Intelligent Quality Management 2 with IntraSpect™ technology for quality control of GEM® Premier™ 5000 blood gas analyzers— A novel application of the patient sample as its own control

James O. Westgard [a,b], Jose Cervera [c,*]

[a] Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medical and Public Health, Madison, WI, USA
[b] President, Westgard QC, Inc., Madison, WI, 53717, USA
[c] R&D Manager – Scientist, Werfen, 180 Hartwell Rd, Bedford, MA, 01730, USA

ARTICLE INFO

ABSTRACT

Background: Quality control (QC) in point-of-care (POC) testing has been greatly improved by automatic control processes, such as the Intelligent Quality Management (iQM®) technology found in GEM Premier blood gas analyzers (Werfen, Bedford, MA). The 2nd generation technology, iQM2, provides additional capabilities, notably the incorporation of IntraSpect software that monitors the response curves of individual tests to detect transient errors caused by micro-clots, micro-bubbles or any event that disturbs the sensor response during sample data acquisition. IntraSpect is a novel form of patient-based, real-time quality control (PBRTQC).
Methods: IntraSpect pattern recognition software monitors the last 15 measurements of each patient-response curve. Control limits for slope coefficients have been established from theoretical models and empirical data. Abnormal measurement behavior is flagged to identify transient errors that invalidate test results. Results from 1,013,391 patient samples were collected on 4,985 GEM Premier 5000 cartridges and 2,765 instruments in clinical use worldwide.
Results and conclusions: Total pre-analytic and transient errors detected by IntraSpect were 1.91% worldwide. iQM2 with IntraSpect technology provides a unique control function detecting transient errors that would otherwise go undetected with traditional QC. Together with the statistical QC technology in iQM2, pre-analytic, analytic, and transient analytic errors are detected much faster—seconds versus hours—than by traditional statistical QC.

1. Introduction

A recent survey of point-of-care (POC) clinicians asked participants to identify the "practices and wants" that would be most helpful for improving the quality of POC testing [1]. The responses prioritized the following activities: more manufacturer-integrated quality and function checks; more in-house training by in-house staff; more training from manufacturers; more supervisor oversight; and more explicit directions from manufacturers. The need for manufacturers to support quality assurance and training activities is clearly

---

critical for achieving high quality POC testing.

The strategic direction for manufacturers should be to provide more integrated controls and a more complete system for monitoring all potential failure modes. These capabilities would reduce the need for operator training and supervision of operations. An example of such a quality system is Werfen's iQM2, the second-generation Intelligent Quality Management system, for the GEM Premier 5000 blood gas analyzer. Nichols et al. [2] recently reported an extensive validation study of iQM2. The methodology for validating detection time was described in the validation of the first generation iQM, approximately 15 years ago [3]. Notably, the average times for detecting errors for each test are much quicker as compared with traditional statistical quality control (SQC) practices. This improved performance has also been demonstrated in field studies in four different institutions, as reported by Toffaletti et al. [4], in two different institutions, as reported by Mion et al. [5], as well as the more recent study in four different institutions, as reported by Nichols et al. [2].

A major improvement in iQM2 is the IntraSpect technology that monitors the measurement response curve of each patient sample. In addition to the analysis of Process Control Solutions (PCSs) that infer the quality of the analytical system before and after sample process, IntraSpect directly detects potential problems for each patient sample within the measurement cycle. During the sample measurement process, IntraSpect collects a series of readings which are then inspected using pattern recognition software to identify unusual sensor behavior due to transient errors. Transient errors are events that happened within the sample measurement due to sample preparation and presentation (e.g., micro-clots, micro-bubbles, interferences). Such errors can lack a recognizable signal, before or after sampling, and are random in nature – making them typically undetectable by traditional QC. Micro-clots and micro-bubbles are particular problems for systems that utilize whole blood samples and are particularly difficult to monitor via traditional, liquid control methods [6].

IntraSpect is a novel form of patient based real-time QC [PBRTQC] with integrated QC for samples. PBRTQC encompasses all types of patient-based control techniques including delta checks for individual patients, limit checks for individual samples, cross-check algorithms, such as anion gap, as well as a wide variety of population statistics such as Average of Normals (AoN). The use of PBRTQC is recommended by the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Committee on Analytical Quality [7]. This committee has stated that "*PBRTQC will become the mainstay of QC in laboratories once the profession sees the advantages of this form of process control, and manufacturers and middleware vendors provide the onboard capability*" [8]. In that context, iQM2 provides an example of a manufacturer's implementation of a PBRTQC technique, along with automation of traditional SQC, and extensive system function checks. The application and performance of IntraSpect technology is described in this report.

2. Methods and materials

2.1 The GEM Premier 5000 system (Werfen, Bedford, MA) is a critical care analyzer that utilizes whole blood to perform tests for blood pH, pCO_2, pO_2, Na^+, K^+, Cl^-, Ca^{++}, glucose, lactate, total bilirubin, hematocrit, total hemoglobin, oxy-hemoglobin, deoxy-hemoglobin, oxygen saturation, carboxy-hemoglobin, and met-hemoglobin. The main components of the GEM Premier 5000 system are a multi-use disposable cartridge (GEM PAK), iQM2 active process control technology, and GEMweb® Plus Custom Connectivity, an automated web-based information management system. The GEM PAK includes all the required analytical components (sensors, solutions, sampler, optical cell, waste bag). iQM2 provides continuous QC monitoring before, during, and after sample measurement, analyzes five PCSs to monitor all the critical decision concentrations, and provides automatic error detection and correction of the
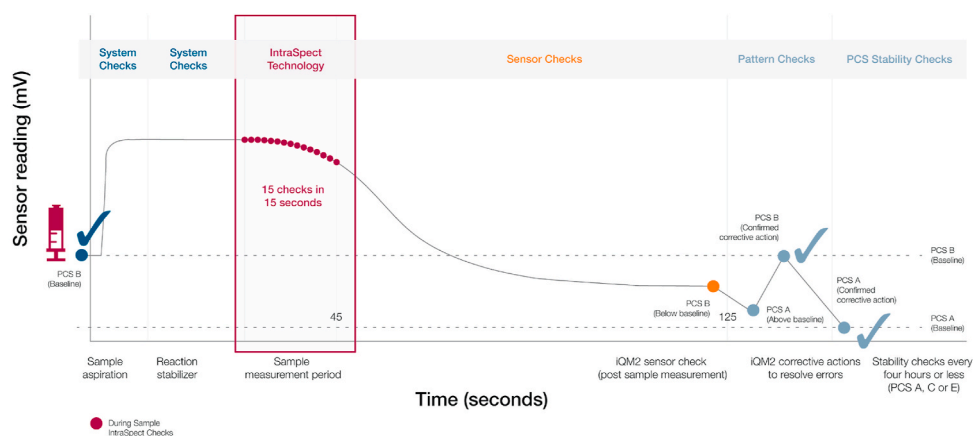

Fig. 1. Description of measurement cycle for GEM Premier 5000 operation and identification of various system checks, IntraSpect monitor, and sensor checks.

The figure summarizes iQM2 Quality Control processes for each sample - before, during, and after - each sample measurement process on the GEM Premier 5000 system. Before - system and sensor functionality checks; During – system fluidic checks in aspiration and InstraSpect; After-sensor pattern checks to ensure sensor recovery after sample washing step and ready for next sample. Corrective action is automatically triggered when a sensor recovery after sample is detected by the iQM2 pattern recognition.

system, complete with documentation of corrective actions.

2.2 Automated SQC was one of the breakthroughs with iQM. Analytic errors could be detected very quickly, as described in an early validation study [3], due to the unique application of SQC. Electronic drift limits, imposed on the sensors for various PCSs, correspond to statistical control limits that can be translated as individual control rules. Error detection capabilities can then be determined from theoretical power curves that describe the probabilities of rejection, as a function of the size of errors occurring. The size of the systematic error that is medically important can be calculated from the allowable total error for a test (e.g., CLIA criteria for acceptable performance) and the precision observed for the methods, which also determine the Sigma-metric (SM) for the test. The SMs for the GEM Premier 5000 with iQM2 are documented in the recent validation report by Nichols et al. [2]. The critical systematic error (ΔSE_{crit}) is related to the Sigma-metric by the equation SM = $\Delta SEcrit$ + 1.65. Given an SM of 6, it follows that the medically important SE is 4.35 times the observed SD. The probability of detecting medically important errors (P_{ed}, probability of error detection) follows the methodology described originally by Westgard et al. [3]. Those probabilities can then be converted into average run lengths (ARL) to characterize the expected time for detection of medically important errors (ARL$_{ed}$ = 1/P_{ed}). Given the frequency of PCS analysis, ARL$_{ed}$ can then be used to estimate the average time for detection of medically important errors. Obviously, the best QC performance will be obtained with the PCS tested most frequently, which is PCS B, analyzed every 2 min. For iQM2, Nichols et al. [2] have characterized the average detection times (ADT) using this approach as 2 min for most analytes, except sodium and glucose which require five and 21 min, respectively.

2.3 The iQM2 Total Quality Process begins with the installation of the GEM PAK on a GEM Premier 5000 system, first validating performance using the five PCSs, then monitoring patient sample analysis, using specific quality checks. These include mechanical, electronic, and fluidic checks, temperature checks, and sensor/CO-Ox checks using the PCSs to monitor change, along with pattern recognition to identify errors and interferences. IntraSpect technology provides real-time monitoring of each patient sample during the measurement cycle to provide additional information about the quality of measurement during the analysis of each patient sample. The process for employing all these checks is illustrated in Fig. 1.

2.4 IntraSpect technology utilizes 15 readings during the last 15 s of patient sample measurements, to monitor the response of the sensor output. Pattern recognition software analyzes the readings for sensor response prior to reporting the patient result. The sample response, in millivolts readings, is fitted to a regression model. The regression equation characterizes the observed sample output and is used to estimate the value of the quantity and to identify abnormal sensor response. The behavior of measurement response curves for normal patient samples was characterized and established from historical clinical sample data collected from GEM Premier systems in the field. Unusual sensor responses are identified by the response shape and coefficients outside of acceptable limits.

Fig. 2 illustrates the expected response of a potassium sensor. The three lower responses represent the stable signals usually observed during the measurement process. The top response curve provides an example of a sample where the signal was disturbed by a factor significantly affecting its behavior – in this case due to a fluidic blockage during the sample aspiration, potentially caused by a blood clot in the sample. The affected sample reported an elevated potassium of 6.8 mmol/L when a value of approximately 4.0 mmol/L was expected. The expected value is based on the initial sensor output response and a retrospective sample analysis run in that same system.

As shown in Fig. 2, the coefficients that, in this case, represent the slope of the signal, are close to zero for normal response curves.
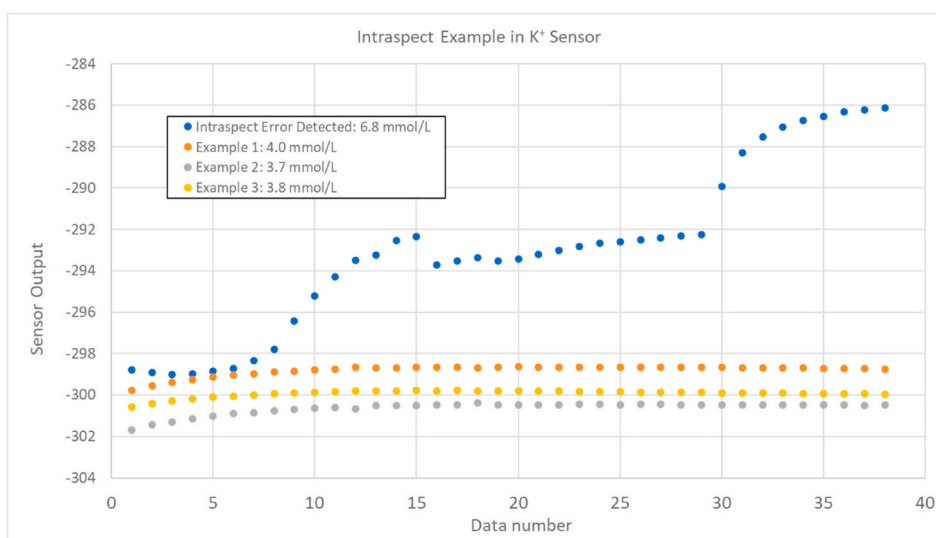

Fig. 2. IntraSpect example response curves for K$^+$ sensor.
Example of a sample affected by a transient event and detected by iQM2. The sensor signal is clearly a different pattern and flagged by IntraSpect (blue), as compared to sample signals of unaffected samples (red, grey and orange). . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Once the nature of the responses for different sensors was determined, a library of coefficients was developed from several thousand samples of known concentrations. This library was then used to optimize the limits for the IntraSpect coefficients for each sensor. During this process, samples were manipulated to cover the reported range of the system for each measured analyte, as well as several known clinical conditions and other known preanalytical conditions (e.g., microbubbles, blood clots, interfering substances). The coefficients were tested and optimized, using data collected in several clinical locations, during the last steps of system development.

Fig. 3 shows the IntraSpect limits for potassium, where most of the coefficient values (slope) are close to zero, and the limits show only minor changes with the analyte concentration. This illustrates the model of "constant" limits that also applies to sodium, chloride, and calcium, though with some empirical modifications, based on the particular analyte, as shown in Fig. 4, parts A, B, and C, respectively.

Fig. 4, part D shows the IntraSpect limits for glucose, where the limits vary with the measured glucose concentration. This "proportional" model applies to sensors with a diffusion component and also applies to lactate (part E). The curvature of the signal is related to the sample concentration. In addition, for glucose and lactate, the signals of these sensors are expected to increase during the sampling process until reaching a plateau (a flat area of the curve) that occurs when an equilibrium between the analyte concentration and the diffusion characteristics of the outer membrane has been achieved (besides other factors). There is no (expected) scenario where the sensor should show a significant decrease in signal during the sampling process. In fact, that event would indicate clear misconduct of the sensor performance, with potential for significantly impacting the analytical result.

Part F of Fig. 4 illustrates the IntraSpect limits for pO_2. As expected on a sensor governed by a diffusion process, the signal is evolving during the data collection time and, therefore, a curved-based equation is used. These limits are influenced by the effects of the concentration of the measured analyte, with respect to the concentration of the same analyte in the PCS before the sample was run (in the case of pO_2, 180 mmHg on PCS B). Therefore, the signal generated by samples with lower pO_2 than PCS B, will create a downward-trend curve generating a coefficient lower than zero, while the opposite is expected for samples with pO_2 higher than PCS B. In addition, more extreme coefficient values are expected at more extreme concentrations, leading to non-linear limits, as shown in Fig. 4. For completeness, Fig. 4 also provides the IntraSpect limits for pH (part G) and pCO_2 (part H).

Once the control limits were established for the different sensors, the performance of iQM2 with IntraSpect technology was assessed from data collected on 4,985 GEM PAKs and 2,765 systems in clinical use worldwide. The control limits were generally less than half of the CLIA-allowable total error for the individual test, and data outside these limits showed that the errors exceeded the CLIA limits in 70% of the cases, as demonstrated by Mansouri et al. [9]. This assessment considered the errors detected on 1,013,391 patient samples tested in multiple settings.

3. Results and discussion

Table 1 summarizes the categories of errors, the errors detected by iQM2, and the prevalence of those errors. Total pre-analytic and transient errors detected by IntraSpect were 1.91% worldwide. Of the total errors observed, approximately half were transient errors detected by IntraSpect (0.80%).

These errors correlated strongly with the detection of micro-clots caused by improper mixing and/or improper anticoagulant. Interferences due to benzalkonium chloride, thiopental, and optical problems from turbidity were observed to be smaller, but still significant, given the 1,013,391 total patient samples in this study.

The prevalence of the errors for the different tests is shown in Table 2. Chloride was highest at 0.39%, followed by pCO_2 at 0.15% and ionized calcium at 0.14%. Due to the numerous factors that could trigger an IntraSpect error detection, such as fluidic and chemical factors, sensor membrane physical properties, or sample properties, the cause of these differences in error prevalence among the sensors are difficult to distinguish.
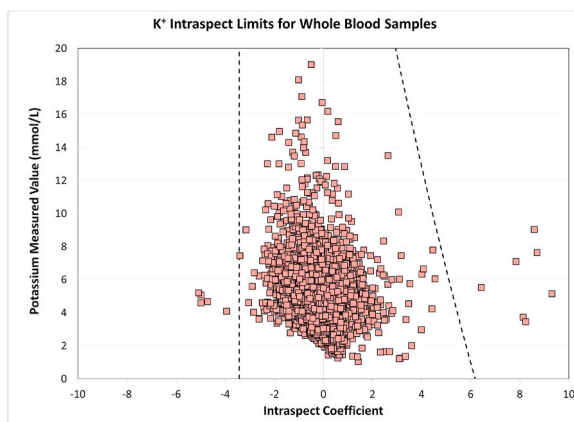

Fig. 3. Distribution of slope coefficients for K$^+$ sensor along with empirical control limits.
Representation of IntraSpect coefficients values against the analyte concentration and the associated thresholds (dotted lines) for flagging transient errors. Samples within the dotted lines are acceptable and outside are flagged or suppressed.
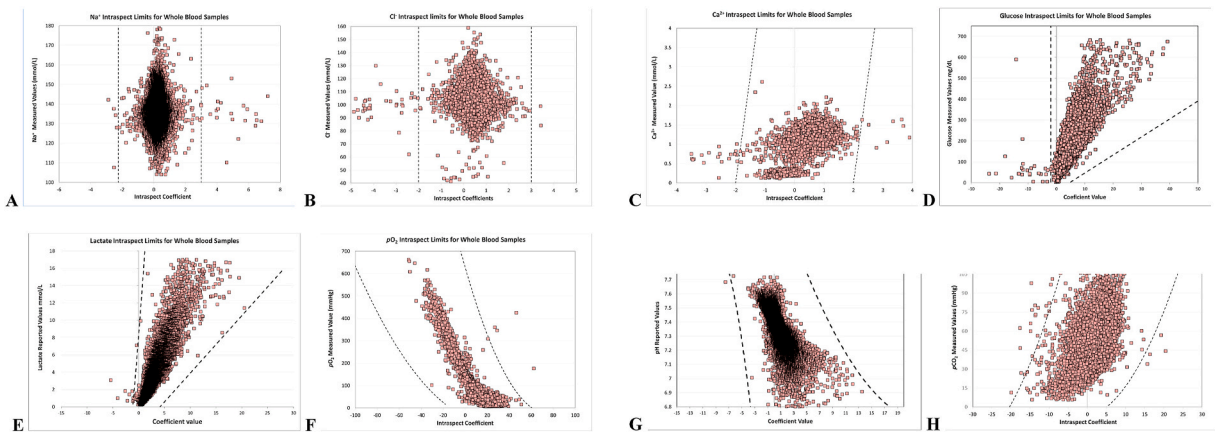
Fig. 4. Distribution of slope coefficients for various tests monitored by IntraSpect. Part (A) shows IntraSpect limits for Na$^+$; (B) Cl$^-$; (C) Ca^{++}; (D) glucose; (E) lactate; (F) pO$_2$; (G) pH; and pCO$_2$.

IntraSpect coefficients values against the sample concentration for all tested sensors. Figures A to C show examples of constant coefficients. Figures D to H show the sensors where coefficients value (or sensor response) are variable to the concentration of measured analyte.

Table 1

Types of errors detected, corrected, and documented in 1,013,391 patient samples, analyzed with 4985 GEM PAKS by 2765 GEM Premier 5000 instruments.

Preanalytical category	Type of error detected by iQM2	Worldwide Error Detection Prevalence (%)
Improper mixing anticoagulant	Micro-clots	0.79
Inadequate sample preparation and/or	Benzalkonium chloride	0.09
patient-specific	Thiopental	0.02
treatment	CO-Ox interferences	0.21
Transient errors	IntraSpect	0.80
Total		1.91

Table 2

Distribution of transient errors by analyte as detected by IntraSpect.

	Distribution of errors among analytes								
	pH	pCO$_2$	pO$_2$	Na$^+$	K$^+$	Cl$^-$	Ca^{++}	Glu	Lac
Results affected (%)	0.02	0.15	0.03	0.10	0.05	0.39	0.14	0.06	0.07

When IntraSpect detects an error for a given analyte, the result for that particular analyte only is suppressed by the system. In combination with IntraSpect during-sample-measurement error detection, the totality of iQM2 processes is able to detect errors where the sensor response was affected before, during or after the sample [2,4,6].

The idea of using patient samples for QC has a long history. Beginning with Hoffman and Waid in the 1960s [10], a method of QC using the "average of normals" has been recommended. In the 1970s, the application of patient data was extended to red cell indices and "Bull's algorithm" [11] which has been widely implemented, possibly representing the most successful application of patient-based QC. In the 1990s, Westgard et al. [12] described how Average of Normals (AoN) could be designed to complement traditional SQC and utilized to measure run length. In the 2000s, interest in patient data QC was kept alive by Parvin [13] and Kazmierczak [14], but laboratories seemed occupied by the trend towards using risk management to design quality control plans. Then, since 2015, a wide variety of studies have been published on the performance of patient-based real-time QC by Bietenbeck [15], Fleming and Katayev [16], Ng et al. [17], Bietenbeck et al. [18], Liu et al. [19], van Rossum and van den Broek [20], and van Rossum and Kemperman [22]. Many of these investigators make up the IFCC PBRTQC group, recommending increased use of patient data for QC.

The IFCC group accepts that there will be a different design and validation models for the various applications of PBRTQC. The particular patient-based technique employed in IntraSpect makes use of the measurement response curve of an individual patient sample, in contrast to the focus on population statistics in most PBRTQC applications. As such, the IntraSpect application functions as a simple SQC model, where control limits are established based on the variation observed in measuring normal samples. Control limits are applied to monitor the shape of the response curve for each individual sample measured. Validation of performance, therefore, follows the principles applied for traditional SQC procedures but also relies on empirical observations of field data. There typically are sets of trial or training data used to establish control limits to minimize false rejections and maximize error detection. In the case of

PBRTQC procedures, it is particularly critical to observe the actual rejection rates in practice to optimize performance [21].

While pre-analytical and transient errors are observed in approximately 1.9% of patient samples globally, they are still very significant when the volume of blood gas testing is considered. It is estimated that approximately 430 million POC blood gas tests will be performed globally, corresponding to 8.17 million patient samples with pre-analytic and transient errors. While these errors would be detected by iQM2 with IntraSpect, they would likely go undetected by traditional QC procedures. Detection of that small percentage of pre-analytic and transient errors still represents a significant improvement in the quality of blood gas testing and demonstrates the usefulness of the new IntraSpect technology in iQM2.

4. Conclusions

Transient errors caused by micro-clots, air bubbles, interferences, or any other event capable of disturbing sensor response, can be detected by monitoring the response curves of each individual patient sample. Studies demonstrate that such transient errors occur in 0.5–1.0% of patient blood gas samples and are likely undetected when using liquid control solutions and traditional quality control (QC). As recommended by the International Federation of Clinical Chemistry and Laboratory Medicine, the addition of patient data QC procedures can complement statistical QC and improve detection of clinically important errors.

Declaration of competing interest

JOW is a Werfen consultant. JC is a Werfen employee.

References

[1] S.A. Westgard, H.M.J. Goldschmidt, S.S. Ehrmeyer, POCT analysts' perspective: practices and wants for improvement, J Appl Lab Med. 5 (2020) 480–493.
[2] J.H. Nichols, T. Cambridge, N. Sanchez, D. Marshall, Clinical validation of a novel quality management solution for blood gas, electrolytes, metabolites, and CO-Oximetry, J Appl Lab Med. 6 (6) (2021) 1396–1408, https://doi.org/10.1093/jalm/jfab053.
[3] J.O. Westgard, K.D. Fallon, S. Mansouri, Validation of iQM active process control technology, Point Care 2 (1) (2003) 1–7.
[4] J.G. Toffaletti, E.H. McDonnell, L.V. Ramanathan, J. Tolnai, R. Templin, L. Pompa, Validation of a quality assessment system for blood gas and electrolyte testing, Clin. Chim. Acta 382 (2007) 65–70.
[5] M.M. Mion, G. Bragato, M. Zaninotto, J. Alessandroni, S. Bernardini, M. Plebani, Analytical performance evaluation of the new GEM Premier 5000 analyzer in comparison to the GEM Premier 4000 and the RapidPoint 405 systems, Clin. Chim. Acta 486 (2018) 313–319.
[6] P. D'Orazio, Effects of blood clots on measurements of pH and blood gases in critical care analyzers, Point Care 4 (2011) 186–188.
[7] T. Badrick, A. Bietenbeck, M.A. Cervnski, A. Katayev, H.H. van Rossum, T.P. Loh, Patient-based real-time quality control: review and recommendations, Clin. Chem. 65 (2019) 962–971.
[8] T. Badrick, A. Bietenbeck, A. Katayev, H.H. van Rossum, M.A. Cervnski, T.P. Loh, Patient-based real time QC, Clin. Chem. 66 (2020) 1140–1145.
[9] S. Mansouri, J. Cervera, E. Kovalchick, Implementing total quality assurance in point of care testing, Clin. Chem. 63 (10) (2017) S225.
[10] R.G. Hoffman, M.E. Waid, The 'average of normals' method of quality control, Am. J. Clin. Pathol. 43 (1965) 134–141.
[11] B.S. Bull, R.M. Elashoff, J. Couperus, A study of various estimators for the derivation of quality control procedures from patient erythrocytic indices, Am. J. Clin. Pathol. 61 (1974) 473–481.
[12] J.O. Westgard, F.A. Smith, P.J. Mountain, S. Boss, Design and assessment of average of normal (AoN) patient data algorithms to maximize run lengths for automatic process control, Clin. Chem. 42 (1996) 1683–1688.
[13] Ye JJ, Ingels SC, Parvin C. Performance evaluation and planning for patient-based quality control procedures. Am J Clin Pathol. 200;113:240–248.
[14] S.C. Kazmierczak, Laboratory quality control: using patient data to assess analytical performance, Clin. Chem. Lab. Med. 41 (2003) 617–627.
[15] A. Bietenbeck, M.A. Cervinski, A. Katayev, T.P. Loh, H.H. van Rossum, T. Badrick, Understanding patient-based real-time Quality Control using simulation modeling, Clin. Chem. 66 (2020) 1072–1083.
[16] J.K. Fleming, A. Katayev, Changing the paradigm of laboratory quality control through implementation of real-time test results monitoring: for patients by patients, Clin. Biochem. 48 (2015) 508–513.
[17] D. Ng, F.A. Polito, M.A. Cervinski, Optimization of a moving average program using a simulated annealing algorithm; the goal is to monitor the process not the patients, Clin. Chem. 62 (2016) 1361–1371.
[18] A. Bietenbeck, M.A. Thaler, P.B. Luppa, F. Klawonn, Stronger together: aggregated z-values of traditional quality control measurements and patient medians improve detection of biases, Clin. Chem. 63 (2017) 1377–1387.
[19] J. Liu, C.H. Tan, T.P. Loh, T. Badrick, Verification of out-of-control situations detected by 'average of normals' approach, Clin. Biochem. 49 (2016) 1248–1253.
[20] H.H. Van Rossum, D. van den Broek, Ten-month evaluation of the routine application of patient moving average for real-time quality control in a hospital setting, JALM 5 (2020) 1184–1193.
[21] H.H. Van Rossum, H. Kemperman, Optimization and validation of moving average quality control procedures using bias detection curves and moving average validation charts, Clin. Chem. Lab. Med. 55 (2017) 218–224.