Results of Continuous Monitoring of the Performance of Rubella Virus IgG and Hepatitis B Virus Surface Antibody Assays Using Trueness Controls in a Multicenter Trial

Tamara Kruk,a Sam Ratnam,a Jutta Preiksaitis,b Allan Lau,a Todd Hatchette,c Greg Horsman,d Paul Van Caeseele,a Brian Timmons,f and Graham Tipplesf

Surveillance and References Services, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba,a Provincial Laboratory for Public Health (Microbiology), Edmonton, Alberta,b Queen Elizabeth II Health Sciences Centre, Halifax, Nova Scotia,c Saskatchewan Disease Control Laboratory, Regina, Saskatchewan,d Cadham Provincial Laboratory, Winnipeg, Manitoba,e and Queen Elizabeth Hospital, Charlottetown, Prince Edward Island,f Canada

We conducted a multicenter trial in Canada to assess the value of using trueness controls (TC) for rubella virus IgG and hepatitis B virus surface antibody (anti-HBs) serology to determine test performance across laboratories over time. TC were obtained from a single source with known international units. Seven laboratories using different test systems and kit lots included the TC in routine assay runs of the analytes. TC measurements of 1,095 rubella virus IgG and 1,195 anti-HBs runs were plotted on Levey-Jennings control charts for individual laboratories and analyzed using a multirule quality control (MQC) scheme as well as a single three-standard-deviation (3-SD) rule. All rubella virus IgG TC results were “in control” in only one of the seven laboratories. Among the rest, “out-of-control” results ranged from 5.6% to 10% with an outlier at 20.3% by MQC and from 1.1% to 5.6% with an outlier at 13.4% by the 3-SD rule. All anti-HBs TC results were “in control” in only two laboratories. Among the rest, “out-of-control” results ranged from 3.3% to 7.9% with an outlier at 19.8% by MQC and from 0% to 3.3% with an outlier at 10.5% by the 3-SD rule. In conclusion, through the continuous monitoring of assay performance using TC and quality control rules, our trial detected significant intra- and interlaboratory, test system, and kit lot variations for both analytes. In most cases the assay rejections could be attributable to the laboratories rather than to kit lots. This has implications for routine diagnostic screening and clinical practice guidelines and underscores the value of using an approach as described above for continuous quality improvement in result reporting and harmonization for these analytes.

In clinical laboratories many different factors, such as test systems employing a variety of different measurement procedures, lot variations of test kits, equipment calibration, and malfunction and other technical and human errors can lead to shifts in the mean test values and affect the assay result and interpretation. Most laboratories subscribe to external proficiency testing programs to ascertain whether their testing method still works and can produce the specified result. While this allows for validating the performance of assays during specific time intervals, thus providing a snapshot of the quality of the assay, variations due to random and systematic errors caused by a variety of factors over time cannot be detected by proficiency testing, whereas continuous monitoring of assays using trueness controls (TC) in conjunction with control charts and the multirule quality control (MQC) scheme as proposed by Westgard et al. (16) can alert workers to changes in assay performance in real time. Such an approach will ensure that corrective measures are taken immediately and errors do not propagate. This is essential in maintaining quality assurance for diagnostic assays and the reproducibility of test results.

International public health guidelines for rubella virus and hepatitis B virus serology have stated that a quantitative level of 10 IU/ml of rubella virus IgG and 10 mIU/ml of hepatitis B virus surface antibody (anti-HBs) measured in a serum sample correlates with a level of immunity that protects from disease (7, 14). These guidelines are based on the assumptions that the quantitative result of analyte measurement is highly reproducible in any measurement procedure and that when an international reference standard is used for calibration, results among different measurement procedures are harmonized (11). In clinical laboratories, rubella virus IgG and anti-HBs are assayed using a variety of commercial systems and test kits with different measurement procedures. While kit controls are included in each run, these may differ from kit to kit and between test systems, and they often do not target quantitative values that correlate with clinically relevant trigger points. Inclusion of standardized external TC traceable to international reference standards in routine assay runs across many laboratories using the same or different commercial systems and test kits can allow for detecting intra- and interlaboratory as well as intra- and inter-measurement system and lot issues that result in variations in result reporting. This may be helpful in monitoring, ensuring, and improving the reproducibility and harmonization of test results on a regional or national basis.

The Canadian National Microbiology Laboratory (NML) in conjunction with the Canadian Public Health Laboratory Network (CPHLN) initiated a multicenter trial to assess the usefulness of including standardized TC in routine runs of rubella virus IgG and anti-HBs assays for continuously monitoring the performance of the assays. The specific purpose of the trial was to assess the extent of variation, if any, between laboratories, test systems,
and kit lots used over time. The TC were bulk purchased and distributed to participating laboratories across the country for inclusion in routine assay runs. The TC measurement results were collected systematically over a period of time, plotted on Levey-Jennings control charts (10) for individual laboratories, and analyzed retrospectively using the MQC scheme (16) to determine assay performance. The results were also analyzed based on a single three-standard-deviation (3-SD or 13σ) rule. In this paper we describe our observations and provide insight on the potential value of continuously monitoring assay performance using TC and utilizing the measurement values of these controls to determine assay run acceptance and rejection.

MATERIALS AND METHODS

Trial outline. The trial utilized standardized TC for rubella virus IgG and anti-HBs traceable to the following World Health Organization (WHO) international reference standards: the first international standard for anti-rubella virus immunoglobulin, human (RUB1-1-94), and the first international standard for hepatitis B virus immunoglobulin (W1042). The concentrations of the rubella virus IgG and anti-HBs TC were 7.2 IU/ml and 9.1 mIU/ml, respectively, determined by calibration of the TC materials to the international reference standard using the Abbott AxSYM system (Abbott Laboratories, North Chicago, IL). The TC materials were custom manufactured from human plasma pools by Acrometrix (Benicia, CA) for the trial and bulk purchased. Seven laboratories comprising both large provincial public health laboratories and smaller hospital-based laboratories across Canada that provide rubella and hepatitis virus serology volunteered to participate in the trial. The participating laboratories were provided with both rubella virus IgG and anti-HBs TC and instructed to include them in routine assay runs for the analytes. Both batch and random access testing were performed by participants, and the TC were included in assay runs once per testing day on a continuous basis. This allowed for monitoring the performance of multiple lots of the test kits used for routine testing during the trial. The results of the TC measurements together with the lot numbers were reported to the NML. The purchase and distribution of the TC and maintenance of a central independent database were undertaken by the NML.

Assay systems. For rubella virus IgG measurement, two different commercial assay systems were used among the seven participating laboratories. Six laboratories utilized the Abbott AxSYM system, and one used the Advia Centaur system (Siemens Healthcare Diagnostics, Los Angeles, CA). For anti-HBs measurement, four different assay systems were used. Four laboratories utilized Abbott AxSYM, two laboratories Abbott Architect, and one laboratory Abbott iMx. One of the laboratories also used Advia Centaur in addition to Abbott Architect. All assays were carried out per standard operating procedures in the participating laboratories and in accordance with the manufacturers’ instructions.

Data analysis. Data were collected during a period of 6 to 8 months in 2006 to 2007 and analyzed retrospectively. Each laboratory used multiple lots of the rubella virus IgG and anti-HBs TC test kits during this time. Measurement results for TC from all laboratories were collated in a central database and, regardless of the kit lots used, were plotted for each laboratory on Levey-Jennings control charts as previously described (10). The means and standard deviations (SD) in the range of ±1 SD, ±2 SD, and ±3 SD were determined for the TC results for each laboratory and added to the charted data. The means and SDs were calculated based on the results reported in the first 30 runs of the TC. The MQC scheme developed by Westgard et al. (16) was used to analyze the individual TC measurement results to determine whether the TC result would designate the run “in control” or “out of control.” We also analyzed the overall data based on a single 3-SD rule. The MQC scheme and its application in our study are briefly described below.

MQC. The MQC scheme is based on a statistical control procedure and has elements that allow for the detection of both random and systematic errors in clinical laboratories. It assumes that the distribution of errors can be described by the mean and standard deviation. If the result for the control is within 2 SD of the mean, it is deemed “in control” and the assay run is accepted. A general initial rule, 1σ, where one control observation exceeds the mean by 2 SD, is used as a warning rule and prompts a more detailed inspection of the control observations using the following rules to determine whether the assay run should be accepted or rejected. The assay is deemed “out of control” if any of the following occurs: (i) one control observation is more than 3 SD from the mean (13σ), (ii) two consecutive control observations are more than 2 SD from the mean (2σ), (iii) four consecutive control observations are more than 1 SD from the mean (4σ), or (iv) 10 consecutive control observations are on the same side of the mean (10σ). That is, if a control observation exceeds a 2-SD limit, it is tested by applying the 1σ, 2σ, 4σ, and 10σ rules. If none of these rules is violated, the run is in control, and if any of them is violated, the run is out of control. For our analysis based on MQC, we determined each rule violated per 1σ observation. The probability associated with each rule has been described elsewhere (17), with the suggestion that a logical sequence of control rules be used in combination. The 1σ rule most often detects random errors, and 2σ, 4σ, and 10σ usually detect systematic errors (16); in the analysis of our results, we assumed these to be the case in all instances.

RESULTS

Rubella virus IgG. A total of 1,095 rubella virus IgG TC results measured using two different test systems, Abbott AxSYM and Advia Centaur, were available for analysis. All assay runs in which these results were produced were reported as valid by the seven participating laboratories based on the manufacturers’ protocols for interpretation of the test run. The mean of TC measurements based on the initial 30 assay runs obtained by the participants using the two test systems ranged from 5.6 to 15.6 IU/ml, with SD ranging from 0.7 to 1.6.

Six participating laboratories using the AxSYM system utilized 11 different lots of the assay kits, with 10 of the lots utilized by two or more laboratories. The laboratory- and lot-specific TC measurement results as analyzed by MQC are summarized in Table 1. While the overall results were found to be satisfactory, with no rejections in the majority of the assay runs by individual laboratories, some runs would have been deemed out of control with both random and systematic errors. However, the rejection rate would have been significantly higher for laboratory A. Specifically, of the four lots (3407M101, 36925M101, 40038M200, and 42632M100) used by at least four of the six laboratories, there would have been rejections in the assay runs with three of the lots, and these were limited to laboratories A, C, E, and F. Laboratory D tested the TC in a continuous series of 85 assay runs using three of the same four lots in addition to four other lots and would have had no rejections. Figure 1 shows the control chart of laboratory D, which would have had no assay run rejections, in comparison with that of laboratory A, which would have experienced significantly higher assay run rejections, both of which used the same AxSYM system and different kit lots. This indicated both intra- and interlaboratory variations. Since some of the kit lots used across laboratories were the same, it may be concluded that the variation observed was not related to reagents supplied in specific kit lots but was likely due to issues related to instrument calibration or other errors made in the measurement procedure within individual laboratories. Regardless, three of the four laboratories using one particular lot (42632M100) would have experienced problems with out-of-control results (Table 1).

Laboratory A used two different AxSYM instruments for ru-
bella virus IgG testing, utilizing the same two lots of kits on both, and reported results obtained with both instruments. In this series, there would have been rejections in the assay runs with both kit lots, with a higher number of rejections occurring with one instrument, with both random and systematic errors identified (Table 2). However, laboratory B, using one of the same kit lots (34241M201) for 24 runs, would not have had any rejections (Table 1). This indicated intra- and interlaboratory variations in the

| Laboratory code (test system used) | Kit lot no. | No. of assay results | 1<sub>2s</sub> | 1<sub>3s</sub> | 2<sub>2s</sub> | 4<sub>1s</sub> | 10<sub>1s</sub> |
|-----------------------------------|------------|---------------------|-------------|-------------|-------------|-------------|-------------|
| A<sup>d</sup> (Abbott AxSYM)       | 34241M201  | 150                 | 32          | 22          | 18          | 16          | 12          |
|                                   | 38592M201  | 2                   | 2           | 1           | 1           |             |             |
|                                   | 40067M100  | 267                 | 51          | 33          | 26          | 22          | 12          |
| Total                             | 419        | 85 (20.3)           | 56 (13.4)   | 45          | 38          | 24          |
| B (Abbott AxSYM)                  | 34241M201  | 24                  | 0           |             |             |             |             |
|                                   | 36925M101  | 25                  | 0           |             |             |             |             |
|                                   | 38592M201  | 47                  | 2           | 1           | 1           |             |             |
|                                   | 39294M100  | 19                  | 4           | 3           | 1           | 1           | 2           |
|                                   | 40038M201  | 27                  | 9           | 5           | 5           | 1           |             |
|                                   | 41287M101  | 18                  | 1           |             |             |             |             |
| Total                             | 160        | 16 (10.0)           | 9 (5.6)     | 6           | 3           | 3           |
| C (Abbott AxSYM)                  | 34241M201  | 2                   | 0           |             |             |             |             |
|                                   | 36925M101  | 9                   | 1           | 1           | 0           | 0           |             |
|                                   | 38595M200  | 5                   | 0           |             |             |             |             |
|                                   | 40038M200  | 6                   | 0           |             |             |             |             |
|                                   | 40038M201  | 9                   | 0           |             |             |             |             |
|                                   | 42632M100  | 7                   | 2           | 0           | 1           |             |             |
| Total                             | 38         | 3 (7.9)             | 1 (2.6)     | 1           | 2           |             |             |
| D (Abbott AxSYM)                  | 36925M101  | 16                  | 0           |             |             |             |             |
|                                   | 38466M200  | 15                  | 0           |             |             |             |             |
|                                   | 40038M200  | 23                  | 0           |             |             |             |             |
|                                   | 41287M100  | 11                  | 0           |             |             |             |             |
|                                   | 42632M100  | 16                  | 0           |             |             |             |             |
|                                   | 44031M100  | 3                   | 0           |             |             |             |             |
|                                   | 44033M100  | 1                   | 0           |             |             |             |             |
| Total                             | 85         | 0                   |             |             |             |             |             |
| E (Abbott AxSYM)                  | 34241M201  | 4                   | 0           |             |             |             |             |
|                                   | 36925M101  | 19                  | 0           |             |             |             |             |
|                                   | 38595M200  | 26                  | 2           | 1           | 1           | 2           | 0           |
|                                   | 39294M100  | 25                  | 1           | 0           | 0           | 1           | 0           |
|                                   | 40038M200  | 24                  | 0           |             |             |             |             |
|                                   | 41287M100  | 17                  | 3           | 1           | 1           | 2           | 1           |
|                                   | 42632M100  | 41                  | 5           | 0           | 3           | 2           | 0           |
| Total                             | 136        | 11 (7.1)            | 2 (1.3)     | 5           | 7           | 1           |
| F (Abbott AxSYM)                  | 40038M200  | 13                  | 0           |             |             |             |             |
|                                   | 41287M101  | 18                  | 0           |             |             |             |             |
|                                   | 42632M100  | 19                  | 3           | 1           | 2           | 2           | 0           |
|                                   | 44031M101  | 17                  | 1           | 0           | 0           | 1           | 0           |
|                                   | 44033M101  | 19                  | 1           | 0           | 0           | 0           | 1           |
|                                   | 44034M100  | 3                   | 0           |             |             |             |             |
| Total                             | 89         | 5 (5.6)             | 1 (1.1)     | 2           | 3           | 1           |
| G (Advia Centaur)                 | 139        | 16                  | 0           |             |             |             |             |
|                                   | 140        | 85                  | 7           | 4           | 5           | 3           | 4           |
|                                   | 142        | 47                  | 4           | 1           | 1           | 1           |             |
| Total                             | 148        | 11 (7.5)            | 5 (3.4)     | 6           | 4           | 5           |

<sup>a</sup> A trueness control was included in assay runs utilizing different lots of AxSYM and Advia kits.

<sup>b</sup> Number (percentage) of assay runs with the 1<sub>2s</sub> warning rule which violated one or more MQCs.

<sup>c</sup> Same as 3 SD referred to in the text.

<sup>d</sup> Results for laboratory A are the total number of assay results produced by two instruments.
results obtained by laboratories A and B. The variation observed by laboratory A could be attributable to problems with both instruments, although one performed better than the other, regardless of the lots used. Laboratory G used Advia Centaur with three lots of the kits, and interlaboratory variation of the assay, if any, could not be determined as there were no other laboratories using the same system. There would have been rejections in the assay runs with two of the three lots used by laboratory G (Table 1). This indicated both intralaboratory and intra- and interlot variations.

Based on MQC analysis of the total number of TC measurement results reported by each of the participating laboratories, the rejection rate for assay runs would have been 20.3% for laboratory A, and it ranged from 5.6% to 10.0% among the five other laboratories (Table 1). The mean rubella virus IgG of the total number of TC measurements obtained with the two assay systems was 7.2 IU/ml (SD, 0.8 IU/ml) for AxSYM and 15.6 IU/ml (SD, 1.0 IU/ml) for Advia.

Anti-HBs assay. A total of 1,195 anti-HBs TC results measured using four different test systems, Abbott AxSYM, Abbott Architect, Abbott IMx, and Advia Centaur, were available for analysis. All assay runs in which these results were produced were reported as valid by the seven participating laboratories based on the manufacturers’ protocols for interpretation of the test run. The mean of TC measurements based on the initial 30 assay runs obtained by the participants using the four test systems ranged from 6.7 to 11.1 mIU/ml, with the SD ranging from 0.4 to 1.8.

With AxSYM, four laboratories used 17 different lots of anti-
HBs kits, and laboratories C and F would have had no rejections of assay runs as analyzed by MQC (Table 3). Of the 17 lots, only four lots (34201LU00, 40295LU02, 41116LU00, and 41273LU01) were used by at least two laboratories. Based on the assays performed using these four lots, there would have been rejections with only one of the lots (41273LU01) and only in the assay runs performed by laboratory D. Overall, the rejection rate of assay runs in laboratory D would have been greater than those in other three laboratories using the AxSYM system regardless of the lot numbers used (Table 3). The control chart for laboratory D showed a trending effect, and this was likely due to instrument error, perhaps combined with lot variations (Fig. 2). These data indicated intra- and interlaboratory variation in assay runs with both random and systematic errors. As with the rubella virus IgG assay, laboratory A used two different AxSYM instruments for anti-HBs detection, using the same three lots of the kits on both. There would have been assay rejections with one of the three lots, with the same instrument that had higher rejection rates in the rubella virus IgG also showing a higher rate of rejection for anti-HBs (Table 2). This highlighted the possibility of an instrument-specific variability in assay performance.

In two laboratories (E and G) using the Architect system, there would have been assay rejections attributable to both random and systematic errors. Since none of the kit lots were utilized by both laboratories, these results could not be directly compared. The control chart for laboratory E showed three distinct TC result clusters, with each cluster associated with a specific kit lot (Fig. 3). Laboratory G also exhibited a trending effect within a single lot, likely attributable to a systematic error. Laboratory G, besides using Architect, also used the Advia system. There would have been rejections in the assay runs with two of the three Advia kit lots. However, the rate of rejection of the Advia assay runs would have been lower than that of the Architect assay runs (Table 3). This suggested variation in assay runs that was test system dependent. Laboratory B was the only one that used the IMx system for anti-HBs detection. Of the seven lots utilized, there would have been rejections in assays performed with two lots, involving both random and systematic errors. This indicated intraintra- and interlot variations in the assays performed. The control chart for this laboratory showed that one kit lot might have led to mean TC measurement results that were higher than those observed for other kit lots.

Based on MQC analysis of the total number of TC measurement results reported by each of the participating laboratories, the assay run rejection rate would have been 19.8% for laboratory D, and it ranged from 3.3% to 7.9% among the four other laboratories that would have had assay run rejections. Based on the single 3-SD rule (Table 3, MQC columns, 13s), the assay run rejection rate would have been 10.5% for laboratory D, and it ranged from 0% to 3.3% for the four other laboratories (Table 3). The mean anti-HBs of the total number of TC measurements obtained with the four assay systems was 9.8 mIU/ml (SD, 0.8 mIU/ml) for AxSYM, 7.3 mIU/ml (SD, 0.6 mIU/ml) for Architect, 11.1 mIU/ml (SD, 1.3 mIU/ml) for IMx, and 10.4 mIU/ml (SD, 1.4 mIU/ml) for Advia.

**DISCUSSION**

In this study, we determined the value of routinely including the same standardized TC in assay runs measuring rubella virus IgG and anti-HBs as a broad quality control measure across multiple laboratories. To analyze the study data, we utilized an MQC scheme, as this is recommended when a stable reference control is used and analyzed repeatedly (16). This type of control procedure was initially described by Shewhart, who introduced the concept of 3-SD (13s) limits for quality control (13). Levey and Jennings (10) adapted this idea to clinical laboratory practice, with Westgard et al. (16) ultimately formulating the MQC scheme to decide whether the measurement results for the quality control material are “in control” and the results from that run are reportable or whether they are “out of control” and rejected. It was indicated that the combination of 13s, 22s, 41s, and 10s rules was better for error detection than 13s alone. Since there is a basis to consider that the use of the 13s single rule alone could be adequate to indicate a real problem with clinical laboratory tests (3), we also applied the 3-SD rule to analyze our study data. Further, whereas the MQC scheme is generally used with 2 or 4 control measurements per run, we utilized TC once per run, as this study was considered a pilot trial to generate some preliminary data to ascertain the extent of variations, if any.

Our study data indicated intra- and interlaboratory, intra- and inter-assay system, and kit lot variation in assay runs of both rubella virus IgG and anti-HBs as judged by both MQC and the 3-SD rules. This occurred despite the fact that all laboratories routinely participated in proficiency testing programs and passed them during the year of the study. Through a retrospective query, it was found that at least two of the participating laboratories included either in-house or other commercial controls in assay runs at the time of this trial. However, the data relating to these control re-

---

**Table 2** MQC analysis of instrument-specific trueness control measurement results for rubella virus IgG and anti-HBs in a single laboratory

| Abbott AxSYM assay | AxSYM instrument 1 | AxSYM instrument 2 |
|-------------------|--------------------|--------------------|
|                    | No. of assay results | No. of MQC violations by: | No. of assay results | No. of MQC violations by: |
| Rubella virus IgG  |                     |                     |                     |
| 34241M201         | 75                  | 12                  | 75                  | 24                  |
| 40067M100         | 138                 | 16                  | 132                 | 35                  |
| Anti-HBs          |                     |                     |                     |
| 34201LU00         | 13                  | 0                   | 15                  | 0                   |
| 42462LU00         | 55                  | 1                   | 57                  | 1                  |
| 40295LU02         | 80                  | 0                   | 42                  | 0                   |

---

1. **Laboratory A utilized two Abbott AxSYM instruments and performed assays using two lots of rubella virus IgG kits and three lots of anti-HBs kits.
2. **Number of assay runs with the 12s warning rule which violated one or more MQCs.
3. **Same as 3 SD referred to in the text.
| Laboratory code (test system used) | Kit lot no. | No. of assay results | 1s | 2s | 3s | 4s | 10s |
|----------------------------------|------------|----------------------|----|----|----|----|-----|
| A (Abbott AxSYM)                | 34201LU00  | 28                   | 0  |    |    |    |     |
|                                 | 40295LU02  | 121                  | 0  |    |    |    |     |
|                                 | 42462LU00  | 112                  | 6  | 1  | 1  | 1  | 3   |
|                                 | 38082LU01  | 57                   | 8  | 1  | 5  | 8  | 6   |
| Total                           | 318        | 14 (4.4)             | 2  | 0.6| 6  | 9  | 9   |
| C (Abbott AxSYM)                | 34201LU00  | 2                    | 0  |    |    |    |     |
|                                 | 38082LU02  | 6                    | 0  |    |    |    |     |
|                                 | 40295LU00  | 6                    | 0  |    |    |    |     |
|                                 | 40295LU02  | 6                    | 0  |    |    |    |     |
|                                 | 41116LU00  | 4                    | 0  |    |    |    |     |
|                                 | 43642LU00  | 4                    | 0  |    |    |    |     |
|                                 | 41273LU01  | 6                    | 0  |    |    |    |     |
| Total                           | 34         | 0                    |    |    |    |    |     |
| D (Abbott AxSYM)                | 30295LU00  | 11                   | 0  |    |    |    |     |
|                                 | 35356LU01  | 20                   | 0  |    |    |    |     |
|                                 | 39488LU00  | 21                   | 5  | 2  | 2  | 1  | 0   |
|                                 | 41116LU00  | 11                   | 0  |    |    |    |     |
|                                 | 43642LU00  | 14                   | 6  | 3  | 4  | 6  | 3   |
|                                 | 41273LU01  | 8                    | 5  | 3  | 3  | 1  |     |
|                                 | 45070LU00  | 1                    | 1  | 1  |    |    |     |
| Total                           | 86         | 17 (19.8)            | 9  | 10.5| 9  | 8  | 3   |
| F (Abbott AxSYM)                | 41116LU00  | 4                    | 0  |    |    |    |     |
|                                 | 41116LU02  | 23                   | 0  |    |    |    |     |
|                                 | 41273LU01  | 21                   | 0  |    |    |    |     |
|                                 | 43640LU00  | 19                   | 0  |    |    |    |     |
|                                 | 46236LU00  | 18                   | 0  |    |    |    |     |
|                                 | 46635LU01  | 3                    | 0  |    |    |    |     |
| Total                           | 88         | 0                    |    |    |    |    |     |
| E (Abbott Architect)            | 34440M100  | 3                    | 0  |    |    |    |     |
|                                 | 35892M100  | 29                   | 0  |    |    |    |     |
|                                 | 37114M100  | 49                   | 7  | 0  | 1  | 3  | 6   |
|                                 | 38090M200  | 38                   | 2  | 0  | 0  | 1  | 2   |
|                                 | 40210M100  | 5                    | 0  |    |    |    |     |
|                                 | 41474M100  | 42                   | 3  | 0  | 0  | 2  | 2   |
| Total                           | 166        | 12 (7.2)             | 0  | 1  | 6  | 10  |
| G (Abbott Architect)            | 33853M100  | 85                   | 1  | 1  | 0  | 0  | 0   |
|                                 | 40016M100  | 94                   | 9  | 0  | 4  | 2  | 8   |
|                                 | 40016M101  | 12                   | 5  | 1  | 3  | 2  | 1   |
| Total                           | 191        | 15 (7.9)             | 2  | 1.0| 7  | 4  | 9   |
| G (Advia Centaur)               | 146        | 113                  | 4  | 4  | 0  | 0  | 0   |
|                                 | 150        | 25                   | 1  | 1  | 0  | 0  | 1   |
|                                 | 153        | 15                   | 0  |    |    |    |     |
| Total                           | 153        | 5 (3.3)              | 5  | 3.3| 0  | 0  | 1   |
| B (Abbott IMx)                  | 37348LU00  | 13                   | 0  |    |    |    |     |
|                                 | 38344LU00  | 37                   | 0  |    |    |    |     |
|                                 | 40213LU01  | 31                   | 3  | 0  | 1  | 0  | 3   |
|                                 | 41178LU00  | 25                   | 0  |    |    |    |     |
|                                 | 43087LU00  | 14                   | 0  |    |    |    |     |
|                                 | 44035LU00  | 27                   | 4  | 1  | 0  | 2  | 3   |
|                                 | 45060LU00  | 12                   | 0  | 0  | 0  | 0  |     |
| Total                           | 159        | 7 (4.4)              | 1  | 0.6| 1  | 2  | 6   |

*a* A trueness control was included in assay runs utilizing various kit lots of different test systems.

*b* Number (percentage) of assay runs with the 12s warning rule which violated one or more MQCs.

*c* Same as 3 SD referred to in the text.

*d* Results for laboratory A are the total number of assay results produced by two instruments.
results were not gathered as part of the trial. Regardless, all assay runs were reported as valid based on the manufacturers’ test protocols and interpretations. It is worth noting that laboratory D would have had no rejections of rubella virus IgG assay runs based on TC measurement results but would have had the highest assay run rejection rate for anti-HBs. The reverse was true for laboratories C and F, both of which would have had no rejections with anti-HBs assay runs but would have had rejected rubella virus IgG assay runs. Interestingly, all of them used the same Abbott AxSYM system for both analytes. Our study indicated that these variations can be detectable only by the continuous monitoring of the performance of assays with the use of common TC and applying an appropriate quality control measure such as MQC or the 3-SD rule.

Bishop and Nix (1) suggested that changes in kit lots of a given assay did not influence the detection of persistent errors through successive runs. In most cases, our data confirm this. We analyzed TC results for individual laboratories broken down by kit lots used so that direct comparisons between laboratories that used the same lots could be made. This allowed for the discernment of whether the assay rejections were due to variations within the laboratory or whether they were due to a problem with the kit lot. In most cases, the assay rejections could be attributable to the laboratories rather than kit lots. Our observations illustrate the

FIG 2 Anti-HBs trueness control measurement results for laboratory D, showing a trending effect. The numbers within the control chart are lot numbers.

FIG 3 Anti-HBs trueness control measurements for laboratory E, showing a trending effect associated with different lots of the test kit. The numbers within the control chart are lot numbers.
The power of the quality control approach we used for distinguishing between these two possibilities and supporting postmarketing surveillance of the quality of commercial test reagents. Problems with the functioning or calibration of instruments could lead to large random fluctuations of the data and can also result in systematic drifting. An example of this was found with laboratory A, where two AxSYM instruments were used for both rubella virus IgG and anti-HBs measurements. The assay run rejection rate would have been higher for rubella virus IgG than for anti-HBs as a whole and would have been greater with one instrument (Table 2). Anti-HBs TC results in laboratories D, E, and G all showed a trending effect below and above the mean, which may have been due to instrument error and lot variations, and these were laboratory specific. Trending or clustering of results that are not close to the calculated mean may lead to both false rejections and acceptance of assay runs.

The application of MQC for our data analysis allowed detection of both random and systematic errors. However, the use of the MQC for routine interpretation of quality control results in serological assay runs may be rather cumbersome and poses practical issues, especially when dealing with a combination of multiple rules. Therefore, one might consider using just the single 3-SD rule, as originally suggested by Shewhart (13) and as recommended by Carroll et al. (3). Also, with the application of the 3-SD rule, the rejection rates will remain within an acceptable range in a routine clinical laboratory setting. The MQC scheme was originally developed for clinical chemistry (16), and how applicable this approach is for quality control of serological tests in clinical microbiology is unknown. Regardless, our data confirm that TC can be very beneficial when run in addition to kit controls to determine assay run validity. In this study, we analyzed the results retrospectively. A web-based system that would allow real-time TC data entry and automated real-time charting and intra- and interlaboratory analysis has even greater potential to increase the pace of quality improvement and test result harmonization.

In this study, we based our analysis on using the mean and SD of the TC results measured from the first 30 runs of the assay that each laboratory performed. As this may pose some practical difficulty, we also looked at the possibility of reducing the number of assays required for this purpose. We found that using the mean and SD derived from the first 5 and 10 runs of the assay would lead to a higher number of rejections, as the values derived from fewer runs tend to exhibit greater variations (data not shown). Therefore, we suggest using the mean and SD measured from a higher number assay runs, and this will require securing an adequate supply of TC to last for several hundred assay runs. In this context, it is notable that the TC we supplied had a 2-year expiry.

The TC for both rubella virus IgG and anti-HBs were calibrated to specific unit values using only the Abbott AxSYM, and these values were reproduced in assays performed on the Abbott AxSYM. However, differences in quantitative results were observed using other systems, as has been observed by others for multiple platforms and kit lot numbers at the point of manufacturing. However, this could not be done due to a number of reasons, and this is considered a limitation of the study. There are inherent variations in precision as well as lot-to-lot variations, as noted in the package inserts of both the AxSYM rubella virus IgG and anti-HBs assays, and these tend to be higher at lower concentrations of the antibodies detected. These could have played a part, alone or in combination with a variety of other factors, in variations in reporting. International reference standards were created to harmonize test results across different measurement procedures. Recently it has been recognized that in order for this to occur, reference material must be commutable (11, 15). Like many other reference standards, the rubella virus IgG and anti-HBs international reference standards have not been formally tested for commutability for any specific method. Although calibration inconsistencies relative to the international standards may account for the variation in results between different systems, it is also possible that the international reference standard is not commutable for some methods because of matrix effects and cannot be used for calibrating some assays. Antibody is a complex biological material which may vary in individual patients depending on exposure to vaccine, natural infection, epitope specificity, and avidity related to stage of infection. Variation in the results of test systems may occur because not all current assay methods are actually measuring the same biological form of antibody, and therefore, harmonization of test results may not be possible, as has been suggested (6). Until further inter-assay system harmonization is achieved, it is important to establish a test system-specific range and the mean for TC or other quality control reagents used for monitoring the assay performance.

The observations made in our trial have significant clinical implications. Recently, investigators have questioned the usefulness of using serology for detecting immunity to rubella virus (2) and hepatitis B virus (7) in highly immunized populations. Based on our observations, a subject being tested for either rubella virus IgG or anti-HBs, in years remote from immunization, could easily be classified as immune or nonimmune depending on the day that the testing is performed in any specific laboratory. Discordant classification may be even more likely if different assay systems in different laboratories are used to screen a population. Clinicians and public health practitioners should be aware of the limitations, as well as the uncertainty associated with rubella virus and anti-HBs laboratory testing as correlates of immunity, and consider using a single test system for public health screening programs within their jurisdictions. In revising clinical guidelines, laboratory testing result correlates of immunity should more accurately reflect the precision, reproducibility, and inter-test system variability in result reporting, even in the presence of an international reference standard. Our data reinforce the need for clinical laboratories to include uncertainty as part of their reported measurement result for rubella virus IgG and anti-HBs. This suggests that laboratories should include a “gray zone” or indeterminate range in their reporting of results and avoid interpretative comments that classify patients as immune or nonimmune when this is inconsistent with the precision and reproducibility of the assay. In this regard, the use of the TC may be helpful to define a gray zone besides helping to determine assay run validity.

ACKNOWLEDGMENTS
We thank Tracy Hickson, Chaleur Regional Hospital, Bathurst, New Brunswick, Canada, for his contributions to the study and Wayne Dimech, National Serology Reference Laboratory, Fitzroy, Victoria, Australia, for reviewing the manuscript and helpful suggestions.

REFERENCES
1. Bishop J, Nix ABJ. 1993. Comparison of quality-control rules used in clinical chemistry laboratories. Clin. Chem. 39:1638–1649.
2. Byrne L, Brant L, Reynolds C, Ramsay M. 2012. Seroprevalence of low rubella IgG antibody levels among antenatal women in England tested by NHS Blood and Transplant: 2004–2009. Is rubella susceptibility increasing? Vaccine 30:161–167.

3. Carroll TA, Pinnick HA, Carroll WE. 2003. Probability and the Westgard rules. Ann. Clin. Lab. Sci. 33:113–114.

4. Dimech W, et al. 1992. Multicenter evaluation of five commercial rubella virus immunoglobulin G kits which report in international units per milliliter. J. Clin. Microbiol. 30:633–641.

5. Ferguson M, Yu MW, Heath A. 2010. Calibration of the second international standard for hepatitis B immunoglobulin in an international collaborative study. Vox Sang. 99:77–84.

6. Heijtink RA, Schneberger PM, Postma B, Crombach W. 2002. Anti-HBs levels after hepatitis B immunisation depend on test reagents: routinely determined 10 and 100 IU/L seroprotection levels unreliable. Vaccine 20:2899–2905.

7. Huang LM, Lu CY, Chen DS. 2011. Hepatitis B virus infection, its sequelae, and prevention by vaccination. Curr. Opin. Immunol. 23:237–243.

8. Huzly D, Schenk T, Jilg W, Neumann-Haefelin D. 2008. Comparison of nine commercially available assays for quantification of antibody response to hepatitis B virus surface antigen. J. Clin. Microbiol. 46:1298–1306.

9. Kinn S, Akhavan S, Agut H, Thibault V. 2011. Performance of the DiaSorin LIAISON® anti-HBs II for the detection of hepatitis B surface antibodies: comparison with the Abbott Architect anti-HBs assay. J. Clin. Virol. 50:297–302.

10. Levey S, Jennings ER. 1950. The use of control charts in the clinical laboratory. Am. J. Clin. Pathol. 20:1059–1066.

11. Miller WG, et al. 2011. Roadmap for harmonization of clinical laboratory measurement procedures. Clin. Chem. 57:1108–1117.

12. Oliveira LC, Okay TS, Kawasato KH, Lima LP. 2007. Comparison of two commercial immunoassays for the detection of anti-rubella IgM and IgG in pediatric patients. J. Imunnoasay Immunochem. 28:297–306.

13. Shewhart WA. 1931. Economic control of quality of manufactured products. D. Van Nostrand Company, Inc., New York, NY.

14. Skendzel LP. 1996. Rubella immunity. Defining the level of protective antibody. Am. J. Clin. Pathol. 106:170–174.

15. Vesper HW, Miller WG, Myers GL. 2007. Reference materials and commutability. Clin. Biochem. Rev. 28:139–147.

16. Westgard JO, Barry PL, Hunt MR, Groth T. 1981. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin. Chem. 27:493–501.

17. Westgard JO, Groth T, Aronsson T, Falk H, de Verdier CH. 1977. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. Clin. Chem. 23:1857–1867.