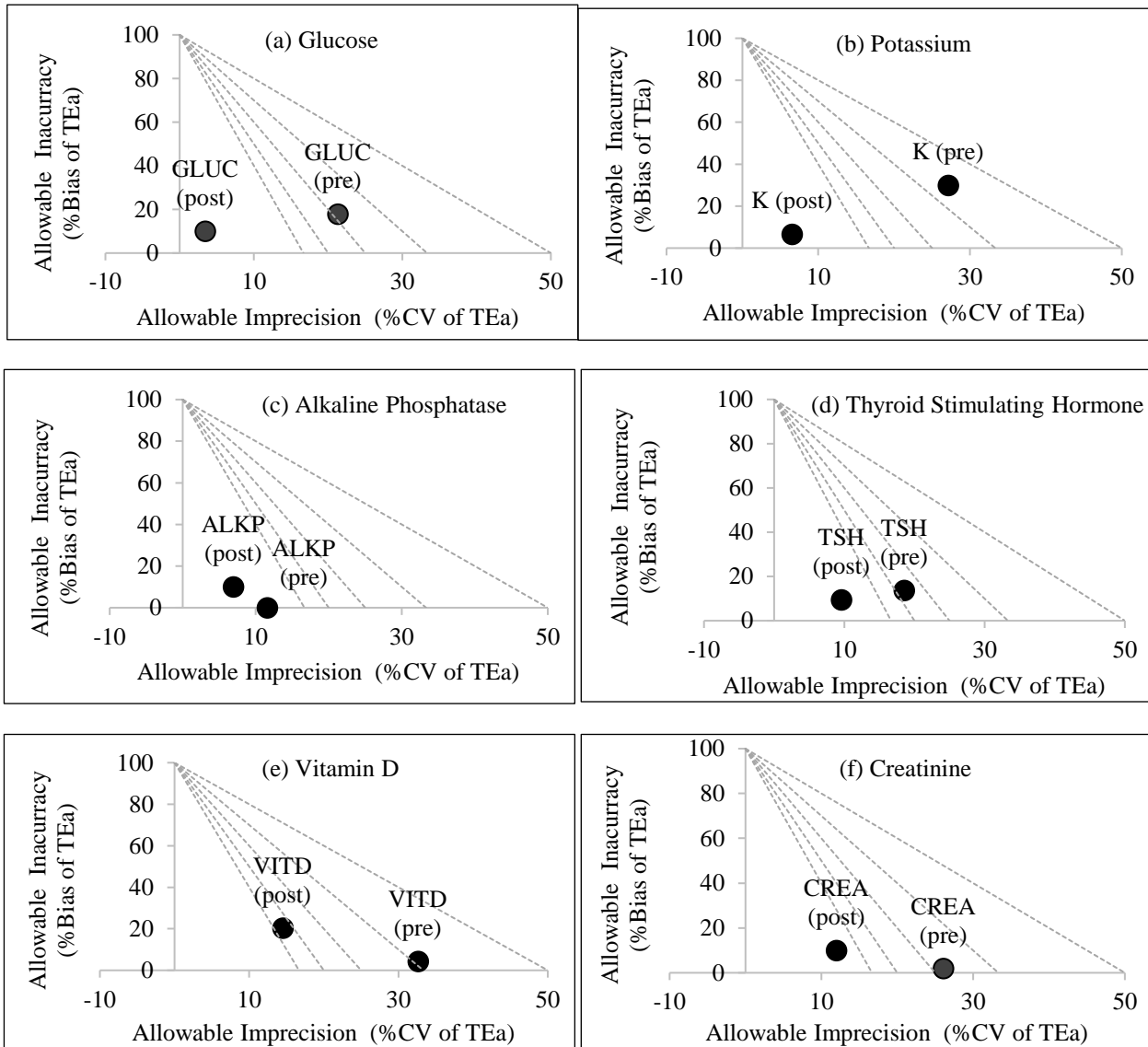

Figure 2: MEDx charts for selected analytes. a) glucose, b) potassium, c) alkaline phosphatase, d) thyroid stimulating hormone, e) vitamin d & f) creatinine.

The largest increase in the sigma metric was shown on GLUC, with a 22.4 point increase from 3.9σ to 26.3σ. Based on this, we can estimate the impact to the error rate on that particular analyte. The DPMO can be calculated from the sigma metric using the standard normal density function (where μ = 0), computed with Microsoft Excel using the NORMSDIST command.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Therefore, for GLUC, we can approximate that the error rate has reduced from 9,224 DPMO to less than 1 DPMO.

However, if we look at another analyte with a smaller sigma metric improvement, for sample VITD which increased only 2.6 points from 2.9σ to 5.5σ, we see that there was an even larger reduction of 75,635 DPMO (from 75,666 down to 31). Therefore, it is important to understand and remember the fact that the sigma scale has a logarithmic relationship with DPMO and thus the impact is much higher in the lower end, meaning that more gains can be made by making small improvements to poor processes than trying to make further improvements on processes that are already running above six sigma quality.

Beyond the magnitude of the increase that was observed in the average sigma value and the DPMO, the results also demonstrate that the improvement was broadly seen across majority of the analytes. In particular, this is reflected in the distribution of the sigma scores shown in Chart 1. On the low end, we see a big reduction in the number of marginally (3σ) and poorly (below 3σ) performing parameters from 9 in the "pre" period to only 3 in the "post" period. On the high end, we see an increase in the number of excellent (5σ) and world-class (6σ) performing assays from 13 in the "pre" period to 20 in the "post" period.

While previous research elsewhere has noted that the sigma values can undergo some fluctuations over time (6), we defined that a change in the sigma value of greater than 1.0 within the 6-month period of the study would be considered a significant change. Using this criterion, in-depth examination of each analyte in Table 3 showed that 17 of the analytes demonstrated a significant increase (GLUC, K, ALKP, TSH, URIC, CREA, AST, TRIG, CHOL, CL, ALB, CAL, TBIL, PHOS, VITD, TT3, and UREA). These analytes range across end-point photometry, kinetic rate, ion-specific electrode potentiometry, and chemiluminescence microparticle immunoassay methodologies-reflecting that the quality improvement is observed across the spectrum of testing. The breadth and magnitude of the quality improvement that we observed in our laboratory are both quite significant and cannot easily be disregarded as random fluctuations.

Five of the analytes (NA, DBIL, TT4, GGT and ALT) did not exhibit a significant sigma value difference. Sodium (NA) is an analyte which has a very tight TEa requirement. Direct bilirubin and total T4 were already performing at over 6σ, so absence of a significant improvement is not unexpected.

There were three analytes that showed a significant decrease in the sigma metric: Total protein (6.9σ to 5.8σ), vitamin B12 (5.3σ to 3.4σ), and HDL cholesterol (19.6σ to 17.1σ). It is worthwhile to note that while the changes greater than 1.0 are considered significant in our present definition, such changes especially for HDL and Total Protein did not lead to a huge increase in the DPMO since the analytes were still performing at "excellent"

(5σ) or "world class" (6σ and above) quality even after the decrease. This could be attributed to many possible explanations including variations in the performance of either the QC and EQA material, calibration, assay reagents, instrument hardware, or operator-related factors. The data does not provide further insights into why these particular analytes did not observe a similar improvement after our error reduction tactics were implemented.

It is important to keep in mind that while both the inaccuracy and the imprecision of each analyte came from actual data collected from within our laboratory, an important variable used in the calculation of the sigma metric comes from a pre-determined set of Total Allowable Error specifications. Despite enormous efforts at the 1999 Stockholm Consensus Conference and the 2014 Milan Consensus Conference, there is no set of standardized TEa specifications for all analytes yet which are universally adopted.[12,13] This makes it difficult for one laboratory to compare and benchmark its sigma metrics against others.

Nonetheless, the majority of the total allowable error specifications used in this study came from cited public sources such as CLIA which have been used in other sigma metric assessments conducted elsewhere around the world.[7,8] While it may still be challenging to make direct comparisons between our laboratory's performance versus other sigma metrics from other publications, this study's innovative use of six sigma tools such as the sigma metric equation, DPMO analysis and MEDx charts to track and guide quality improvement internally within our own institution has uncovered additional insight into the usefulness and power of six sigma methodology in a novel application.

Furthermore, different analyzer systems from different manufacturers have been demonstrated to have varying levels of analytical performance when we use the same evaluation criteria and calculate with the same set of Total Allowable Error specifications.[14] Therefore, while the one-time decision to select an analyzer system remains a prime determinant of the laboratory's level of performance, laboratory managers should also consider reliable day-to-day actions that can be implemented to make further continuous quality improvements.

This quality improvement study has brought multiple benefits to our laboratory. First, in objectively demonstrating that our laboratory is achieving world class performance across the highest volume clinical chemistry and immunoassay methods, we can enjoy an elevated level of confidence in the reliability of our results. Second, the quality improvements have enabled the majority of our analytes to reach 5σ or higher. This high level of analytical quality means that our quality control processes can be streamlined and simplified to use less QC runs and QC rules while still maintaining the same level of error detection and analytical quality

assurance as outlined by the Westgard Sigma Rules and CLSI C24 guideline recommendations.[15,16] Third, with improved and more efficient analytical processes, streamlined QC planning, and higher operator confidence, our laboratory reaps the additional benefits of less errors, less unnecessary reruns and troubleshooting, and better turnaround time for our clinicians.

Last and most important, the impressive quality improvement that we experienced in our laboratory came after much teamwork and close collaboration between our dedicated laboratory staff and the knowledgeable technical advisors from the manufacturer. Six sigma concepts and tools are not only useful in helping the laboratory to select the appropriate methods for each analyte, but can also be used to guide and monitor quality improvement initiatives over time.

CONCLUSION

The changes in sigma metric had a significant positive impact on the DPMO. It reinforced the reliability of our test results. It showed that our quality control processes can be streamlined and simplified further, to optimize the frequency of internal quality control, while still maintaining the same level of error detection and analytical quality assurance. The analysis also provided additional benefits of achieving lesser errors, fewer sample reruns and troubleshooting, and improved turnaround time, for better clinician and patient satisfaction.

ACKNOWLEDGEMENTS

REFERENCES

1. Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L. Evaluating laboratory performance on quality indcators with the six sigma scale. Arch Pathol Lab Med. 2000;(124): 516-519.
2. Westgard J. Six sigma quality design and control: Desirable precision and requisite QC for laboratory measurement processes Madison, WI: Westgard QC; 2001.
3. Flatland B, Breickner L, Fry M. Analytical performance of a dry chemistry analyzer designed for in-clinic use. Vet Clin Pathol. 2014;43(2):206-17.
4. Lenters-Westra E, English E. Understanding the Use of Sigma Metrics in Hemoglobin A1c Analysis. Clin Lab Med. 2017;37(1):57-71.
5. Farr A, Freeman K. Quality control validation, application of sigma metrics, and performance comparison between two biochemistry analyzers in a commercial veterinary laboratory. J Vet Diagn Invest. 2008;20(5):536-44.
6. Harrison H, Jones J. Using sigma quality control to verify and monitor performance in a multi-instrument, multi-site integrated health care network. Clin Laborat Medi. 2017;37(1):207-41.
7. Taher J, Cosme J, Renley B, Daghfal D, Yip P. A novel Sigma metric encompasses the global multi-site performance of 18 assays on the Abbott Alinity system. Clin Biochem. 2018;63:106-112.
8. Fasano T, Bedini J, Fle P, Jlaiel M, Hubbert K, Datta H, et al. Multi-site performance evaluation and Sigma metrics of 20 assays on the Atellica chemistry and immunoassay analyzers. Clin Chem Lab Med. 2019;58(1):59-68.
9. Carobene A, Aarsand A, Bartlett W, Coskun A, Diaz-Garzon J, Fernandex-Calle P, et al. The European Biological Variation Study (EuBIVAS): a summary report. Clin Chem Lab Med. 2022;60(4):505-17.
10. Ricos C, Alvarez V, Cava F, Garcia-Lario J, Hernandez A, Jimenez C, et al. Current databases on biologic variation: pros, cons and progress. Scand J Clin Lab Invest. 1999;59:491-500.
11. Westgard J. A method evaluation decision chart (MEDx chart) for judging method performance. Clin Lab Sci. 1995;8(5):277-83.
12. Hyltoft Petersen P, Fraser C, Kallner A, Kenny D. Strategies to set global analytical quality specifications in laboratory medicine. Scand J Clin Lab Invest. 1999;59(7).
13. Sandberg S, Fraser C, Horvath A, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med. 2015;53(6):833-5.
14. El Sharkawy R, Westgard S, Awad A, AbdelKarem Omneya A, El Hadidi I, 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 Med (Zagreb). 2018;28(2).
15. Westgard J, Westgard S. Basic quality management systems Madison WI; 2014.
16. Clinical Laboratory Standards Institute. Statistical quality control for quantitative measurement procedures: Principles and Definitions Wayne, PA; 2016.