Rev. sci. tech. Off. int. Epiz., 2005, **24** (3), 1005-1016

Quality assurance/quality control of foot and mouth disease solid phase competition enzyme-linked immunosorbent assay – Part II. Quality control: comparison of two charting methods to monitor assay performance

N. Goris & K. De Clercq

Department of Virology, Veterinary and Agrochemical Research Centre, Groeslenberg 99, 1180 Brussels, Belgium

Submitted for publication: 12 April 2005
Accepted for publication: 26 July 2005

Summary
Diagnostic laboratories are increasingly required to meet stringent quality standards, and validated assays are needed to achieve formal accreditation. Validation of test methods is often considered to be finalised when the assay parameters and characteristics have been established. However, like any process, diagnostic assays are subject to random variation resulting in shifts in the mean test values. Continuous monitoring of assays using control charts will alert the interpreter of changes in performance. For this purpose, several charting methods have been developed and implemented. This paper compares the Shewhart and the exponentially weighted moving average (EWMA) control charts with respect to the day to day monitoring of internal quality control samples for the foot and mouth disease solid phase competition enzyme-linked immunosorbent assay. Both chart types are equally sensitive to shifts, but the EWMA method seems to provide the best balance between false rejection and false acceptance.

Keywords
Accreditation – Control chart – Enzyme-linked immunosorbent assay – Exponentially weighted moving average – Foot and mouth disease – Quality control – Shewhart – Validation – Westgard rule.

Introduction

The foot and mouth disease virus (FMDV) is the causal agent of a devastating disease that affects all susceptible cloven-hoofed animals. Several virological (3, 7, 26) and serological (1, 6, 9, 15, 19) techniques have been developed to detect FMDV infection and/or to monitor the immune response in various species. Serological assays, however, have been and still remain the primary diagnostic method for international trade purposes (33). For the detection of antibodies to FMDV, one of the methods prescribed by the World Organisation for Animal Health (OIE) for international trade of animals and animal products is the solid phase competition enzyme-linked immunosorbent assay (SPCE) (18). International collaborative efforts, sponsored by the European Commission for the control of foot and mouth disease (FMD) of the Food and Agriculture Organization of the United Nations (FAO-EUFMD), to standardise and harmonise the SPCE protocol at a global level are ongoing (13, 16, 20, 21, 22, 23).

However, no matter how well an assay performs in the diagnostic laboratory, it is of little value unless trading partners accept the results with confidence (31). Therefore, the issue of quality control (QC) in veterinary laboratories, including FMD reference laboratories, has rapidly gained importance over the last few years (5). A validated assay (standard or laboratory-developed) is a prerequisite in any laboratory that utilises an external QC scheme, such as the International Organization for Standardization/ International Electrotechnical Commission (ISO/IEC) International Standard 17025 (11). However, assay performance is a dynamic process and the characteristics of an assay (diagnostic sensitivity and specificity, repeatability, and reproducibility) should be regularly updated. Method validation should, therefore, be conceived as an ongoing process that requires vigilance and maintenance (12). When the major sources of variability (biological reagents, buffers, and physicochemical parameters) have been minimised through the establishment of a standard operating procedure, operational QC provides a measure of the inherent variability of the method (31) and thus quantifies the robustness (precision) of the assay during routine use (variability may occur due to small differences in ambient temperature, different operators, etc.). Internal QC (IQC) procedures are performed on data generated from control samples incorporated in daily diagnostic runs.

For enzyme-linked immunosorbent assay (ELISA) tests, working or tertiary standards may function as control samples for IQC purposes (32). Since ISO/IEC 17025 obliges regular use of reference standards to estimate the accuracy of a test (11), to minimise the use of precious space on the ELISA plate, it is preferable to monitor well-developed working standards rather than developing additional IQC references and incorporating multiple types of reference samples (primary, working, and/or IQC) (8). When reference materials are used for dual purposes (accuracy and precision) in an assay, 'standard' and 'control' are functional terms that are used synonymously (32).

Ensuring the validity of an assay by constantly monitoring the performance of the incorporated controls using control charts was first described by Walter Shewhart (28) and subsequently introduced in clinical chemistry by Levey and Jennings (14). The charts are, therefore, sometimes referred to as 'Shewhart charts' and other times as 'Levey-Jennings charts'. This powerful tool monitors IQC data variation in time and visualises trends and process improvement. It was not until 1998 when Jacobson (12), and later Rebeski and colleagues (25), described a QC method for diagnostic ELISAs for infectious diseases. Paiba *et al.* (19) subsequently employed the Levey-Jennings charts to validate the FMD SPCE serotype O.

In the past, several criteria have been proposed for determining whether a run (plate) should be accepted or rejected. For example, according to the initial criteria, a run should be rejected if one of the control values exceeds, in either way, the control limit of three standard deviations from the mean (28, 31). The criteria were later extended to include additional conditions (unirule versus multirule and univariate versus multivariate procedures), called the Westgard rules (29).

Other types of control charts have since been developed and used, such as cumulative sum (CUSUM) charts (24), combined Shewhart-CUSUM charts (2, 30), and exponentially weighted moving average (EWMA) charts (27) and are said to be equally or even more sensitive in error detection than the conventional Shewhart charts and Westgard rules (17, 24, 30). However, in spite of the theoretically convincing concepts on which the more recent charts were developed, the mathematical prerequisites needed are more complex than the requirements for the Shewhart chart and have hindered implementation in diagnostic laboratories (17).

Fortunately, due to the development of computerised software programmes and an easy-to-implement design strategy (4), EWMA charts are no longer fancy tools for statisticians but can readily be applied by laboratory personnel. This paper describes the application of the IQC procedure to the FMD SPCE using the conventional Shewhart chart (with the different Westgard rules) and the EWMA chart. A comparison of both control-charting methods was performed on IQC data generated by previously developed working standards/controls.

Materials and methods

The solid phase competition enzyme-linked immunosorbent assay for the detection of antibodies to foot and mouth disease virus

The SPCE used in this study is a modified version of the protocol published by Mackay and colleagues (15) and has previously been described (8).

Internal quality controls: selection and normalisation

Working standards for the SPCE were prepared following the method described by Goris and De Clercq (8). Briefly, candidate primary reference sera, originating from the FAO-EUFMD Collaborative Study Phases XV to XVII (13, 16, 21, 23), were used as reference values for the production of in-house secondary standards. Working standards were subsequently calibrated against the secondary standards, and a strong positive (SP), weak positive (WP), cut-off (C/O), and negative (NC) working standard were developed. All four standards were incorporated in daily diagnostic runs to guarantee the accuracy of the assay.

The carefully developed standards covered the entire detection range of the assay and, therefore, constituted a potential IQC. They were, thus, redefined, for IQC purposes, as (working) controls for monitoring the performance of the SPCE over time and estimating the precision of the assay.

On each ELISA plate, four replicates of the no serum (Co), C/O, and WP controls were incorporated. The NC and SP control sera were run in duplicate. The median raw optical density (OD) readings, taken at 490 nm, of the Co, C/O, WP, and NC controls were normalised against the mean $OD_{490\ nm}$ of the SP control. The median rather than the mean of the four replicates was chosen for calculation purposes as this value more closely approached the true control value by excluding bias due to dispersion of the four replicates (25). The normalised QC data were subsequently plotted on both Shewhart and EWMA control charts.

Shewhart control chart

A reference pool of results derived from 20 tests (12) was needed to establish the control data range (25). Based on these results, Shewhart control charts for the normalised Co, C/O, WP, and NC controls were produced. Using the defined pool, a mean target value (m) and a corresponding standard deviation (s) were calculated for the normalised control values and used to establish the warning (m ± 2.09 s) and action limits (m ± 3.58 s) of the respective control charts. The limits were calculated using the Student's t probability, with 19 degrees of freedom and a probability of 5% and 0.2%, respectively.

Although Shewhart charts can easily be produced using a simple Microsoft Excel sheet, an automated process that makes use of all the Westgard rules for chart interpretation is advisable to exclude human error. Therefore, the Shewhart charts were established through a computerised software programme, called MedLabQC (copyright Philippe Marquis, www.multiQC.com).

MedLabQC generates multirule Shewhart control charts using conventional Westgard rules as the criteria to accept or reject runs. Warning and action signals are applied within and/or between control values (for abbreviations refer to Table I): 1:2 s (within), 1:3 s (within), 2:2 s (within and between), R:4 s (within and between), 4:1 s (within and between), and 10:m (within and between). According to Westgard *et al.* (29), a plate should be rejected when a violation of the 1:3 s rule occurs. The 1:2 s rule is a warning rule indicating the need for additional control data inspection. Following subsequent inspection, if one or more additional Westgard rules are violated, the assay should be considered 'out of statistical control' because of the occurrence of random or systematic errors, and the plate should be rejected. In all other instances, the plate can be accepted and the assay is regarded as being 'in statistical control'.

Exponentially weighted moving average chart

To make a meaningful comparison between both charting methods, the same 20 data point reference pool was used to determine the mean target value (m) and standard deviation (s) for the Co, C/O, WP, and NC normalised control values. However, control limits, data interpretation, and acceptance and rejection rules differed between the two charting methods.

Contrary to the Shewhart control charts, EWMA charts not only take the immediate control value into consideration, but also use the control values from previous runs to judge whether the current run (plate) is 'in statistical control'. The control value obtained at a certain moment in time (x_t) is multiplied by a weighting factor (λ) and added to the weighted sum of all former measurements of the same control sample multiplied by a factor of $(1 - \lambda)$. Thus, at each time point (t = 1, 2, …, n) the test statistic $EWMA_t$ is calculated according to the following formula:

$$EWMA_t = \lambda\, x_t + (1 - \lambda)\, EWMA_{t-1}, \text{ with } \lambda \in\,]0;1]$$

To produce an EWMA chart for the Co, C/O, WP, and NC controls, the Crowder four-step procedure (4) was implemented. Briefly, chart parameters, such as average run length (ARL) (a measure of the false rejection rate or α-error) and λ, were determined according to the following steps:

a) select the smallest acceptable ARL, which should be obtained under 'statistically in-control conditions'. This corresponds to fixing the false alarm rate

b) select the desired magnitude of shift from the mean in the assay that must be readily detected by the EWMA charting method. This implies fixing the sensitivity of the charting method. Based on this sensitivity, λ is determined using the graphs provided by Crowder (4)

c) calculate the control limits

d) perform a sensitivity analysis by varying the arbitrarily chosen ARL and λ to obtain the best overall performance.

Due to the complex calculations that were needed to update the charts after each run, charts were generated by the MultiQC software programme (copyright Philippe Marquis, www.multiQC.com), using predetermined parameter values for ARL and λ, and control values were then entered for each run (ELISA plate).

MultiQC is an EWMA chart-generating programme that, when used in the multivariate QC mode, has the ability to determine the correlation between all IQC samples on a plate by calculating the Hotelling T^2 value (for an overview see reference 10). The correlation values are plotted on a separate T^2 chart and integrated in the decision process. Additionally, MultiQC has a built in Student's t table for automatic calculation of the control limits. The associated

Table I
The Westgard rules

Abbreviation	Explanation	Example
1:2 s	One control value exceeds the control limit, which is set at approximately m ± 2 s. It is a warning rule indicating that additional inspection of control data is required	The normalised median weak positive (WP) internal quality control value (in short WP value) < m – 2 s
1:3 s	One control value exceeds the control limit, which is set at approximately m ± 3 s. This is the usual rejection criterion on a Shewhart control chart	Negative (NC) value > m + 3 s
2:2 s	Two consecutive control values exceed the same limit, which is set at approximately m + 2 s or m – 2 s). 'Consecutive' refers to two different control values on the same plate or the same control value on two different plates	1 plate: WP value > m + 2 s AND (Co) value > m + 2 s 2 plates: plate 1: cut-off (C/O) value < m – 2 s AND plate 2: C/O value < m – 2 s
R:4 s	Within a plate, the range (R) between the two control values exceeds 4 s Between two successive plates, the difference (R) for the same control value exceeds 4 s: one of the values exceeds the approximated m + 2 s, the other value exceeds the approximated m – 2 s	1 plate: WP value < m – 2 s AND C/O value > m + 2 s 2 plates: plate 1: NC value < m – 2 s AND plate 2: NC value > m + 2 s
4:1 s	Four consecutive control values exceed the same limit, which is set at approximately m + 1 s or m –1 s. This can occur between two different control values on two successive plates, as well as within one control value on four successive plates	2 plates: plate 1 & 2: WP value > m + 1 s AND plate 1 & 2: C/O value > m + 1 s 4 plates: plate 1, 2 ,3 & 4: Co value < m – 1 s
10:m	Ten consecutive control values fall on the same side of the m. The trend occurs on five successive plates for two different controls or on ten successive plates for the same control value	5 plates: plate 1→5: C/O value > m AND plate 1→5: WP value > m 10 plates: plate 1→10: NC value < m

control interval is defined as m ± t*s with the Student's t being dependent upon the size of the control interval, the chosen ARL, and the number of degrees of freedom of s. A run (ELISA plate) is rejected when one or more controls at a given time point exceeds the defined upper and/or lower control limit and/or when the correlation between the control values is violated (T^2 is greater than or equal to 1).

Results

Internal quality control samples

Working reference sera were developed for the SPCE for FMDV strains O_1 Manisa, A_{22} Iraq 24/64, A Iran 1996, C_1 Noville, and Asia 1 Shamir as previously described (8). Development of working standards was based on the characteristics of the candidate primary reference sera originating from the FAO Collaborative Study Phases XV to XVII (13, 16, 21, 23). The resulting normalised Co, C/O, WP, and NC IQC values for the SPCE type A Iran 1996 will be discussed in depth. Similar results were obtained for the other FMDV strains mentioned above.

Table II summarises the 85 values for each of the A Iran 1996 SPCE controls collected from 7 January to 26 August, 2004. The values were subsequently analysed using both the Shewhart and EWMA control chart software programmes.

Shewhart control chart

Figure 1 illustrates the Shewhart control charts generated by the MedLabQC software programme for the IQC data obtained for the SPCE type A Iran 1996. The first 20 observations (7 January to 10 March) constituted the reference pool and were used to estimate the mean target

Table II
The normalised internal quality control values for the solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 collected from 7 January to 26 August 2004

Date	Plate	C/O	WP	NC	Co
7/01/2004	1	2.97	1.94	4.35	4.51
8/01/2004	2	3.25	2.04	4.81	5.05
14/01/2004	3	3.55	2.11	5.08	5.25
23/01/2004	4	3.55	1.96	4.67	5.24
4/02/2004	5	3.02	1.93	4.66	4.64
5/02/2004	6	2.65	1.81	3.70	3.94
10/02/2004	7	2.85	1.83	4.13	4.27
12/02/2004	8	2.76	1.75	4.38	4.76
16/02/2004	9	3.12	1.93	4.59	4.87
19/02/2004	10	2.84	1.88	4.46	4.68
3/03/2004	11	3.62	2.19	5.49	5.62
3/03/2004	12	3.49	2.15	5.20	5.55
4/03/2004	13	2.89	1.76	4.05	4.71
4/03/2004	14	3.75	2.17	5.60	5.97
8/03/2004	15	3.78	2.11	5.50	5.76
8/03/2004	16	3.29	1.96	5.25	5.39
9/03/2004	17	3.24	1.87	5.36	5.52
9/03/2004	18	3.46	1.95	5.84	6.05
10/03/2004	19	3.09	1.90	4.55	4.84
10/03/2004	20	3.11	1.90	4.43	4.59
12/03/2004	21	3.39	1.90	5.67	5.89
12/03/2004	22	3.50	2.14	5.65	5.93
16/03/2004	23	3.14	1.91	4.72	4.79
16/03/2004	24	3.67	1.96	5.33	5.50
16/03/2004	25	3.19	2.04	4.29	4.61
17/03/2004	26	3.80	2.43	5.00	5.11
17/03/2004	27	3.47	2.06	5.62	5.80
19/03/2004	28	3.25	2.03	4.22	4.52
22/03/2004	29	2.85	1.95	3.87	3.99
25/03/2004	30	3.23	1.96	4.50	4.95
29/03/2004	31	2.82	1.90	3.79	3.97
1/04/2004	32	3.00	1.88	4.02	4.36
1/04/2004	33	3.22	1.90	4.30	4.60
5/04/2004	34	3.52	1.88	4.34	4.84
5/04/2004	35	2.94	1.88	4.19	4.24
6/04/2004	36	3.26	2.07	4.34	4.52
7/04/2004	37	3.32	2.04	4.52	4.54
8/04/2004	38	3.52	1.95	5.06	5.22
16/04/2004	39	3.37	2.03	4.88	5.09
20/04/2004	40	3.41	1.96	4.78	5.08
22/04/2004	41	3.28	1.99	4.40	4.56
22/04/2004	42	3.45	2.06	4.76	5.01
27/04/2004	43	2.61	1.80	3.93	4.16
28/04/2004	44	2.97	1.86	4.00	4.07
28/04/2004	45	3.12	1.94	4.07	4.20
3/05/2004	46	3.30	2.03	4.63	4.66
6/05/2004	47	2.92	1.81	4.03	4.23
6/05/2004	48	3.43	2.01	4.70	4.96
10/05/2004	49	3.52	2.02	5.13	5.30
18/05/2004	50	3.62	2.06	5.91	6.02
25/05/2004	51	3.68	2.00	5.72	6.11
26/05/2004	52	3.38	1.83	4.36	5.31
27/05/2004	53	3.37	1.99	5.15	5.16
7/06/2004	54	2.96	1.93	4.24	4.36
7/06/2004	55	2.92	1.84	4.68	4.90
8/06/2004	56	2.88	1.80	4.04	4.42
9/06/2004	57	3.10	1.99	4.32	4.44
14/06/2004	58	2.96	1.88	4.28	4.58
14/06/2004	59	3.03	1.91	3.96	4.66
14/06/2004	60	3.29	1.94	4.69	5.06
15/06/2004	61	2.68	1.79	3.80	4.17
15/06/2004	62	3.13	1.90	4.10	4.30
15/06/2004	63	3.02	1.89	4.23	4.63
28/06/2004	64	2.53	1.68	3.70	3.79
29/06/2004	65	3.12	1.85	4.62	4.90
1/07/2004	66	2.76	1.75	3.99	4.01
1/07/2004	67	2.71	1.79	3.79	3.86
1/07/2004	68	3.22	1.91	4.41	4.76
1/07/2004	69	3.14	1.90	4.34	4.85
1/07/2004	70	2.84	1.85	4.16	4.40
2/07/2004	71	2.75	1.78	4.00	4.13
2/07/2004	72	2.92	1.81	4.20	4.48
2/07/2004	73	2.94	1.88	4.20	4.47
2/07/2004	74	2.87	1.85	4.32	4.55
5/07/2004	75	3.04	1.91	4.19	4.35
5/07/2004	76	2.67	1.79	4.24	4.48
5/07/2004	77	3.19	1.99	4.19	4.10
5/07/2004	78	2.58	1.63	4.30	4.44
6/07/2004	79	3.21	1.85	4.88	5.03
8/07/2004	80	3.27	2.03	5.00	4.96
8/07/2004	81	3.56	2.07	5.27	5.28
8/07/2004	82	3.03	1.79	4.26	4.47
9/07/2004	83	3.92	2.23	6.10	6.39
20/08/2004	84	3.44	2.05	4.79	4.95
26/08/2004	85	3.14	2.00	4.58	4.70

Co: no serum control
C/O: cut-off control
NC: negative control
WP: weak positive control

value (m) and the standard deviation (s) of each control as well as to fix the warning and action limits (Table III).

An additional 65 IQC values for the C/O, WP, NC, and Co controls (12 March to 26 August, 2004) were subsequently collected, entered into the MedLabQC software programme, and evaluated. The Westgard rules were violated on various occasions (Table IV). However, apart from four occasions on which the test was regarded as 'out of statistical control' (17 March plate 26, 28 June, 5 July plate 78 and 9 July), the majority of violations were considered minor violations and did not require any actions to be taken. The plates that were classified as 'out of statistical control' were rejected; however, when the runs were repeated at a later date the test was found to be 'in statistical control'.

Table III
The Shewhart chart parameters for the solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 based on a reference pool of 20 internal quality control (IQC) values (7 January to 10 March 2004)

IQC sample	Mean (m)	Standard deviation (s)	Warning limits (m ± 2.09s)	Action limits (m ± 3.58s)
Cut-off	3.214	0.338	2.506 – 3.922	2.003 – 4.425
Weak positive	1.957	0.132	1.680 – 2.234	1.484 – 2.430
Negative	4.805	0.582	3.587 – 6.023	2.722 – 6.888
No serum (Co)	5.061	0.574	3.859 – 6.262	3.006 – 7.115

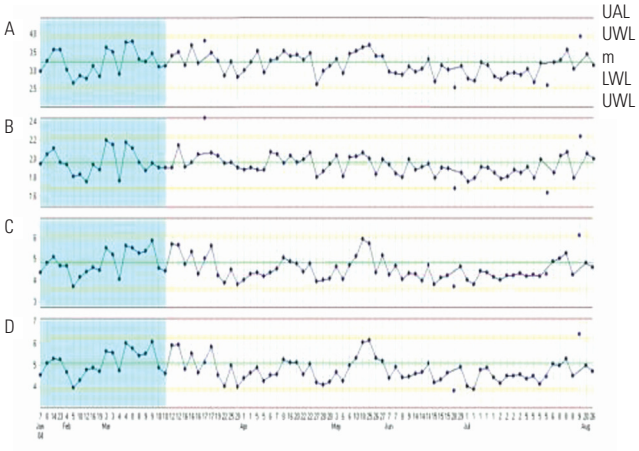

Fig. 1
The Shewhart charts for the foot and mouth disease solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 obtained using the MedLabQC software programme
Chart A represents the 85 data points obtained for the cut-off control.
Chart B is based on the data for the weak positive control.
Chart C displays the negative control and chart D the no serum (Co) control values.
The blue zone represents the reference pool collected from 7 January to 10 March (n = 20). The mean target value (m) is designated by the green line. The yellow lines are indicative of the upper and lower warning limits, and the red lines represent the upper and lower action limits. The dark blue dots are the accepted control values, and the light blue dots are the rejected values

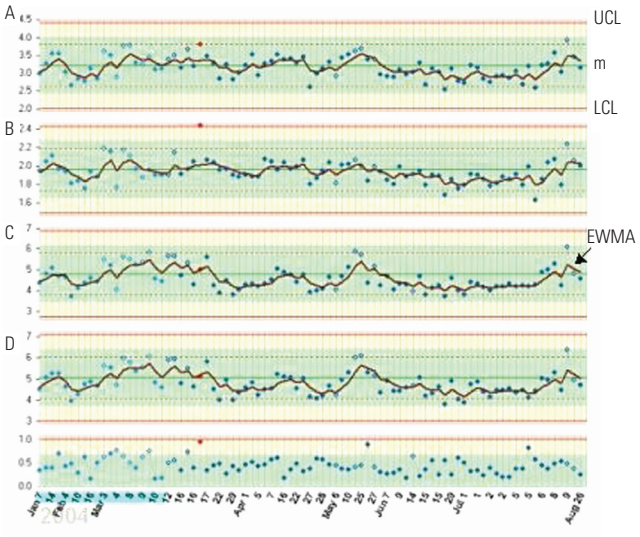

Fig. 2
The exponentially weighted moving average (EWMA) charts for the foot and mouth disease solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 obtained using the MultiQC software programme with an average run length of 370 and a λ value of 0.4
Chart A represents the 85 data points obtained for the cut-off control.
Chart B is based on the data for the weak positive control.
Chart C displays the negative control and chart D the no serum (Co) control values.
The red line running through charts A to D represents the $EWMA_t$. The calculated Hotelling T^2 value is represented in chart E (T^2 equal to or greater than 1 results in plate rejection). The light blue dots represent the reference pool collected from 7 January to 10 March (n = 20). The mean target value (m) is designated by the green line, and the upper and lower red lines represent the upper and lower control limits. The dark blue dots are the accepted control values, and the red dots are the rejected values

Exponentially weighted moving average chart

Figures 2 and 3 represent the EWMA control charts generated by the MultiQC software programme for the IQC data obtained for the SPCE type A Iran 1996. The first 20 observations (7 January to 10 March) constituted the reference pool and were used to estimate the mean target value (m) and the standard deviation (s) of each control.

To further compare the Shewhart and EWMA charts, two different values for the ARL were chosen. Assuming a Gaussian distribution, the ARL equalled the $1/\alpha$ error of the Shewhart charts. When applied to the EWMA charts,

Fig. 3
The exponentially weighted moving average (EWMA) charts for the foot and mouth disease solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 obtained using the MultiQC software programme with an average run length of 100 and a λ value of 0.5
Chart A represents the 85 data points obtained for the cut-off control.
Chart B is based on the data for the weak positive control.
Chart C displays the negative control and chart D the no serum (Co) control values.
The line running through charts A to D represents the $EWMA_t$.
The calculated Hotelling T^2 value is represented in chart E (T^2 equal to or greater than 1 results in plate rejection). The light blue dots represent the reference pool collected from 7 January to 10 March (n = 20). The mean target value (m) is designated by the green line, and the upper and lower red lines represent the upper and lower control limits. The dark blue dots are the accepted control values, and the red dots are the rejected values

this resulted in fixing the ARL to 370, which was comparable to the use of control limits of approximately 3 s in Shewhart charts (α equals approximately 0.0027) (17). Employing the Westgard algorithm (i.e. reliance on the Westgard multirules for Shewhart control chart interpretation), a corresponding ARL of 100 for 'in statistical control' conditions was calculated (17). Control limits were fixed based on the Student's t calculations, with a 20 point reference interval and ARLs of 370 and 100 (Table V). The λ values that resulted in maximum protection against a 2 s shift from the mean were determined for both of the ARLs and found to be 0.4 and 0.5, respectively.

The additional 65 IQC values for the C/O, WP, NC, and Co controls (12 March to 26 August, 2004) were subsequently collected, entered into the MultiQC software programme, and evaluated. Only one plate was rejected using the EWMA charts (17 March plate 26). The plate was rejected because the value of the WP control exceeded the upper control limit. The corresponding calculated Hotelling T^2 value was 0.93, indicating poor correlation between the four controls. None of the other plates were rejected under either of the conditions (ARL = 370 with λ = 0.4 or ARL = 100 with λ = 0.5).

Two downward trends were also detected: 12 March to 5 April and 26 May to 1 July. In addition, from 6 May to 25 May and between 5 July and 9 July, the EWMA showed upward drifts under both of the conditions.

Discussion

The normalised IQC data obtained with the no serum control and the working C/O, WP, and NC reference sera were successfully utilised in two easy-to-implement computerized charting programmes: one that generated Shewhart charts and another that produced EWMA charts. Both assay monitoring tools indicated that the test performed on 17 March (plate 26) needed to be rejected because the WP control value exceeded the upper action or control limit. The correlation between the four controls, calculated by the MultiQC programme, indicated that at the specified point in time a random error had occurred for the WP control, while the data for the C/O, NC, and Co controls were all within two thirds of the control interval (i.e. m + 2 s for the Shewhart charts).

When the two charting methods were further compared, the Shewhart charts using the Westgard multirules seemed more sensitive at first glance. Plates were rejected on three additional occasions and several violations of the decision rules were reported when the Shewhart charts were used. However, when the data for the EWMA charts was inspected more closely, all of the observed Westgard violations corresponded to upward and downward trends on the EWMA charts. For example, the violations on 12 March and from 1 April to 7 April were visualised as the downward trend noted from 12 March to 5 April, and the violations from 14 June to 1 July were part of the downward trend from 26 May to 1 July. For the latter, the trend was detected even earlier than the first Westgard violation (26 May versus 14 June). Thus, performing trend analysis on EWMA charts makes the control procedure, at least, equally as sensitive as the Shewhart charts to possible shifts in assay accuracy.

Furthermore, closer analysis of the plates rejected on 28 June and 9 July revealed a proportional downward and upward shift, respectively, from the mean target value for all controls, indicating that the WP, C/O, and NC control sera reacted in a similar fashion. This conclusion was further emphasised by the calculated Hotelling T^2 correlation values of 0.54 and 0.48, respectively, which indicated that routine data interpretation was not

Table IV
Westgard rules violations for the solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 internal quality control values collected from 12 March to 26 August 2004

Date	Plate	C/O	WP	Across	NC	Co	Across	Action
12/03/2004	22						4:1 s	None
17/03/2004	26		1:2 s, 1:3 s					Reject plate
1/04/2004	32						4:1 s, 10:m	None
1/04/2004	33						10:m	None
5/04/2004	34						10:m	None
5/04/2004	35						10:m	None
6/04/2004	36						10:m	None
7/04/2004	37				10:m	10:m	10:m	None
28/04/2004	44						4:1 s	None
28/04/2004	45						4:1 s, 10:m	None
3/05/2004	46						10:m	None
6/05/2004	47						10:m	None
6/05/2004	48						10:m	None
25/05/2004	51						4:1 s	None
14/06/2004	58						10:m	None
14/06/2004	59						10:m	None
14/06/2004	60						10:m	None
15/06/2004	61						10:m	None
15/06/2004	62						4:1 s, 10:m	None
15/06/2004	63				10:m	10:m	10:m	None
28/06/2004	64	1:2 s			10:m	1:2 s, 10:m	10:m	Reject plate
29/06/2004	65				10:m	10:m	10:m	None
1/07/2004	66			10:m	10:m	10:m	10:m	None
1/07/2004	67				4:1 s, 10:m	10:m	4:1 s, 10:m	None
1/07/2004	68		10:m		10:m	10:m	10:m	None
1/07/2004	69		10:m		10:m	10:m	10:m	None
1/07/2004	70		10:m		10:m	10:m	10:m	None
2/07/2004	71		10:m		10:m	10:m	4:1 s, 10:m	None
2/07/2004	72		10:m		10:m	10:m	10:m	None
2/07/2004	73		10:m	10:m	10:m	10:m	10:m	None
2/07/2004	74		10:m	10:m	10:m	10:m	10:m	None
5/07/2004	75		10:m	10:m	10:m	10:m	10:m	None
5/07/2004	76		10:m	10:m	10:m	10:m	10:m	None
5/07/2004	77				10:m	10:m	10:m	None
5/07/2004	78	10:m	1:2 s		10:m	10:m	10:m	Reject plate
6/07/2004	79	10:m				10:m		None
8/07/2004	80					10:m		None
9/07/2004	83				1:2 s	1:2 s	2:2 s	Reject plate

Co: no serum control
C/O:cut-off control
NC: negative control
WP: weak positive control

hindered. The tested sera had proportionally higher and lower $OD_{490\ nm}$ values than on previous occasions, but the sera were scored correctly with respect to the presence or absence of FMDV specific antibodies (fit for purpose). Rejecting the plate and repeating the test at a later date was a waste of precious laboratory time and money, as confirmed by the retest results. The same applied to the 5 July test, which had a calculated correlation value of 0.57. It is likely that a diagnostic laboratory performing the test would have reached the same conclusion based on common sense interpretation of the data generated by the different controls.

Table V
The exponentially weighted moving average chart parameters for the solid phase competition enzyme-linked immunosorbent assay type A Iran 1996 based on a reference pool of 20 internal quality control (IQC) values (7 January to 10 March 2004)

IQC sample	Mean (m)	Standard deviation (s)	Action limits (ARL 370 and l 0.4)	Action limits (ARL 100 and l 0.5)
Cut-off	3.214	0.338	2.019 – 4.409	2.222 – 4.206
Weak positive	1.957	0.132	1.490 – 2.424	1.570 – 2.344
Negative	4.805	0.582	2.749 – 6.861	3.099 – 6.511
No serum (Co)	5.061	0.574	3.033 – 7.088	3.378 – 6.743

ARL: average run length

Quality control today is as important as when it was first described by Walter Shewhart in 1931 (28). Selection of one charting method and the supporting statistics over another comes down to choosing between two types of errors: false rejection (α or often denoted as type I error) and false acceptance (β or type II error). Although the decision criteria of the Shewhart charts, based on the Westgard rules, are appealing because of the simplicity (statistics and automated software programmes are not required and the data is easy to interpret), it is the authors' opinion that for this specific assay the EWMA charts provide the best balance between error detection and false rejection. Even when more stringent control limits for the EWMA charts are implemented, based on the Westgard algorithm (m ± 2.93 s), a lower false rejection rate is observed than for the Shewhart charts (m ± 3.58 s). Furthermore, in addition to being as sensitive as the conventional Shewhart charts, the EWMA charts are also more flexible since different smoothing factors can be selected to increase the trend and shift detection (4, 17). Moreover, the correlation between the different incorporated controls is also calculated and taken into account when accepting or rejecting plates and supplements common sense interpretation of the data. Therefore, based on the data generated on 28 June, 5 July plate 78, and 9 July, the risk of false acceptance is low when using the EWMA charts.

This paper promotes the use of control charts, in particular EWMA charts, as a tool for visualising potential assay problems. The charting method is easy to implement, and the parameters can be optimised to best suit the predetermined test requirements. Assay performance monitoring is not only useful for day to day IQC and test validation and accreditation, it is also essential for mutual recognition of test results between trading partners and could result in the elimination of trade barriers that may exist due to a lack of confidence in laboratory testing.

Conclusions

Once the three conventional stages of method validation (feasibility, development and standardisation, and characterisation of an assay performance) have been successfully finalised and the parameters of the assay have been characterised, constant monitoring, maintenance, and enhancement of the assay are imperative through the use of control charts and participation in proficiency testing schemes (12). Combining an IQC procedure based on EWMA charts with external quality assurance through the use of calibrated assays and the development of working standards/controls, as presented in this paper, will consistently provide statistically valid and reliable test results and increase the confidence of decision makers in the field.

Acknowledgements

This study was funded by the Federale Diensten voor Wetenschappelijke, Culturele en Technische Aangelegenheden. The technical assistance of Miss Sofie Diez (Veterinary and Agrochemical Research Centre, Brussels, Belgium) is greatly acknowledged.

L'assurance qualité et le contrôle de la qualité appliqués à l'épreuve immuno-enzymatique de compétition en phase solide pour le diagnostic de la fièvre aphteuse – II. Contrôle de la qualité : comparaison de deux méthodes cartographiques pour vérifier les performances des tests

N. Goris & K. De Clercq

Résumé
Les laboratoires de diagnostic étant désormais soumis à des normes de qualité de plus en plus strictes, leur accréditation officielle est soumise à une validation préalable des épreuves qu'ils utilisent. La validation des méthodes de test est généralement considérée comme achevée une fois établis les paramètres et les caractéristiques de l'épreuve. Or, comme tout processus, les épreuves de diagnostic sont sujettes à des variations aléatoires se traduisant par des altérations des valeurs moyennes du test. Une vérification permanente des tests au moyen de cartes de contrôle permet d'attirer l'attention de l'analyste sur une éventuelle variation des performances du test. Plusieurs méthodes cartographiques ont été mises au point à cette fin. Les auteurs comparent les cartes de contrôle de Shewhart et à moyenne mobile avec pondération exponentielle (EWMA) pour la vérification de routine de la qualité interne des échantillons utilisés pour l'épreuve immuno-enzymatique de compétition en phase solide. Si la sensibilité aux variations des deux méthodes s'avère équivalente, la carte EWMA semble offrir un meilleur équilibre entre les faux rejets et les fausses acceptations.

Mots-clés
Accréditation – Carte de contrôle – Contrôle de la qualité – Épreuve immuno-enzymatique – Fièvre aphteuse – Moyenne mobile à pondération exponentielle – Règle de Westgard – Shewhart – Validation.

Garantía y control de calidad del ensayo inmunoenzimático de competición en fase sólida para la fiebre aftosa. Parte II. Control de calidad: comparación de dos métodos gráficos para determinar el rendimiento del ensayo

N. Goris & K. De Clercq

Resumen
Los laboratorios de diagnóstico, a los que se exige cada vez más que cumplan rigurosos criterios de calidad, necesitan ensayos validados para conseguir una acreditación u homologación oficial. A menudo se considera que el proceso de validación de un método de prueba acaba cuando se definen los parámetros y características del ensayo de que se trate. Sin embargo, como cualquier otro

proceso, las pruebas de diagnóstico están sujetas a una parte de azar que genera variaciones en los valores medios de los resultados. El seguimiento continuo de los ensayos mediante gráficas de control advertirá al intérprete de todo cambio que se produzca en las propiedades de rendimiento. Con tal idea se han elaborado y aplicado diversos métodos gráficos. Los autores comparan sendas gráficas de control, la media móvil ponderada exponencialmente y la de Shewhart, con respecto al seguimiento cotidiano de muestras de control interno de calidad del ensayo inmunoenzimático de competición en fase sólida utilizado para diagnosticar la fiebre aftosa. Las dos clases de gráfica presentan igual sensibilidad a los cambios, pero la media móvil ponderada exponencialmente parece ofrecer un mejor equilibrio entre falsos negativos y falsos positivos.

Palabras clave
Acreditación – Control de calidad – Ensayo inmunoenzimático – Fiebre aftosa – Gráfica de control – Media móvil ponderada exponencialmente – Regla de Westgard – Shewhart – Validación.

References

1. Bergmann I.E., Malirat V., Neitzert E., Panizutti N., Sanchez C. & Falczuk A. (2000). – Improvement of serodiagnostic strategy for foot and mouth disease virus surveillance in cattle under systematic vaccination: a combined system of an indirect ELISA-3ABC with an enzyme-linked immunoelectrotransfer blot assay. *Arch. Virol.*, **145**, 473-489.

2. Blacksell S.D., Gleeson L.J., Lunt R.A. & Chamnanpood C. (1994). – Use of combined Shewhart-CUSUM control charts in internal quality control of enzyme-linked immunosorbent assays for the typing of foot-and-mouth disease virus antigen. *Rev. sci. tech. Off. int. Epiz.*, **13** (3), 687-699.

3. Callens M. & De Clercq K. (1997). – Differentiation of the seven serotypes of foot-and-mouth disease virus by reverse transcriptase polymerase chain reaction. *J. virol. Meth.*, **67** (1), 35-44.

4. Crowder S.V. (1989). – Design of exponentially weighted moving average schemes. *J. Qual. Technol.*, **21**, 155-162.

5. De Clercq K. (1998). – Implementation of quality assurance in national foot and mouth disease laboratories, based on the guidelines of the Office International des Epizooties. *Rev. sci. tech. Off. int. Epiz.*, **17** (3), 786-795.

6. De Diego M., Brocchi E., Mackay D. & De Simone F. (1997). – The use of the non-structural polyprotein 3ABC of FMD virus as a diagnostic antigen in ELISA to differentiate infected from vaccinated cattle. *Arch. Virol.*, **142** (10), 2021-2033.

7. Ferris N.P. & Dawson M. (1988). – Routine application of enzyme-linked immunosorbent assay in comparison with complement fixation for diagnosis of foot-and-mouth and swine vesicular diseases. *Vet. Microbiol.*, **16** (3), 201-209.

8. Goris N. & De Clercq K. (2005). – Quality assurance/quality control of foot and mouth disease solid phase competition enzyme-linked immunosorbent assay – Part I. Quality assurance: development of secondary and working standards. *Rev. sci. tech. Off. int. Epiz.*, **24** (3), 995-1004.

9. Hamblin C., Barnett I.T.R. & Hedger R.S. (1986). – A new enzyme-linked immunosorbent assay (ELISA) for the detection of antibodies against foot-and-mouth disease virus. I. Development and method of ELISA. *J. immunol. Meth.*, **93** (1), 115-121.

10. Hotelling H. (1947). – Multivariate quality control. *In* Techniques of statistical analysis (C. Eisenhart, M. Hastay & W.A. Wallis, eds). McGraw-Hill, New York, 111-184.

11. International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) (1999). – General requirements for the competence of testing and calibration laboratories, 1st Ed. ISO/IEC International Standard 17025, Geneva.

12. Jacobson R.H. (1998). – Validation of serological assays for diagnosis of infectious diseases. *Rev. sci. tech. Off. int. Epiz.*, **17** (2), 469-486.

13. Kitching P., Rendle T. & Newman B. (2000). – FAO collaborative study phase XVI: establishing reference standards for FMD serology. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 5-8 September, Borovets, Bulgaria. FAO, Rome, 224-230.

14. Levey S. & Jennings E.R. (1950). – The use of control charts in the clinical laboratories. *Am. J. clin. Pathol.*, **20**, 1059-1066.

15. Mackay D.K.J., Bulut A.N., Rendle T., Davidson F. & Ferris N. (2001). – A solid-phase competition ELISA for measuring antibody to foot-and-mouth disease virus. *J. virol. Meth.*, **97** (1-2), 33-48.

16. Mackay D.K.J., Rendle T. & Kitching R.P. (1998). – FAO collaborative study Phase XV: standardisation of FMD antibody detection. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 14-18 September, Aldershot, UK. FAO, Rome, 148-158.

17. Neubauer A.S. (1997). – The EWMA control chart: properties and comparison with other quality-control procedures by computer simulation. *Clin. Chem.*, **43** (4), 594-601.

18. OIE (World Organisation for Animal Health) (2004). – Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, 5th Ed. OIE, Paris.

19. Paiba G.A., Anderson J., Paton D.J., Soldan A.W., Alexandersen S., Corteyn M., Wilsden G., Hamblin P., Mackay D.K. & Donaldson A.I. (2004). – Validation of a foot-and-mouth disease antibody screening solid-phase competition ELISA (SPCE). *J. virol. Meth.*, **115** (2), 145-158.

20. Paton D.J. (2004). – Proposals for Phase XIX [abstract]. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 12-15 October, Chania, Crete. FAO, Rome, 94.

21. Paton D.J., Armstrong R.M. & Anderson J. (2003). – An update on progress with the FAO collaborative studies for FMD serology standardisation, phases XVII and XVIII. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 16-19 September, Gerzensee, Berne, Switzerland. FAO, Rome, 102-115.

22. Paton D.J., Armstrong R.M., Fernandez R., Hamblin P.A., Turner L., Corteyn M., Gibson D., Parida S., Wright C. & Anderson J. (2004). – FAO Phase XVIII FMD serological standardisation; progress and future prospects. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 12-15 October, Chania, Crete. FAO, Rome, 77-93.

23. Paton D.J., Armstrong R.M., Turner L.S., Hamblin P.A., Corteyn M., Gibson D. & Anderson J. (2002). – FAO collaborative study phase XVII: standardisation of FMD antibody detection. *In* Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the control of foot and mouth disease, 17-20 September, Cesme, Izmir, Turkey. FAO, Rome, 226-234.

24. Pearson D. & Cawte S.A. (1997). – Long-term quality control of DXA: a comparison of Shewhart rules and CUSUM charts. *Osteoporos Int.*, **7** (4), 338-343.

25. Rebeski D.E., Winger E.M., Ouma J.O., Kong Pages S., Büscher P., Sanogo Y., Dwinger R.H. & Crowther J.R. (2001). – Charting methods to monitor the operational performance of ELISA method for the detection of antibodies against trypanosomes. *Vet. Parasitol.*, **96** (1), 11-50.

26. Reid S.M., Forsyth M.A., Hutchings G.H. & Ferris N.P. (1998). – Comparison of reverse transcription polymerase chain reaction, enzyme-linked immunosorbent assay and virus isolation for the routine diagnosis of foot-and-mouth disease. *J. virol. Meth.*, **70** (2), 213-217.

27. Roberts S.W. (1959). – Control charts based on geometric moving averages. *Technometrics*, **1**, 239-250.

28. Shewhart W.A. (1931). – Economic control of quality of manufactured product. Van Nostrand, New York.

29. Westgard J.O., Barry P.L., Hunt M.R. & Groth T. (1981). – A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin. Chem.*, **27** (3), 493-501.

30. Westgard J.O., Groth T., Aronsson T. & de Verdier C.H. (1977). – Combined Shewhart-CUSUM control chart for improved quality control in clinical chemistry. *Clin. Chem.*, **23** (10), 1881-1887.

31. Wright P.F (1987). – Enzyme immunoassay: observations on aspects of quality control. *Vet. Immunol. Immunopathol.*, **17**, 441-452.

32. Wright P.F., Nilsson E., Van Rooij E.M.A., Lelenta M. & Jeggo M.H. (1993). – Standardisation and validation of enzyme-linked immunosorbent assay techniques for the detection of antibody in infectious disease diagnosis. *Rev. sci. tech. Off. int. Epiz.*, **12** (2), 435-450.

33. Wright P.F & Zhou E.-M. (1999). – Developments in international standardization. *Vet. Immunol. Immunopathol.*, **72** (1-2), 243-248.