Original Article

Plot3asic, three level QC data plotting and analysis freeware, for addressing cost of quality issues in diagnostic laboratories at resource constrained settings

Rajarshi Gupta

From, Assistant Professor, Department of Microbiology, Shri Ramkrishna Institute of Medical Sciences & Sanaka Hospitals, West Bengal, India.

Correspondence to: Dr. Rajarshi Gupta, Department of Microbiology, Shri Ramkrishna Institute of Medical Sciences & Sanaka Hospitals, Malandighi, Kanksha, Durgapur, West Bengal, India- 713212. Email: rguptamicrobiology@gmail.com

ABSTRACT

Objectives: In the rapidly evolving field of laboratory medicine, Quality Control (QC) is paramount in ensuring result accuracy and precision. However, QC procedures incur high cost of quality (COQ) which is often a hurdle to laboratories in resource constrained settings. Proprietary QC software is either expensive or tethered to proprietary platforms. Open source statistical libraries require considerable coding expertise often unavailable with laboratory personnel. Therefore, Plot3asic was developed as a freeware for generating publication quality QC plots with Westgard rule violation checks across three control levels. **Methods:** Plot3asic was developed using Python3 with 'plotly', an open source visualization toolbox. QC rules implemented by Plot3asic are adapted from Westgard rules. Mathematical calculations were independently verified by two investigators. Artificial datasets with known QC violations were analyzed using Plot3asic. Commercial QC material run data for the month of June 2019 were simultaneously analyzed on MultiXLv2017.01A(EM200) platform (Transasia/Erba Mannheim) and Plot3asic. **Results:** The results were generated as HTML files. The application was able to identify all QC violations as claimed. There was 100% corroboration between the violations flagged by MultiXLv2017.01A(EM200) platform and Plot3asic. The application also calculates sigma metric should the user choose to do so. **Conclusion:** To the best of our knowledge, Plot3asic provides features that no other freeware does till date. Determination of sigma metric helps streamline QC and reduce COQ. Plot3asic does not interface with proprietary platforms and is particularly helpful in semi-automated and manual assay data analysis in resource constrained settings.

Key words: analysis, freeware, plot3asic, quality, sigma

In the rapidly evolving field of laboratory medicine, the need for Quality Control (QC) is paramount. Robust QC ensures that the results generated by a laboratory are consistent on a day to day basis and that the results generated by one laboratory are comparable to those generated by other laboratories [1]. The QC system as described by Badrick in 2008 encompasses an understanding of analytical error, use of QC materials and a set of QC rules to decide whether a run is in or out of control and the processes to follow if the run is out of control [2]. Though QC procedures exist for both qualitative and quantitative testing systems yet a majority of QC rules pertain to the interpretation of quantitative test results. Westgard multi-rule procedure is the most commonly used system wherein different control rules are combined to increase the probability of error detection for a particular level of false rejection [3, 4]. It is imperative that all laboratories, basic and advanced, in developing or in developed countries implement these QC rules to improve the quality of results generated. In recent years, Six Sigma Metrics have become increasingly popular in all sectors including the health-care sector [5]. Since

laboratory generated results are vital in clinical decision making, laboratory performance for each test parameter should be evaluated using sigma metrics. However, during thorough literature search, we found that majority of software applications that guide a laboratory physician in implementation of the multi-rule QC are proprietary and are interfaced to their respective proprietary testing platforms. Free applications that are available do not cover a comprehensive gamut of metrics required for interpretation of QC data. Statistical software packages that can generate plots and determine QC violations require considerable amount of coding expertise which is not usually available with all laboratory physicians.

Plot3asic, a simple freeware application conceived to address these issues, allows users to instantly generate publication quality data plots with QC rule violation flags and sigma metrics. Plot3asic source code and the Plot3asic logo are both licensed under the Creative Commons Attribution – NonCommercial - NoDerivatives 4.0 International Public License. The application accepts QC data in a specific Microsoft Excel spreadsheet (.xlsx) file which is bundled with the application. QC data is recorded daily by the user on the spreadsheet file and analyzed instantly to generate QC data plots with custom choice of control rule combinations. The application does not interface with any particular proprietary system. Data from any testing platform can be entered manually in the spreadsheet and analyzed using Plot3asic.

MATERIAL AND METHODS

The application was developed using Python 3 programming language. For data plot generation the application uses package 'plotly' which is an open source graphing library and data visualization toolbox for Python. The QC data file and the Controls data file which the users are required to use, are both Microsoft Excel spreadsheets (.xlsx format). The application is cross platform and can run on any system where it is possible to install Python 3 (tested on Windows 10 and Linux based OS such as Ubuntu, Debian, openSuse and CentOS). The application requires some basic platform specific setup before it can be used to generate plots. The setup instructions are made available with the application.

The predefined QC data file which the users are meant to populate with their test values is shown in Figure 1A. Each QC data file is designed to contain data for all test parameters for a particular month. Each of the sheets of the QC data file is meant for a single test parameter. The "Controls" data file contains details of the control materials being used (Figure 1B). The details of how to use these files and how to use the application in general are mentioned in the help document that is available with the application. A sample QC data source file and a Controls data file are available bundled with the application, for the users' reference.
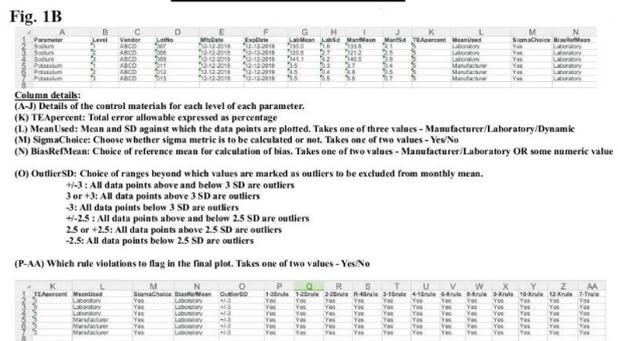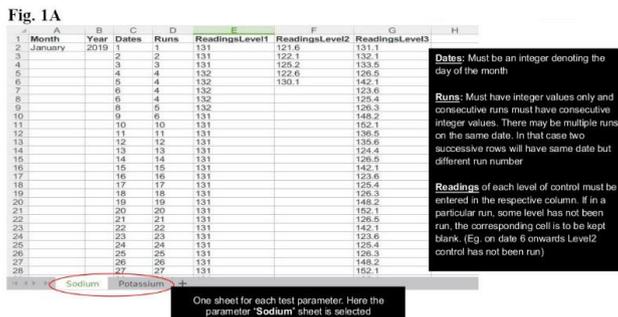
Figure 1: [Original] A: The QC data source file B: The Controls data file

In brief, the QC data file and the application can handle three levels of controls for each parameter (Figure 2). The 'Dates' column takes integer values for days of the month. The 'Runs' column takes integer values for identifying a particular analytical run. Consecutive runs must have consecutive integer values. An analytical run is "for purposes of QC, an interval, that is, a period of time or series of measurements, within which the accuracy and precision of the measuring system is expected to be stable" [2, 6]. This indicates that there may be multiple runs for a particular observation date (see observation dates 6 in Figure 2). Each run generates a certain value for Level1, Level2 and Level3 controls. These values are entered in their respective columns ('ReadingsLevel1', 'ReadingsLevel2' and 'ReadingsLevel3'). If a particular run involves testing of only one level of control, the appropriate cell for the other levels is then left blank (see date 6 onwards in Figure 2 where only Level1 and Level3 controls have been tested). The "Controls" data file contains data on the control materials such as vendor names, lot numbers, manufacturer provided means and standard deviations, laboratory derived means and standard deviations and total allowable error percentages. It also contains data on mean against which points are plotted, choice of sigma metric calculation, reference mean to be used for bias calculation, definition of outliers and

individual rule violations that need to be flagged for each level of control (Figure 1B).

Figure 2: Sample QC data sheet for parameter "Sodium" [Original]

After uploading the QC data file and the Controls data file to the application interface (Figure 3), data plots with summary statistics (mean, standard deviation, Coefficient of variation or CV%), various performance metrics (bias, total error, sigma) and QC rule violation flags can be generated using any combination of rules (see below). From the control data file, user can additionally choose to calculate sigma metrics at the end of each month and can even specify which values to mark as outliers and exclude from monthly mean, coefficient of variation (CV) and bias calculations. As an advanced function, users are even allowed to modify outliers.

The QC rules implemented by Plot3asic (stated below) are adapted from Westgard rules [7]. Plot3asic can flag QC violations both within and across three levels of controls.

- 1-3S: One point beyond 3 standard deviations
- 1-2S: One point beyond 2 standard deviations
- 2-2S: Two points consecutively beyond 2 standard deviations on same side of the mean OR two points (in a particular run) beyond 2 standard deviations on same side of the mean (within level and across levels)
- R-4S: Two points consecutively(within run), beyond 2 standard deviations on opposite sides of the mean (within level or across levels)
- 3-1S: Three points consecutively beyond 1 standard

deviation on same side of the mean OR three points (in a particular run) beyond 1 standard deviation on same side of the mean(within level and across levels)

- 4-1S: Four points consecutively beyond 1 standard deviation on same side of the mean OR four points (in a particular run) beyond 1 standard deviation on same side of the mean(within level and across levels)
- 6-X: Six points consecutively on same side of the mean OR six points (in a particular run) on same side of the mean (within level and across levels)
- 8-X: Eight points consecutively on same side of the mean OR eight points (in a particular run) on same side of the mean (within level and across levels)
- 9-X: Nine points consecutively on same side of the mean OR nine points (in a particular run) on same side of the mean (within level and across levels)
- 10-X: Ten points consecutively on same side of the mean OR ten points (in a particular run) on same side of the mean (within level and across levels)
- 12-X: Twelve points consecutively on same side of the mean OR twelve points (in a particular run) on same side of the mean (within level and across levels)
- 7-T: Seven points consecutively showing increasing or decreasing trend (within level and across levels)
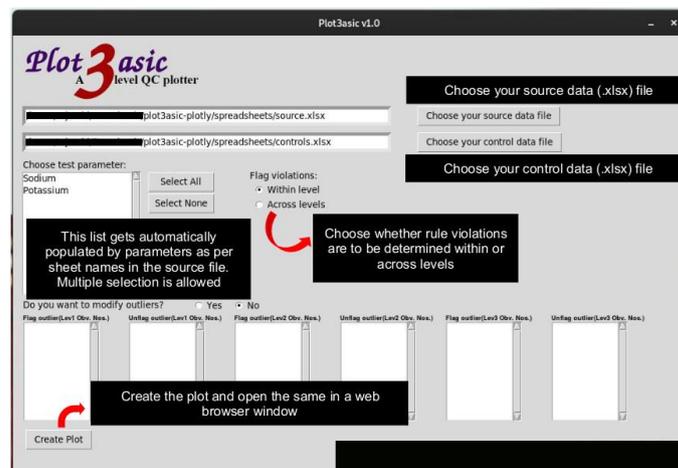
Figure 3: The application interface [Original]

Mean and standard deviation were calculated using the following formulae [8].

$$\text{Mean } (\bar{x}) = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (SD) = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x} - x_i)^2}{n-1}}$$

Bias, CV, total error and sigma metric are calculated using the following formulae [9–11].

Bias% = (Average absolute deviation from the target value/Target) × 100 or, Bias% = (|Monthly mean −

Reference Mean|)/Reference Mean x 100

Coefficient of variation (CV) = SD/Mean, CV% = (SD/Mean) x 100, Total error% = 1.65 × CV (%) + Bias (%), Sigma metric = (TEa − |Bias%|) / CV%

(TEa – allowable total error which the user decides as per prevailing industry standards and records in the appropriate cell of the "Controls" sheet)

Artificial datasets and commercial QC material run data have been used instead of actual patient sample data to test various aspects of the software. Since the tests parameters for which this application is meant to be used are all numeric, any artificial dataset or commercial QC material run data can actually simulate the results of analysis of real life values. Further, it is highly unlikely that real life QC run data sets will contain all QC rule violations. Hence both artificial datasets and commercial QC material run data have been used to check the validity of each rule violation implemented by the application.

Table1: The artificial dataset (a summary) [Original]

Month: January | Year: 2019
Level of control: 1 | Mean: 130.0 | SD: 0.5
Level of control: 2 | Mean: 125.0 | SD: 0.5
Level of control: 3 | Mean: 140.0 | SD: 0.5

Dates	Runs	ReadingsLevel1	ReadingsLevel2	ReadingsLevel3
1	1	130.3	125.1	139.4
2	2	130.5	124.8	139.5
3	3	130.6	126.4	139.6
3	3	129.5	123.6	139.7
5	4	130.5	125.1	139.8
6	5	130.4		139.9
6	6	129.6		140.0
8	7	129.8		140.1
9	8	132.0		139.8
10	9	130.3		140.2
11	10	130.5		140.3
12	11	130.6		139.7
13	12	129.5		139.9
14	13	130.5		140.1
15	14	130.4		140.2
16	15	129.6		140.1
17	16	129.8		140.3
18	17	131.2		140.4
19	18	131.3		140.4
20	19	130.3		140.3
21	20	130.5		140.2
22	21	130.6		140.2
23	22	129.5		140.3
24	23	130.5		140.1
25	24	130.4		140.2
26	25	129.6		139.8
27	26	130.6		140.2
28	27	130.7		140.1
29	28	130.7		139.9
30	29	130.6		140.2
31	30	130.2		139.9

A summary of the artificial dataset used is shown in Table 1. It is summarized from one of the many QC data

source files and a Controls data files used to test the application. This particular dataset illustrates the capability of Plot3asic using the parameter "Sodium". It contains all rule violations that Plot3asic is supposed to flag.

To further demonstrate the practical utility of this application, commercial (Transasia / Erba Mannheim) QC material (Lot No. B071836) for parameters "serum bilirubin (direct)" and "serum alanine aminotranferase" were run on Transasia / Erba Mannheim EM 200 Automated Clinical Chemistry Analyzer during June 2019. The results were analyzed on both the proprietary MultiXL v2017.01A (EM200) platform (Transasia / Erba Mannheim) and Plot3asic. Since a newly opened batch of control material was used, the analysis was made against the manufacturer provided mean and standard deviation. The implementation of mathematical calculations has been independently verified by two investigators.

RESULTS:

For artificial dataset, Plot3asic was set to identify all available QC rule violations. The results illustrated in Figure 4 are generated as HTML files which open automatically in the default web browser. Plot3asic has been able to flag all the violations that the artificial dataset contains. The points where rules have been violated are marked in red while others are marked in blue. The line only touches those points which are not designated as outliers. As illustrated in Figure 4, the application also provides the value of sigma metric if the user chooses so. Figure 5 further illustrates the extended capability of this application where the user can individually choose to flag or unflag outliers at individual data points. Modifying outliers is an advanced feature. If outliers are changed so does the monthly mean.
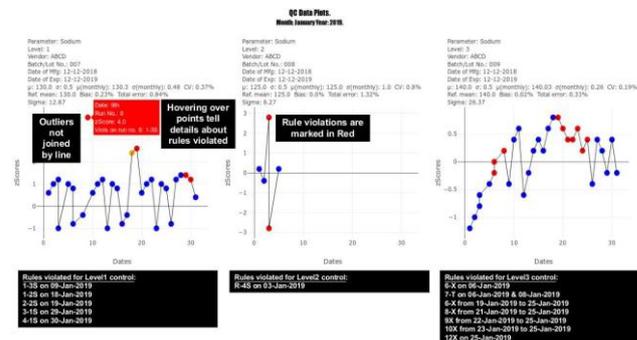
Figure 4: Sample plot with QC violation data (artificial dataset) [Original]

For commercial QC material, both the data analysis applications were set to identify these QC rules; 1-3S, 1-

2S, 2-2S, 4-1S and 10X. The result for "serum bilirubin (direct)" has been summarized in Table 2. Both the applications flagged 1-3S violation on 07-Jun-2019 and 4-1S violation on 08-Jun-2019. The result for "serum alanine aminotransferase" has been summarized in Table 3. Both the applications flagged 1-2S violation on 02-Jun-2019 and 18-Jun-2019; 2-2S violation on 03-Jun-2019 and 04-Jun-2019; 4-1S violation on 05-Jun-2019, 06-Jun-2019 and 08-Jun-2019. There is 100% corroboration between the rule violations flagged by the two platforms.
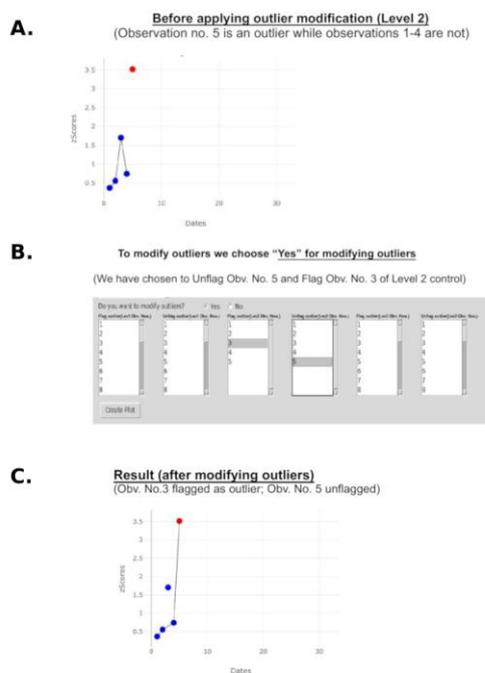
Figure 5: Manual flagging and unflagging of outliers [Original]

DISCUSSION

To the best of our knowledge Plot3asic being a free application provides features that no other freeware does till date. In provides a framework to directly read QC data from Microsoft Excel files which the user may employ to store QC data. There is no need to copy and paste data into the application interface for every instance of analysis. It is successfully able to flag a gamut of QC rule violations that are required for day to day monitoring of laboratory performance. By allowing users to easily calculate sigma metric, Plot3asic assists users in gauging their laboratory performance as per current industry benchmarks (namely the six sigma scale) [9, 10, 12]. Determination of sigma metric further allows the users to streamline their QC requirements according to Westgard sigma rules thus

enabling lower cost of quality (COQ).[12] Plot3asic urges users to employ the cost of quality (COQ) approach in determining best quality control practices for their laboratories.

Table2: Comparison of analysis of serum bilirubin (direct) results on MultiXL v2017.01A [EM200] platform (Transasia / Erba Mannheim) and Plot3asic [Original]

Month: June | Year: 2019
Level of control: 1 | Mean: 0.597 | SD: 0.06

Dates	Runs	Readings Level1	Violations flagged (MultiXL v2017.01A)	Violations flagged (Plot3asic)
2	1	0.59		
2	2	0.61		
3	3	0.6		
4	4	0.58		
4	5	0.69		
5	6	0.68		
6	7	0.66		
7	8	1.42	1-3S	1-3S
8	9	0.66	4-1S	4-1S
9	10	0.61		
10	11	0.61		
11	12	0.6		
12	13	0.57		
13	14	0.6		
14	15	0.59		
15	16	0.6		
16	17	0.58		
17	18	0.55		
17	19	0.49		
18	20	0.49		
19	21	0.55		
20	22	0.57		
21	23	0.61		
22	24	0.64		
23	25	0.55		
24	26	0.61		
25	27	0.58		
26	28	0.59		
27	29	0.62		
28	30	0.61		
29	31	0.61		
30	32	0.59		

Elbireer et al have in a resource constrained setting estimated that approximately 32% of total laboratory expenses were spent on creating and following processes that ensure good laboratory practices (GLP) and contribute to high-quality outcomes. About 70% of the total COQ was estimated to be the cost of quality appraisal [13]. Plot3asic does not interface with any proprietary testing platform and is thus particularly helpful in analysis of QC data generated from semi-automated and manual assays that are common in resource constrained settings. Quality is a never ending process. Plot3asic flags all probable QC violations as per aforementioned rules. What rules the user chooses to implement, depends on the level of stringency in Quality control he/she requires. The final record is generated in the form of a Hypertext Markup Language

(HTML) file with plots and QC rule violation information. Space for additional comments has been provided so that the user can record additional statements if necessary. Plot3asic source code and the Plot3asic logo are both licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License. Plot3asic will always be free and open to modifications based on user feedback.

Table 3: Comparison of analysis of serum alanine aminotransferase results on MultiXL v2017.01A [EM200] platform and Plot3asic [Original]

Month: June | Year: 2019
Level of control: 1 | Mean: 38.007 | SD: 3.48

Date	Runs	Readings Level1	Violations flagged (MultiXLv2017.01 A)	Violations flagged (Plot3asic)
2	1	42.0		
2	2	45.4	1-2S	1-2S
3	3	45.2	2-2S	2-2S
4	4	47.5	2-2S	2-2S
4	5	44.0	4-1S	4-1S
5	6	44.4	4-1S	4-1S
6	7	43.7	4-1S	4-1S
7	8	155.4	1-3S	1-3S
8	9	42.5	4-1S	4-1S
9	10	37.1		
10	11	38.2		
11	12	38.1		
12	13	37.0		
13	14	38.1		
14	15	36.3		
15	16	37.4		
16	17	38.6		
17	18	36.3		
18	19	29.8	1-2S	1-2S
19	20	35.0		
20	21	38.9		
21	22	39.8		
22	23	41.7		
23	24	35.9		
24	25	42.5		
25	26	40.0		
26	27	39.7		
27	28	41.3		
28	29	35.3		
28	30	37.2		
29	31	35.0		
30	32	39.0		

CONCLUSION

Following the creative vision of 'Resource constraint must not be a hurdle to QC implementation, Plot3asic will always be free and open to user feedback. To the best of our knowledge, Plot3asic provides features that no other freeware does till date. Determination of sigma metric helps streamline QC and reduce COQ. Plot3asic does not interface with proprietary platforms and is particularly helpful in semi-automated and manual assay data analysis in resource constrained settings.

REFERENCES

1. Pacey AA. Quality assurance and quality control in the laboratory andrology. Asian J Androl. 2010;12:21–5.
2. Badrick T. The quality control system. Clin Biochem Rev. 2008;29 Suppl 1:S67–70.
3. Harel O, Schisterman EF, Vexler A, et al. Monitoring quality control: can we get better data? Epidemiology. 2008;19:621–7.
4. Badrick T. Quality leadership and quality control. Clin Biochem Rev. 2003;24:81–93.
5. Mao X, Shao J, Zhang B, et al. Evaluating analytical quality in clinical biochemistry laboratory using Six Sigma. Biochem medica. 2018;28:20904.
6. Westgard JO BP. Cost-effective quality control: managing the quality and productivity of analytical processes. Washington (DC): AACC Press; 1986.
7. Westgard J. "Westgard Rules" and Multirules - Westgard (Internet). Westgard QC. 2009 (cited 2018 Oct 24). Available from: https://www.westgard.com/mltirule.htm
8. Lee DK, In J, Lee S. Standard deviation and standard error of the mean. Korean J Anesthesiol. 2015;68:220–3.
9. Westgard S, Bayat H, Westgard JO. Analytical Sigma metrics: A review of Six Sigma implementation tools for medical laboratories. Biochem medica. 2018;28:20502.
10. Guo X, Zhang T, Gao X, et al. Sigma metrics for assessing the analytical quality of clinical chemistry assays: a comparison of two approaches. Biochem medica. 2018;28:20708.
11. Biswas SS, Bindra M, Jain V, et al. Evaluation of imprecision, bias and total error of clinical chemistry analysers. Indian J Clin Biochem. 2015;30:104–8.
12. Westgard JO, Westgard SA. Quality control review: implementing a scientifically based quality control system. Ann. Clin. Biochem. 2016.
13. Elbireer A, Jackson JB, Gable AR. Cost of Quality at a Clinical Laboratory in a Resource-Limited Country. Lab Med. 2010;41:429–33.