Iranian Journal of Pathology (2008)3 (3), 113- 118

Original Article

Remodeling of Average of Patients QC Method to Maximize Lengths of Analytical Runs in Regional Reference Laboratories

Peyman Mohammadi Torbati[1], Hamid Seyed Javadi[2]

1. Dept. of Pathology, Labbafinejad Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
2. MDS laboratories/Auto lab Systems, Etobiocoke, ON, Canada.

ABSTRACT

Background and Objective: **Improved and modified automation will require the development of smart process control systems that provide on-line decisions to release patients' test results based on high analytical quality assurance formula.**

Materials and Methods: **We collected patients' test results from 10840 healthy subjects based on 1.96z as truncation limit for 29 common haematochemical analytes at a regional reference laboratory. Computer simulation studies by EZ rules™ and EZ runs™ software were performed to generate operating specification charts (OPSpces) that consider truncation limits set at 3(s_{pop}) and control limits set at 3 $s_{pop/n^{1/2}}$ and number of patient subgroups which varied from 10 to 480 depending on the s_{pop}/s_{meas} ratio that varied from 1.58 to 19.75.**

Results: **On the basis of the test parameters defined and the workload expected in our regional laboratory, average of patients (AOP) algorithms would be expected to be useful for monitoring run length on analytical systems that test for ALP, ALT, AST, total bilirubin, calcium, creatinine, glucose, hematocrit, hemoglobin, potassium, sodium, TSH and urea. These tests provide high potential capability indicating lowP_{fr}, highP_{ed} and high analytical quality assurance (AQA) with low control observations for applying AOP algorithms to monitor run length.**

Conclusion: **Our investigation revealed that approximately fifty percent of commonly requested haematochemical tests could achieve high capability in order to establish AOP method to maximize analytical run length.**

Key words: Average of patients, Regional reference laboratory, Analytical run length

Introduction

The remodeling of laboratory processes should consider the need for optimizing and improving quality control procedures and practices. Advanced automation will require the development of smart process control systems that provide on-line decisions to release patient test results, repeat analyses of patient specimens, and maximize the cost-effective operation of the testing process. Providing these capabilities will require strictly designed control procedures that optimize the frequency of false rejections, probability of error detections, and maximize the analytical run length.

Received: 16 March 2008
Accepted: 9 April 2008
Address communicatins to: Dr. Peyman Mohammadi Torbati, Department of Pathology, Labbafinejad Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
Email: p2000torbati@yahoo.com

The development of quality control (QC) planning tools such as power function graphs, critical error graphs, and operating specification charts (OPSpecs) facilitate the selection of control rules and number of control measurements which are selected given to measured imprecision and inaccuracy of an analytical system (1;2). In this regard, QC planning is based on analytical and quality requirements that are needed for a particular analyte or test (3;4).

Recent studies have demonstrated the application of QC planning models for a variety of multi-test chemical analyzers and immunoassay methods providing a systematic QC planning process to optimize the detection of clinically significant errors (5;6).

Nevertheless, there has been little progress in developing a reliable and reproducible method for maximizing the analytical run length. In most laboratories, run lengths are established on the basis of regulatory requirements, manufacturers' recommendations, or laboratory experience with analytical system (7). Regulatory requirements may set a maximum run length of 24 hours and manufacturers may recommend an eight hours work shift. In practice, laboratories may define shorter periods such as 2 or 4 hours according to the stability of the method and measurement system.

Since the productivity of analytical system and the cost per test depend on the approximate numbers of standards, controls, and repeat patient samples, it would be desirable to develop an automatic process for selecting statistical QC procedures to determining the appropriate maximum run length (8). In 1994, Cembrowski (9) introduced the average of patients (AOP) QC method, in which the average of patients test results which are detected within the reference interval used to monitor variations in testing process. Hoffman et al assessed the expected performance and the practical usefulness of such procedure (10). They used powerful computer simulation programs to develop power curves that describe the probability of false rejection as a function of the size of the systematic error (11). They showed that the power of AOP QC method depends on the ratio of population to analytical standard deviation, the width of the truncation limits and control limits and finally the number of patient test results. In this paper, given to patient test results which are gathered from a regional reference laboratory, we describe how AOP procedures can be designed to provide the error detection needed to monitor the length of analytical run.

Materials and Methods

We collected patient test results of 10840 healthy subjects based on 1.96z as truncation limit for results of 29 common haematochemical analytes at a regional reference laboratory (Labbafinejad hospital) and all of the patients with test results which were located out of this control limit were excluded. Computer simulation studies by EZ rules™ and EZ runs™ software were performed to generate patient test values and patient population control statistics. Power curves were determined for ratios of the population standard deviation (s_{pop}) to the analytical standard deviation (s_{meas}) of 1.58 to 19.75, truncation limits set at $3(s_{pop})$, and control limits set at $3s_{pop}/_{n^{1/2}}$, where n is the number of patient test results, which varied from 10 to 480 depending on s_{pop}/s_{meas} ratio. The program allows user specification of the population mean, biological standard deviation s_{biol}, s_{meas}, n, truncation limits, control limits, systematic error, and number of runs at each level of systematic error.

The program simulates biological and analytical variation independently which are scaled appropriately and combined with the population mean and systematic error to yield a simulated patient test results. Patient test results that exceeded the truncation limits were omitted from calculations of the AOP statistics and from the sample count. For each s_{pop}/s_{meas} and n, we executed 20 simulated runs at various levels of systematic error and plotted power curves showing the probability for rejection vs. the size of systematic errors. For calculation of operating specifications, the sizes of systematic errors for which probabilities of rejection of 0.9, 0.5, and 0.25 would be achieved by the AOP procedures were estimated by interpolation of the points related to specified probabilities.

OPSpecs charts and graphs showing the critical-size systematic error superimposed on power curves were prepared with the QC Validator™ program (version 2.0). QC procedures assessed on the basis of the ratio determined for the test of interest. The critical size systematic error was calculated as follows: $\Delta SE_{crit} = [(TE_a - bias_{meas})s_{meas}]$**-1.65**, where

TE_a is the analytical allowable total error, $bias_{mes}$ is the observed inaccuracy or stable systematic error of the measurement procedure, and 1.65 is a Z-value. When the mean of the patient test results causes 5% of individual patient test results to have errors exceeding the total error requirement, the run will be considered unstable. For example, given an analytical quality requirement of 10%, observed imprecision of 2%, observed inaccuracy of 0.0%, and s_{pop}/s_{meas} ratio of 4, an AOP algorithm requires 40 patient test results to effectively monitor stability of the analytical run. Figure 1 shows an OPSpecs chart for AOP algorithms having n from 200 to 40, which are represented by the lines from top to bottom. The observed imprecision and inaccuracy are showing the operating point. An effective AOP algorithm is one whose limits of inaccuracy and imprecision are above the operating point. The solid line identifies an AOP algorithm with n=40 as being appropriate for this application. Figure 2 shows the critical-size error imposed on the power curves to further illustrate the error detection available with different n values. Given a test that has a different s_{pop}/s_{meas} ratio of 8 but the same analytical quality requirement and the same observed imprecision and observed inaccuracy, the AOP algorithm requires 120 patient test results to effectively monitor the stability of the analytical run. Figure 3 shows an OPSpecs chart for n from 450 to 120, as represented by the lines from top to bottom. The observed imprecision and inaccuracy are again shown by the operating point and the solid line identifies the AOP algorithm with n=120 as being appropriate for this application. Figure 4 provides comparable information in the form of the critical-error graph for systematic error.

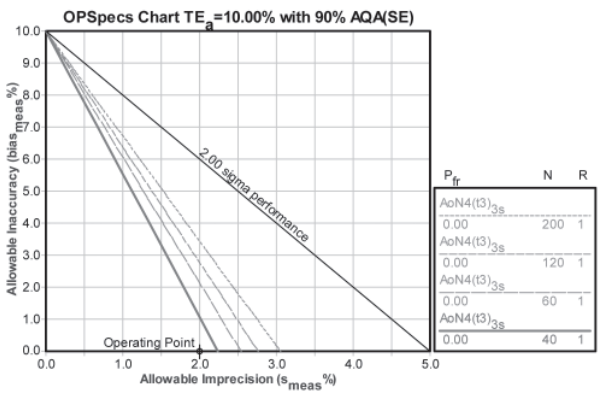

Figure 1: OPSpecs chart for AOP algorithm for a test possessing a ratio of 4

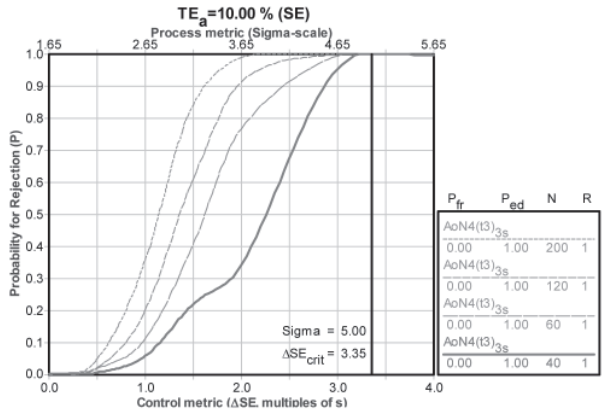

Figure 2: Critical-error graph for AOP algorithms for a testing possessing a ratio of 4

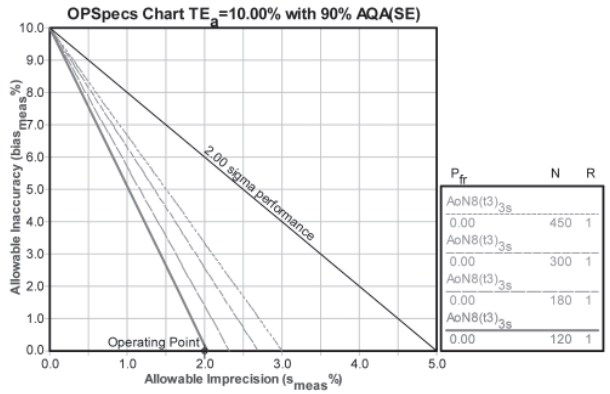

Figure 3: OPSpecs chart for AOP algorithm for a test possessing a ratio of 8

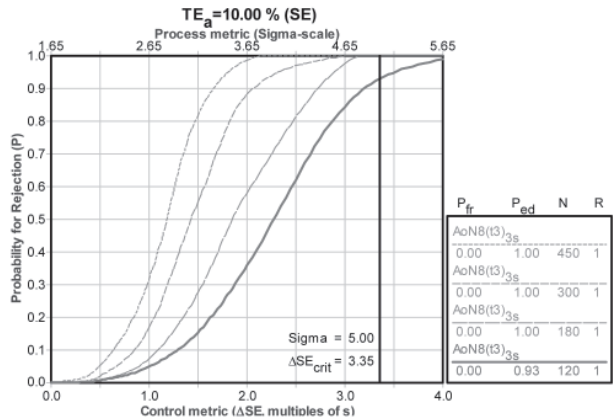

Figure 4: Critical-error graph for AOP algorithm for a test possessing a ratio of 8

Results

Table 1 summarizes important test characteristics and quality-planning parameters for the assessment and design of AOP algorithm for a variety of common laboratory tests. The number of tests expected per day varies from 25 to 960, depending on the analyte. These test volumes represent usual daily workloads in

a regional reference laboratory (pathology laboratory of Labbafinejad hospital, Tehran, Iran). Estimates of the population and method standard deviations represent observations on current analytical systems, resulting in s_{pop}/s_{mea} ratios from as low as 1.58 for sodium to as high as 19.75 for FSH and 32.5 for triglyceride. The analytical total error requirements (TEa) represent CLIA PT criteria for acceptability, these criteria are often more demanding for many chemistry tests [e.g. 5% vs. 10% for calcium, 15% vs. 30% for ALP, 15% vs. 20% for ALT and AST, 10% vs. 15% for creatinine, 7% vs. 10% for glucose] and less demanding for a couple of hematologic tests [10% vs. 6% for hematocrit]; thus these assessments and recommendations for AOP applications depend on the quality requirements that must be achieved by the particular laboratory.

On the basis of the test parameters defined and the workload expected in our regional laboratory, AOP algorithms would be expected to be useful for monitoring run length on analytical systems that test for ALP, ALT, AST, total bilirubin, calcium, creatinine, glucose, hematocrit, hemoglobin, potassium, sodium, TSH and urea (Table 1). AOP algorithms would have only moderate potential for monitoring run length for tests such as ferritin, LDH, phosphate, T3, T3 uptake, and urate (Table 1). For most of these tests, an increased daily workload would make AOP procedures more useful for monitoring run length for tests such as amylase, total and HDL cholesterol, CK, FSH, LH, T4, total protein, and triglyceride (Table 1). The overall findings point out the need to assess the performance of AOP algorithms for each laboratory on the basis of the quality requirements.

Table 1: Test characteristics and performance criteria to design AOP algorithm

Analyte	n, expected per day	subgroup	S_{pop}/S_{meas}	TE_a	SE_{crit}	P_{ed}
High	**potential**	**for AOP**	**application**			
ALP	450	450	*13.83*	15	3.77	*0.99*
ALT	255	40	*5.37*	15	4.53	*0.98*
AST	255	40	*2.2*	15	4.35	*1.00*
Total bilirubin	300	120	*8.6*	20	10.85	*0.96*
Calcium	270	60	*1.84*	5	1.21	*0.91*
Creatinine	850	40	*4.44*	10	3.91	*0.99*
Glucose	960	90	*6.00*	7	3.73	*0.99*
Hematocrit	480	80	*5.00*	10	8.35	*1.00*
hemoglobin	480	120	*9.09*	10	8.35	*0.96*
Potassium	600	100	*2.78*	5	1.68	*0.98*
Sodium	600	60	*1.58*	3	.85	*0.93*
TSH	450	300	*8.89*	30	2.97	*0.99*
Urea	850	450	*11.09*	10	2.61	*0.91*
Moderate	**potential**	**for AOP**	**application**			
Ferritin	260	120	*7.89*	24	2.35	*0.94*
LDH	450	60	*6.59*	8	3.68	*0.80*
Platelet	500	180	*11.59*	20	6.18	*0.91*
Phosphate	270	90	*7.65*	10	3.91	*0.93*
Total T3	350	10	*2.18*	24	3.15	*0.94*
T3 uptake	320	20	*3.00*	12	3.15	*0.76*
Uric acid	280	300	*13.59*	10	5.02	*0.97*
Low	**potential**	**for AOP**	**application**			
Amylase	25	60	*8.54*	15	5.85	*0..63*
Cholesterol	950	450	*10.34*	5	0.26	*0.12*
HDL	760	300	*7.00*	5	0.02	*0.16*
CK	64	20	*4.93*	15	2.85	*0.31*
FSH	120	180	*19.75*	24	1.78	*0.27*
LH	120	180	*15.27*	24	2.71	*0.27*
Total protein	80	80	*4.94*	5	0.98	*0.09*
Total T4	170	60	*6.10*	24	3.15	*0.18*
Triglyceride	920	480	*32.50*	5	0.98	*0.19*

Discussion

AOP algorithms can be designed to monitor the stability of tests and systems, with the objective of providing statistical evidence of instability to signal the end of a stable period of operation, or the end of an analytical run. The amount of instability that can be allowed before ending an analytical run is the critical systematic error calculated from the TEa and the observed imprecision and inaccuracy of the measurement procedure (12).This instability can be detected by AOP algorithms that depend on the ratio of the population and analytical standard deviations for the particular test and include an appropriate number of patient samples.

Cembrowski et al. (9) guidelines for implementing AOP procedures utilized a nomogram for selecting an appropriate number of patient samples on the basis of the ratio and the probability for detecting a shift equivalent to 2 times .Table 1 shows that the critical systematic error varies greatly ,from as low of 0.02 for HDL cholesterol to a high of 10.85 for total bilirubin; thus a more quantitative approach is needed to design the AOP algorithm for the exact conditions that must be monitored for individual tests in a particular laboratory.

To provide a quantitative design approach, power curves were determined for different ratios and a wide range of n values, then incorporated in the table of candidate QC procedures in a QC Validator 2.0. This program (11) enabled the preparation of OPSpecs charts, as well as critical error graphs, to facilitate the selection and design of appropriate AOP algorithms. This approach makes it practical to assess the expected performance of AOP algorithms quickly and easily for many individual tests on different analytical systems (13;14).

Results show that it may be difficult to obtain a sufficient numbers of patient samples to monitor the run lengths of some tests, even in a regional reference laboratory. For practical implementation, it may be necessary to determine whether the AOP algorithm for a single test can be used as a measure of stability for a group of tests or for a whole analytical system. Such applications should still be an improvement over the current practices for defining run length based on past experience with the instrument systems, rather than current performance data about system instability. Other instrument information may also be combined with AOP information to refine the assessment of instability and run length (15).

Implementation requires computer support to process the numbers of patient test results that are needed to provide the desired sensitivity for monitoring analytical runs. Conventional laboratory information systems and QC software may not provide the capability to accumulate large numbers of patient results for QC purposes, trim or select only those within specified truncation limits, calculate appropriate statistics to estimate the mean values, and then assist operators in interpretation the results and making decision on process performance (16).

Therefore strategies for implementation may need to consider alternative ways to calculate and display the data. The means of subgroups may be plotted on conventional means chart to display changes and trends and multi-rule decision criteria could be applied to evaluate the data. Cusum charts could also be used, with the means of subgroups as the data points. Exponentially weighted moving averages and control charts can be used to evaluate changes in process (17).

As laboratories reengineer their testing process and move towards higher levels of automation, improved process control software will be needed to assure both the quality of laboratory test results and productivity of laboratory testing processes. A recent College of American Pathologists Q-probe on laboratory QC practices (18) revealed that most laboratories that are accredited by IFCC are using the same QC practices as 10 years ago. These QC practices, which are already outdated by the introduction of newer generations of analytical systems, need to be improved to balance the error detection and false rejection characteristics of statistical QC procedures, as well as to maximize run length and the number of control measurements (11). Improved process control software will be necessary to support this optimization and also to implement improved QC designs for individual tests performed by multi-test systems.

Conclusion

In the future , process control software may evolve into the next generation laboratory information system because it will provide active control of analytical systems for the purpose of acquiring high quality patient test results at low cost. High quality will mean accurate test results with as short turnaround time as possible; low cost will mean effective utilization of available testing processes and associated laboratory resources. To achieve high quality and low cost,

laboratories will have to develop more quantitative approaches for managing the quality and productivity of their testing processes. Improved management should include careful selection of statistical process control rules and numbers of control measurements for stable control materials and careful application of AOP algorithms to utilize patient data for maximizing the length of the analytical run.

References

1. Westgard JO. Assuring analytical quality through process planning and quality control. Arch Pathol Lab Med 1992 Jul;116(7):765-9.

2. Westgard JO. Charts of operational process specifications ("OPSpecs charts") for assessing the precision, accuracy, and quality control needed to satisfy proficiency testing performance criteria. Clin Chem 1992 Jul;38(7):1226-33.

3. Westgard JO, Wiebe DA. Cholesterol operational process specifications for assuring the quality required by CLIA proficiency testing. Clin Chem 1991 Nov;37(11):1938-44.

4. Petersen PH, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, et al. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. Eur J Clin Chem Clin Biochem 1996 Dec;34(12):983-99.

5. Mugan K, Carlson IH, Westgard JO. Plannign QC procedures for immunoassays. J Clin Immunoassay 1994;17:216-22.

6. CLSI EP15-A2. User Verification of Performance for Precision and Trueness. Clinical Laboratory Standards Institute, Wayne PA, USA 2005.

7. CLSI C24-A3. Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definition; Approved Guideline. Clinical and Laboratory Standards Institute, Wayne PA, USA. 2006.

8. Westgard JO. The need for a system of quality standards for modern quality management. Scand J Clin Lab Invest 1999 Nov;59(7):483-6.

9. Cembrowski GS, Chandler EP, Westgard JO. Assessment of "Average of Normals" quality control procedures and guidelines for implementation. Am J Clin Pathol 1984 Apr;81(4):492-9.

10. Hoffmann Rg, Waid ME. The "Average of Normals" method of quality Control. Am J Clin Pathol 1995;43:134-41..

11. Westgard JO, Stein B, Westgard SA, Kennedy R. QC Validator 2.0: a computer program for automatic selection of statistical QC procedures for applications in healthcare laboratories. Comput Methods Programs Biomed 1997 Jul;53(3):175-86.

12. Wiebe DA, Westgard JO. Cholesterol--a model system to relate medical needs with analytical performance. Clin Chem 1993 Jul;39(7):1504-12.

13. Westgard JO, Stein B. Automated selection of statistical quality-control procedures to assure meeting clinical or analytical quality requirements. Clin Chem 1997 Feb;43(2):400-3.

14. Westgard JO. Quality Control: How labs can apply Six Sigma principles to quality control planning. Clin Lab News 2006(Jan):10-12.

15. Koch DD, Oryall JJ, Quam EF, Feldbruegge DH, Dowd DE, Barry PL, et al. Selection of medically useful quality-control procedures for individual tests done in a multitest analytical system. Clin Chem 1990 Feb;36(2):230-3.

16. NCCLS Document C24-A. Internal quality control testing: principles and definitions. Villanova, PA: National Committee for Clinical Laboratory Standards. Clinical and Laboratory Standards Institute, Wayne PA, USA.2006.

17. Smith FA, Kroft SH. Exponentially adjusted moving mean procedure for quality control. An optimized patient sample control procedure. Am J Clin Pathol 1996 Jan;105(1):44-51.

18. Terault GA, Steindel SJ. Q-Probe 98-08. Daily quality control exception practices. Chicago , IL: College of American Pathologists,1994.