Quality Control of Multichannel Hematology Analyzers: Evaluation of Bull's Algorithm

Article *in* American Journal of Clinical Pathology · April 1985

DOI: 10.1093/ajcp/83.3.337 · Source: PubMed

2 authors:

George S. Cembrowski
Alberta Health Services
217 PUBLICATIONS **2,826** CITATIONS
SEE PROFILE

James O. Westgard
University of Wisconsin–Madison
189 PUBLICATIONS **5,494** CITATIONS
SEE PROFILE

Quality Control of Multichannel Hematology Analyzers: Evaluation of Bull's Algorithm

GEORGE S. CEMBROWSKI, M.D., PH.D. AND JAMES O. WESTGARD, PH.D.

Bull's algorithm has been evaluated by computer simulation studies. Varying amounts of systematic analytic error were simulated in either hemoglobin (Hgb), red blood cell count (RBC), or mean corpuscular volume (MCV) with the resulting red blood cell indices averaged in batches of 20 using Bull's algorithm. The number of average indices outside the limits of 0.97 $\bar{x}$ and 1.03 $\bar{x}$ ($\bar{x}$ = stable patient mean index) was tabulated and plotted against the size of the systematic shift, expressed in multiples of the long-term analytic standard deviation (SD). The resulting plots, called power functions, show that Bull's algorithm can detect large shifts effectively and that its power increases with increasing batch number. Shifts less than 2 SD rarely are detected. The minimum error that is detected 50% of the time after nine consecutive batches is shown below:

	Parameter Averaged		
Analyte	MCHC	MCH	MCV
Hgb (SD = 0.1 g/dL)	3.6 SD	3.9 SD	—
RBC (SD = 0.05 $\times 10^{12}$/L)	2.4 SD	2.6 SD	—
MCV (SD = 0.60 fL)	4.7 SD	—	4.7 SD

The simulation of populations with outlying indices, *e.g.*, neonates and oncology patients, resulted in both decreased and increased power, depending on the proportion of outliers averaged, the index averaged, and the direction of the shift. (Key words: Quality control; Bull's algorithm; Statistics; Hematology analyzers; Trend analysis; Power functions) Am J Clin Pathol 1985; 83: 337-345

Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, and Department of Pathology and Laboratory Medicine, University Hospital and Clinics, Madison, Wisconsin

IN 1963, Dorsey described the use of the daily average of patient mean corpuscular hemoglobins for the quality control of the Coulter Counter®.[8] The use of average red blood cell indices did not become popular until after 1974, when Bull and associates reported that a novel technic for averaging consecutive patient red blood cell indices might become the primary form of quality control for hematology counters.[2] Bull and associates recommended that this averaging technic (now commonly known as Bull's algorithm) be used to average groups of 20 consecutive patient red blood cell indices. An average outside its 3% limits (0.97–1.03 times the accepted stable mean patient index) would require investigation and, if necessary, corrective instrument maintenance and/or recalibration. Bull's algorithm is now in widespread use, with most of the manufacturers of blood counters, including Coulter Electronics, Ortho Diagnostic Systems, and Toa Medical Electronics, providing online averaging of red blood cell indices. Some hospital hematology laboratories are using averages of patient red blood cell indices as their primary form of quality control.[9]

The advantages of averages of patient indices over commercial whole blood controls include economy and the frequent monitoring of analytic performance.[3] Whole blood controls are expensive and thus are analyzed infrequently relative to the number of patient specimens. Also, because commercial controls have a short shelf life, a large amount of time and effort must be expended in repeatedly determining the mean and acceptable range of new lots of hematology control material. A potential disadvantage of the algorithm is shifting of the average indices outside their 3% limits when certain groups of patients are analyzed, *e.g.*, neonates. Bull and Korpman have suggested that samples be randomized prior to analysis, so no more than one-third of the specimens in one batch of 20 come from cancer chemotherapy wards, neonatal units, or iron deficiency anemia clinics.[3]

Two groups have compared the performance of commercial controls to that of averages of patient indices. Lappin and associates[10] correlated commercial controls with patient indices for 1,900 consecutive patient samples analyzed by the Coulter S®. Unfortunately, the control limits for the commercial control were expressed as multiples of the mean, 0.97 to 1.03 times the long-term control mean, rather than multiples of the long-term standard deviation from the control mean. Because of this, it is difficult to compare the performance of commercial controls with that of average indices. Cavill and associates compared the performance of averages of patient mean corpuscular volume (MCV) with one commercial control analyzed every 20 patient speci-

Received May 3, 1984; received revised manuscript and accepted for publication July 27, 1984.

Address reprint requests to Dr. Cembrowski: Department of Pathology and Laboratory Medicine, William Pepper Laboratory, Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, Pennsylvania 19104.

mens.[5] Two-thousand, one hundred patient specimens were run, with decision limit cusum used to analyze the control data. On all ten occasions that the average MCV was outside its limits, the control samples analyzed by cusum did not indicate any error. Cavill used approximately 2% limits for the patient MCV, which differed from Bull's recommendation of 3% limits for the indices.

A long-term evaluation of Bull's algorithm by Koepke and Protextor[9] reported four years of experience with the system. Forty-two significant instrument malfunctions were correlated to different patterns of deviations of red blood cell indices. The mean corpuscular hemoglobin concentration (MCHC) was affected in 83% of the malfunctions, mean corpuscular hemoglobin (MCH) in 60%, and MCV in 36%. The authors discontinued the use of commercial controls during the study and did not report any comparison with the performance of the commercial controls.

The sensitivity (proportion of runs with analytic error that are detected to have error) and false positive rate (proportion of runs without analytic error that are falsely detected to have error) of Bull's algorithm cannot be assessed readily from actual Coulter runs in which controls and Bull's algorithm are correlated. Westgard and co-workers have used computer simulation to evaluate the sensitivity and false-positive rate of various quality control procedures that use reference samples.[12] The false positive rate and sensitivity may be expressed graphically by plotting the probability of rejection *versus* the size of error, either systematic error or random error. These plots, called power functions, can be used to compare various quality control technics. Westgard has shown that the sensitivity of traditional quality control procedures in detecting moderately sized errors is rather small. For example, the probability of detecting a shift of 2 SD is 12% when 3 SD limits are used as error limits for single controls. The probabilities of detecting the same shift are 20 and 48% if two or four controls are used, respectively, with the same 3 SD limits. Westgard and co-workers have shown that the sensitivity can be increased by increasing the number of control observations (pooling observations from different control materials) and by selecting combinations of control rules.[12]

We recently have used computer simulations of clinical chemistry instruments to study the usefulness of chemistry patient data for quality control.[1,6,7] Computer simulation of multichannel hematology analyzers and associated systematic analytic error should likewise provide information about the utility of Bull's algorithm for quality control. In this article we present the development of a model for simulation of the Coulter Counter, its verification, and the use of Coulter Counter simulations to determine the efficacy of Bull's algorithm.

Materials and Methods

The Coulter Counter directly measures hemoglobin (Hgb), red blood count (RBC), and mean corpuscular volume (MCV). Hematocrit (Hct), mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCH) are calculated from the measured quantities: $Hct = RBC \times MCV$, $MCH = Hgb/RBC$, $MCHC = Hgb/(RBC \times MCV)$. The long-term analytic standard deviations that were used for the simulation of systematic error were obtained from the daily analysis of Coulter whole blood controls at University of Wisconsin Hospital ($SD_{Hgb} = 0.1$ g/dL; $SD_{RBC} = 0.05 \times 10^{12}$/L; $SD_{MCV} = 0.60$ fL). Use of these standard deviations was later substantiated with group data reports that summarized the intralaboratory performance of a large number of Coulter Counters.

Model Development and Validation

The relationships between the directly measured quantities Hgb, RBC, and MCV, were determined from the analysis of consecutive sets of patient Coulter data, as measured during a one-day period (Period 1) by the Coulter Counter. Regression analysis indicated that there was a strong linear relationship between Hgb and RBC and also between MCV and the quotient Hgb/RBC. The relationships showed that the tests could not be simulated independently of each other and required that RBC be calculated from Hgb and MCV then be calculated from the quotient Hgb/RBC.

The validity of the model for the simulation of patient hematology data was tested in two ways. First, the stability of the interrelationships of Hgb, RBC, and MCV was evaluated by analyzing a second set of patient data obtained over another one-day period (Period 2) approximately one month after Period 1. Second, the means and standard deviations of the simulated Hgb, RBC, MCV, Hct, MCH, and MCHC were compared with the means and standard deviations of Period 1.

Simulation of Populations with Outlying Indices

To test the response of Bull's algorithm to populations with outlying indices, the Coulter Hgb, RBC, and MCV values of two different populations, 150 neonatal intensive care patients (0–4 weeks old) from another hospital and 100 consecutive oncology outpatients seen at University of Wisconsin Hospital were analyzed with multiple regression to obtain coefficients for calculating RBC from Hgb and MCV from Hgb/RBC. The validity of the regression equations for the simulation of the outlying populations was evaluated by comparing the means and standard deviations of the hematology parameters of the simulated populations with those of the original populations.

Table 1. Comparison of Monthly Intralaboratory Standard Deviations of Hgb, RBC, and MCV for Four Different Coulter Blood Counters Compiled from Fisher Diagnostics Group Data Reports 200 (11/5/83–12/2/83) and 012 (3/3/84–3/30/84)

	Coulter Model			
	S Senior	S+	S + II	S + 4
Hgb				
Number of instruments	133	50	21	13
Average Hgb (g/dL)	12.4	12.4	12.4	12.4
Average SD_{Hgb} (g/dL)	0.12	0.12	0.14	0.14
Range of SD_{Hgb} (g/dL)	0.04–0.27	0.07–0.27	0.07–0.27	0.07–0.33
RBC				
Number of instruments	133	48	23	13
Average RBC ($\times 10^{12}$/L)	4.06	4.12	4.10	4.15
Average SD_{RBC} ($\times 10^{12}$/L)	0.04	0.04	0.04	0.04*
Range of SD_{RBC} ($\times 10^{12}$/L)	0.02–0.12	0.02–0.08	0.03–0.07	0.03–0.08*
MCV				
Number of instruments	130	47	21	13
Average MCV (fL)	84.9	88.5	86.6	86.4
Average SD_{MCV} (fL)	1.00	0.68	0.55	0.60
Range of SD_{MCV} (fL)	0.44–3.11	0.23–1.63	0.51–0.97	0.28–0.80

* After exclusion of three outlying values of SD_{RBC}.

Description of Simulation Program

The program was written in FORTRAN and required 130,000 words of memory for execution on a Univac 1100® computer. Patient Hgb data were simulated with a random Gaussian number generator and the Period 1 Hgb mean and standard deviation. The regression equation for computing RBC from Hgb then was used to calculate RBC. The random Gaussian number generator was used to add random error to the calculated RBC values with the means set to the individual RBC values and the standard deviation as the standard error of the estimate $S_{RBC/Hgb}$. MCV was calculated from the regression equation using the quotient Hgb/RBC.

Analytic shifts were simulated separately in Hgb, RBC, and MCV. Thus, for Hgb, multiples of 0, 0.5, 1.0, . . . , 5.0 of the long-term standard deviation of Hgb were added to the calculated Hgb value. Similarly, analytic shifts in either RBC or MCV were simulated by adding multiples of the long-term standard deviation of RBC or MCV to the calculated RBC or MCV value, respectively. Once the individual Hgb, RBC, and MCV were simulated, the Hct, MCH, and MCHC were computed. At each error level, 500 groups of 20 patients were simulated, with their indices averaged by Bull's algorithm.[2] The proportion of average indices that exceeded the limits of 0.97 to 1.03 times the mean index then was calculated and tabulated. Power function curves were constructed by plotting the proportion or probability of rejection *versus* the size of systematic error.

Because Bull's algorithm incorporates information from preceding batches, it was important to study the response of the algorithm to increasing numbers of patient batches. Up to nine successive batches of 20 patients were simulated to determine the long-term response to error. Finally, simulations were performed that measured the response of Bull's algorithm to different proportions of patients with outlying indices. The proportion of either the simulated neonatal or oncology populations was varied from 0.10 to 0.50 with the response to both positive and negative systematic errors measured.

Results

Model Validation

The long-term intralaboratory standard deviations of Hgb, RBC, and MCV for different Coulter Counters

Table 2. Coefficients of Regression (±SD) for RBC = a + b Hgb

	Period 1 (N = 144)	Period 2 (N = 138)	Neonatal Population (N = 150)	Oncology Outpatients (N = 100)
a	0.134 ± 0.142	0.118 ± 0.119	0.789 ± 0.184	0.456 ± 0.267
b	0.315 ± 0.011	0.320 ± 0.010	0.241 ± 0.011	0.284 ± 0.021
R^2	0.840	0.880	0.763	0.662
$S^*_{RBC/Hgb}$	0.318	0.311	0.360	0.400

* Standard error of the estimate.

Table 3. Coefficients of Regression (±SD) for MCV = c + d(Hgb/RBC)

	Period 1	Period 2	Neonatal Population	Oncology Outpatients
c	14.6 ± 1.6	14.9 ± 2.2	8.00 ± 2.54	7.68 ± 1.50
d	23.7 ± 0.5	23.8 ± 0.7	27.5 ± 0.7	27.0 ± 0.5
R^2	0.933	0.891	0.906	0.971
$S^*_{MCV/(Hgb/RBC)}$	1.64	1.96	2.51	1.52

* Standard error of the estimate.

shown in Table 1 are very close to the long-term standard deviations used in the simulations. Tables 2 and 3 compare the Period 1 and 2 coefficients of regression for the relationships between RBC and Hgb, and MCV and the ratio Hgb/RBC. There are no significant differences between the Period 1 and 2 coefficients. R^2 is uniformly high for the relationships, indicating that RBC and MCV can be computed from Hgb and Hgb/RBC, respectively. Tables 2 and 3 also show the coefficients of regression for the computation of RBC and MCV for the neonatal and oncology outpatient populations.

Table 4 shows the means and standard deviations of the Coulter red blood cell parameters for the three different populations—the usual hospital population and the neonatal and the oncology outpatient populations. The neonatal intensive care indices are markedly different from the usual hospital population indices, with MCV increased by 18%, MCH increased by 12.6%, and MCHC decreased by 4.6%. Although the direction of the deviations is the same as in term neonates,[11] their red blood cell parameters cannot be compared rigorously with literature values because many of the neonates had been transfused. The deviations of the oncology population are in the same direction as the neonates but are smaller in magnitude with MCV increased by 6.0%, MCH increased by 2.5%, and MCHC decreased by 3.4%. In the oncology population, the increased MCV is consistent with chemotherapy and the decreased MCHC with the anemia of chronic disease. Also shown in Table 4 are the means and standard deviations of the three simulated populations. With the exception of the difference between the standard deviation of the simulated and actual neonatal MCHC (0.50 *vs.* 0.82, respectively), there is excellent agreement between the simulated and actual indices for all three populations.

Response to Analytical Error

Figure 1 shows plots of families of power functions, in which the probability of detecting an average red blood cell index outside its 3% limits is plotted on the y-axis *versus* the size of systematic error (SE) on the x-axis. The size of the systematic error is expressed in multiples of the long-term standard deviation. The labeled lines on individual plots correspond to the number of consecutive batches of 20 patients averaged after introduction of the systematic error. Figure 1*A*, for example, is a plot of the probability of detecting an average MCH outside its 3% limits when systematic error, expressed in multiples of SD (0.1 g/dL), is introduced into the Hgb measurement. The probability of Bull's algorithm detecting a moderately large shift (3 SD or 0.3 g/dL) is low after one batch of 20 patients,

Table 4. Comparison of Actual* and Simulated† Population

	Hospital Population		Neonate Population		Oncology Population	
	Actual	Simulated	Actual	Simulated	Actual	Simulated
Hgb (g/dL)	12.10	12.12 ± 0.09	16.52	16.56 ± 0.10	12.86	12.90 ± 0.06
SD_{Hgb} (g/dL)	2.29	2.27 ± 0.05	2.68	2.67 ± 0.05	1.96	1.98 ± 0.03
RBC ($\times 10^{12}$/L)	3.95	3.96 ± 0.04	4.77	4.78 ± 0.02	4.11	4.13 ± 0.03
SD_{RBC} ($\times 10^{12}$/L)	0.79	0.78 ± 0.02	0.74	0.74 ± 0.02	0.68	0.68 ± 0.02
MCV (fL)	87.66	87.62 ± 0.27	103.49	103.51 ± 0.27	92.91	92.63 ± 0.26
SD_{MCV} (fL)	6.29	6.30 ± 0.20	8.14	7.87 ± 0.12	8.84	8.72 ± 0.20
Hct (%)	34.48	34.52 ± 0.27	49.35	49.36 ± 0.28	37.88	37.99 ± 0.17
SD_{Hct} (%)	6.66	6.45 ± 0.15	8.33	7.89 ± 0.15	5.48	5.80 ± 0.11
MCH (pg)	30.80	30.81 ± 0.11	34.67	34.72 ± 0.08	31.58	31.52 ± 0.09
SD_{MCH} (pg)	2.56	2.69 ± 0.08	2.81	2.88 ± 0.05	3.22	3.27 ± 0.06
MCHC (g/dL)	35.14	35.13 ± 0.01	33.51	33.54 ± 0.01	33.95	33.95 ± 0.02
SD_{MCHC} (g/dL)	0.83	0.79 ± 0.02	0.82	0.50 ± 0.01	0.63	0.65 ± 0.03

* Period 1.

† Mean ± 1 SD of nine groups of simulations with 1,000 patients simulated per group.

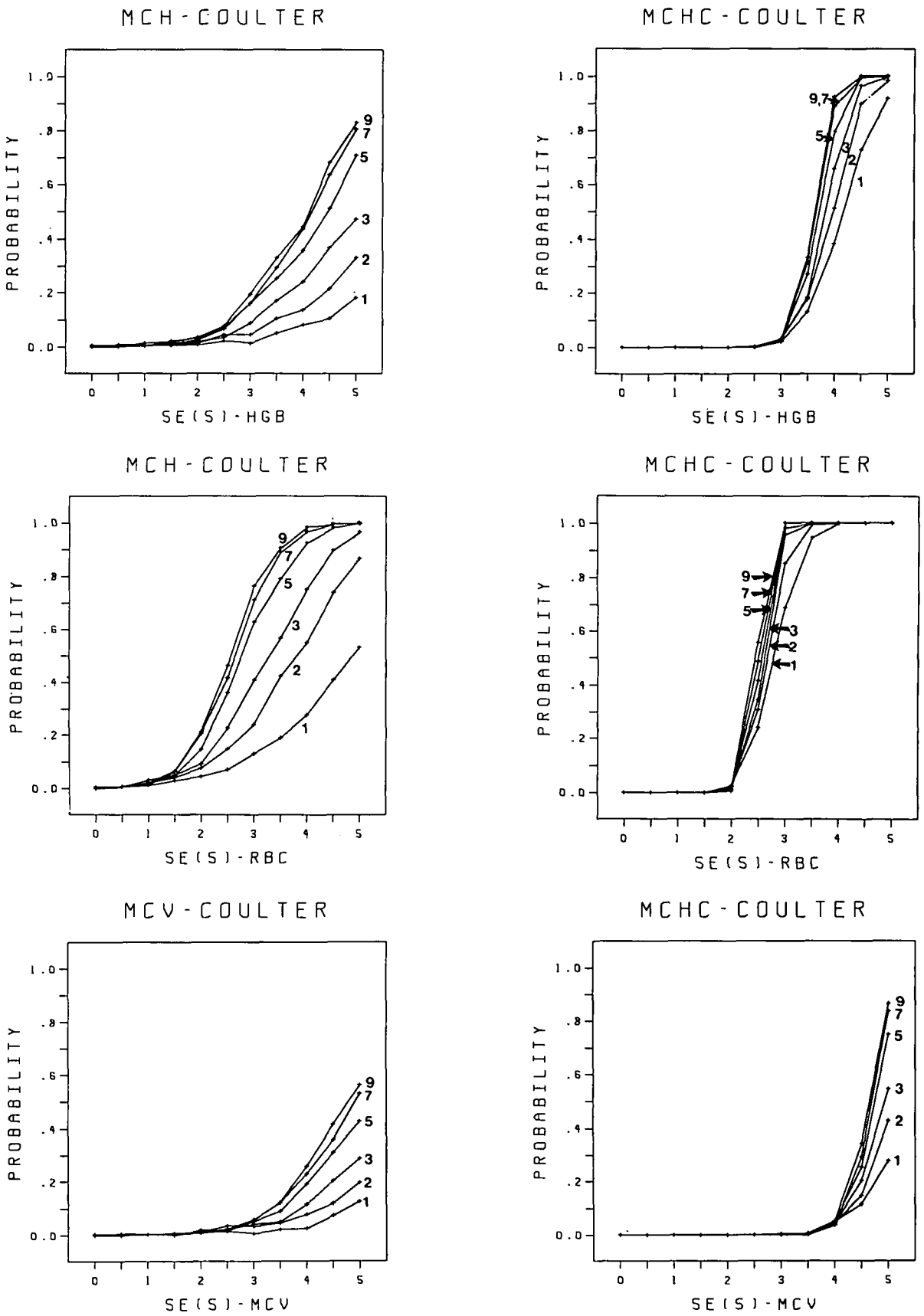

FIG. 1. Power functions for Bull's algorithm for the detection of error in hemoglobin (*A, upper, left, B, upper, right*) red blood cell count (*C, center, left, D, center, right*), and mean corpuscular volume (*E, lower, left, F, lower, right*). The probability of rejection is plotted against the size of the systematic error (SE) expressed in multiples of the long-term standard deviation (S). The numbers 1, 3, 5, 7, and 9 correspond to the number of batches of 20-patient specimens analyzed.

averaged, approximately 2%. After five batches of 100 patients this probability increases to 17%. The probability of average MCHC detecting systematic error in Hgb is shown in Figure 1*B*. Errors of less than 3 SD have a very low probability of being detecting (about 3%) even with large batch numbers. Shifts of greater than 4 SD, however, have a high probability of detection.

The probability of detecting systematic errors in RBC is shown in Figures 1*C* and 1*D*. Even after one batch, the probability of detecting a 3 SD shift in RBC is very high when MCHC is averaged, approximately 70%. After nine batches the probability of detecting a 3 SD shift in RBC using averaged MCH or MCHC is at least 80%. The performance of the average indices after nine batches is summarized in Table 4, the minimum error detected at least 50% of the time. Figures 1*E* and 1*F* show that only large shifts (greater than 4 SD) in MCV can be detected with moderate probability when either MCV or MCHC is averaged.

Response to Populations with Outlying Indices

Figure 2 shows the effect of varying the proportions of neonatal specimens analyzed in a batch of 20 specimens. In each of the six plots, Figures 2*A–F*, lines labeled with "B = 1" or "B = 9" correspond to the power functions of Bull's average for 1 and 9 consecutive batches, respectively. The proportion of neonatal red blood cell indices averaged with the usual hospital population was varied from 0.10 to 0.50. Because the outlying populations affected the power functions nonsymmetrically, the effects of the outlying populations are shown for positive and negative systematic errors.

To illustrate the use of these power functions, Figure 2*A* shows the probability of detecting shifts in Hgb when MCH is averaged. If a small proportion of the population averaged is composed of neonates (0.1), then the probability of detecting a large positive shift increases when compared with the power function for Bull's average of the first batch. As the proportion of neonatal indices averaged increases to 0.5, the probability of detecting any positive shift becomes very large. In fact, if one-half of the population is neonates and no systematic error exists, approximately 30–40% of the time Bull's average of MCH will be outside its 3% limits. This high false positive rate is due to the averaging of the elevated neonatal MCH values. The probability of detecting large negative shifts in Hgb, conversely, is decreased tremendously, with neonatal proportions of 0.10–0.20 resulting in almost a zero probability of detecting a large negative shift. The opposite phenomenon occurs with errors in Hgb when MCHC is averaged. Figure 2*B* shows that the probability of detecting a negative shift is increased when there is a significant proportion of neonates averaged; however, the probability of detecting a positive shift decreases.

With the inclusion of a significant proportion of neonates, positive shifts in MCV result in an increased probability of detection if either MCV or MCHC is averaged (Figs. 2*E* and 2*F*). If the proportion of neonates exceeds 0.30, there is a high probability of the average MCV being outside its 3% limits without an error being present (high false postive rate). Negative shifts, in the presence of even a small proportion of neonates (0.10–0.20), result in a very low probability of error detection.

Figure 2*C* shows an increased probability of detection of negative shifts in RBC if MCH is averaged. There is again a high false positive rate for proportions of neonates exceeding 0.50. The probability of detecting a positive shift in RBC is decreased if MCH is averaged. The averaging of MCHC (Fig. 2*D*) has the opposite effect—a lower probability of detecting a negative shift and a higher probability of detecting a positive shift. The effects of averaging variable proportions of an oncology outpatient population are shown in the power functions of Figure 3. The effects on the power functions are similar to those shown in Figure 2 but are decreased in amplitude.

Discussion

Cavill and Jacobs[4] recommend that the maximum acceptable difference between two daily measurements for Hgb and RBC should be approximately 6%, or twice the maximum physiologic variation. This suggests that analytic errors larger than 3% should be detected by the control procedure to provide medically useful information. With respect to MCV, Cavill and Jacobs suggest that the "stability of the 'absolute' red blood cell indices is a hematologic axiom," suggesting that small errors may limit the interpretation of the measurements.

Inspection of Figure 1 indicates that when 3% control limits are employed, small to moderate shifts cannot be detected by averaging patient data. In the absence of populations with outlying indices, Hgb shifts less than 3.5 SD (0.35 g/dL), RBC shifts less than 2 SD (0.10 $\times 10^{12}$/L), and MCV shifts less than 3.5 SD (2.1 fL) will not be detected. Table 5 indicates that even after nine batches, the systematic error must be very large (in multiples of the long-term standard deviation) to be detected even 50% of the time when it occurs. The sizes of the smallest errors that are detected 50% of the time are 0.36 g/dL for Hgb, 0.12 $\times 10^{12}$/L for RBC, and 2.8 fL for MCV.

When patient populations with outlying indices are present, *e.g.*, neonates and oncology patients, the size of the errors that may escape detection can approach 5–6 SD (Figs. 2 and 3). The averages of patient indices are

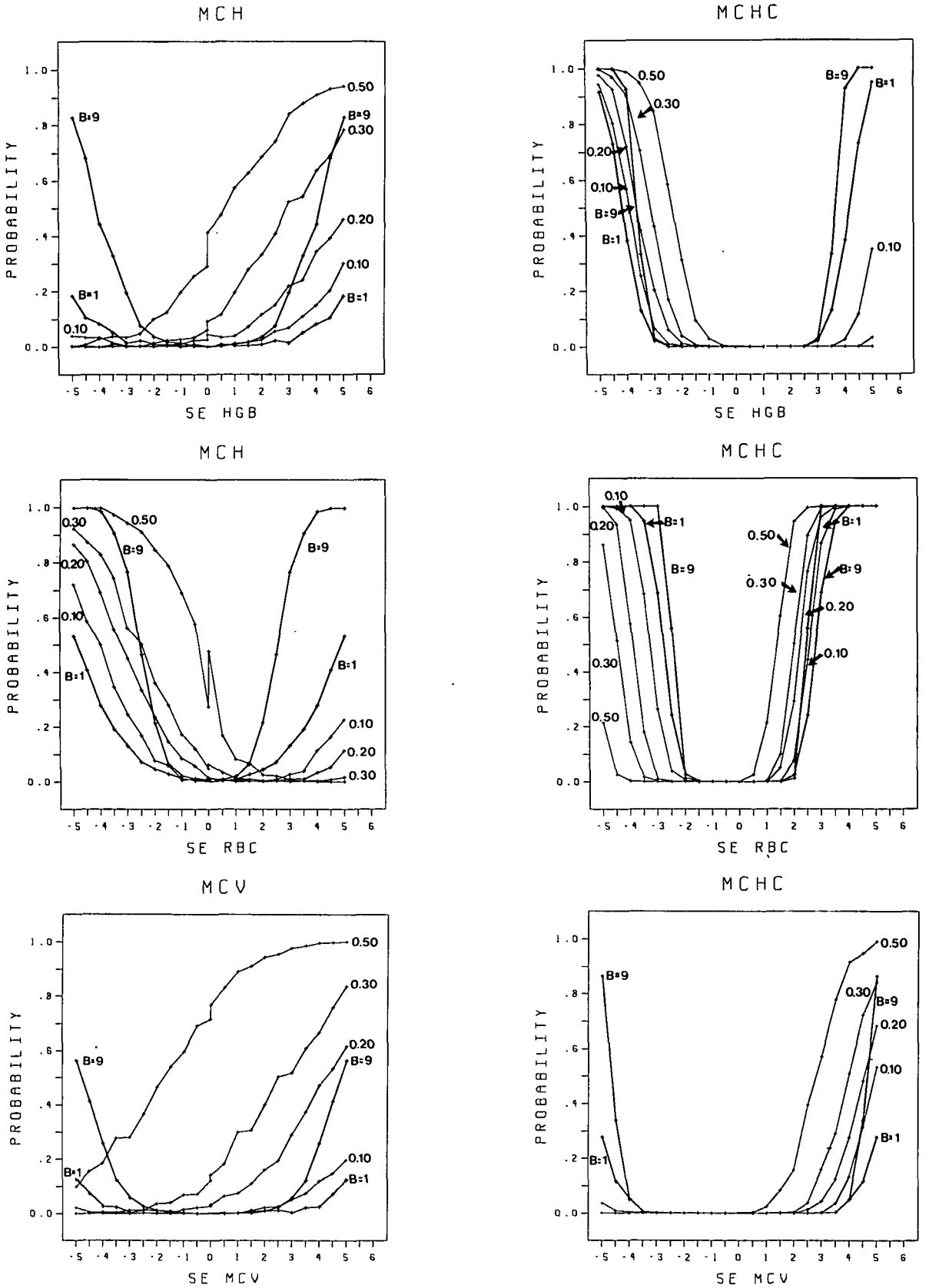

FIG. 2. Effect of averaging different proportions of neonatal patients on the probability of rejection for errors in hemoglobin (*A, upper, left, B, upper, right*), red blood cell count (*C, center, left, D, center, right*), and mean corpuscular volume (*E, lower, left, F, lower, right*). The probability of rejection is plotted against the size of the systematic error (SE), expressed in multiples of the long-term standard deviation. The systematic error can either be positive or negative. The thick lines marked B-1 and B-9 correspond to the probability of rejection of a run without neonatal patients after one and nine batches, respectively. The proportion of neonates is varied from 0.10 to 0.50.

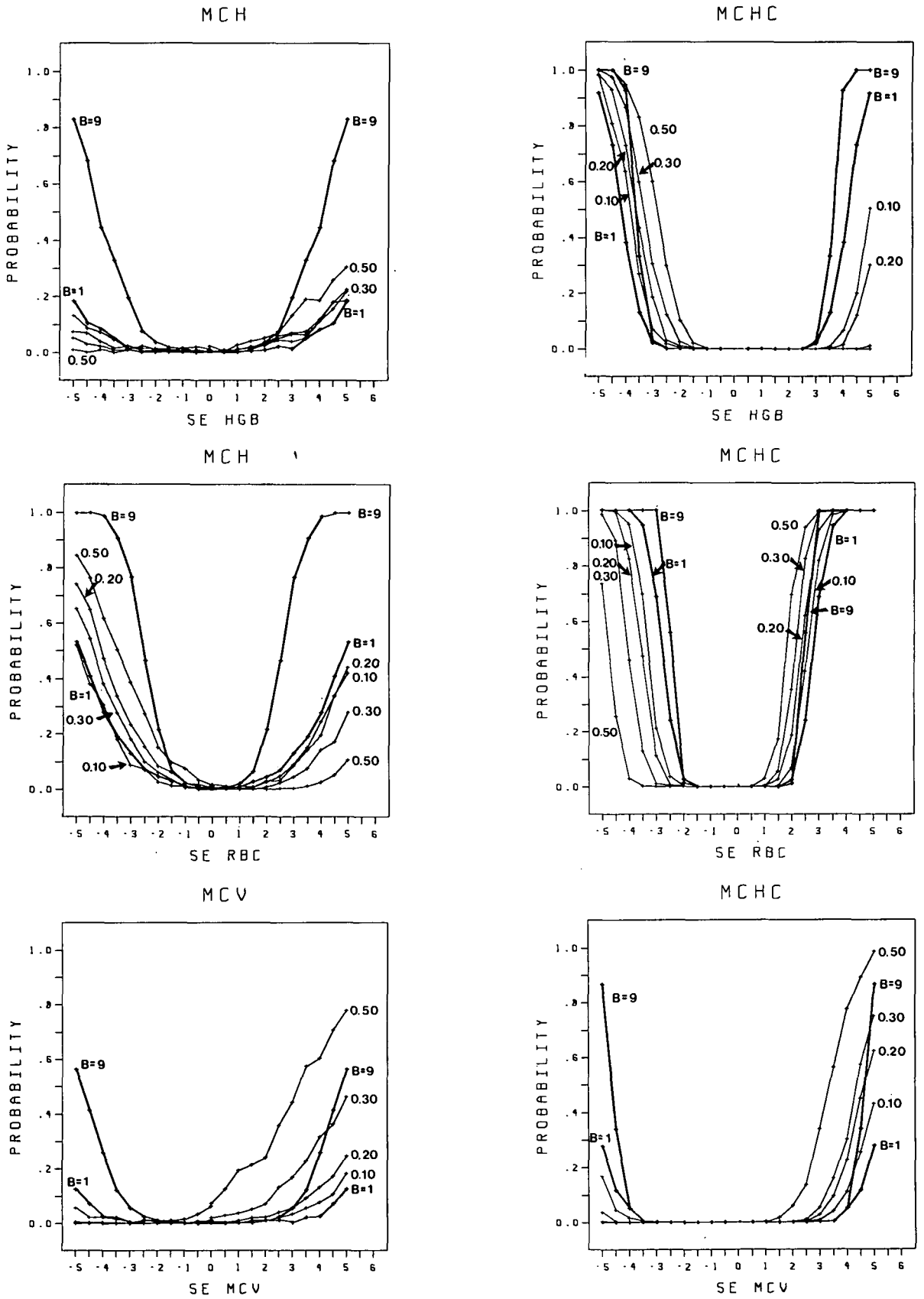

FIG. 3. Effect of averaging different proportions of oncology outpatients on the probability of rejection for errors in hemoglobin (*A, upper, left, B, upper, right*), red blood cell count (*C, center, left, D, center, right*), and mean corpuscular volume (*E, lower, left, F, lower, right*). The probability of rejection is plotted against the size of the systematic error (SE), expressed in multiples of the long-term standard deviation. The systematic error can either be positive or negative. The thick lines marked B-1 and B-9 correspond to the probability of rejection of a run without oncology outpatients after one and nine batches, respectively. The proportion of oncology outpatients is varied from 0.10 to 0.50.

Table 5. Minimum Error Detected 50% of the Time (nine batches averaged)

Analyte	Averaged Parameter		
	MCHC	MCH	MCV
Hgb (SD = 0.1 g/dL)	3.6 SD	3.9 SD	—
RBC (SD = 0.05 × 10^{12}/L)	2.4 SD	2.6 SD	—
MCV (SD = 0.60 fL)	4.7 SD	—	4.7 SD

thus most useful when patients with homogeneous indices are analyzed. Because patients with outlying indices limit the usefulness of instrument-based calculations of mean patient indices, it would be more effective if the laboratory computer checked the demographic data and eliminated any neonates and oncology patients from calculations of average indices.

For comparison purposes, the error detection rates for control procedures based on stable materials also can be estimated from power function graphs.[12] Error detection will depend on the number of control observations and the control rules (charting technic) employed. Use of 3 SD control limits and one to two control observations per run will provide 12–20% detection of 2 SD shifts. Use of a multirule control procedure[13] with four observations provides 60% detection of 2 SD shifts and 100% detection of 3 SD shifts. When the averages of patient red blood cell indices are employed, 2 SD shifts seldom will be detected (probability near zero), regardless of the number of batches that are averaged. When nine batches are averaged, detection of 3 SD shifts increases to 100% for RBC, 20% for Hgb, and 5% of MCV.

These simulation studies suggest that quality control by averaging patient red blood cell indices does not offer as good error detection as that available from conventional charting technics using stable control materials. In addition, the use of stable control materials permits the monitoring of random errors, whereas the averages of patient data cannot provide any detection of random errors. Control materials can be analyzed at any time, without any requirement to accumulate a batch of specimen results, and thus are especially useful when starting up an analyzer and after maintenance and recalibration. The disadvantages of control materials are, of course, the relatively short period of stability (one to three months), their cost, and the fact that they are occasionally defective.

We recommend caution in discontinuing the use of control materials. We think that a balanced use of control materials and patient data is advisable at present and that continuing work is needed to further optimize control technics for hematology applications.

Acknowledgments. The authors thank Judy Gosse for her assistance with the computer simulations and plotting and Caroline Stieghorst and Audrey Faust of University Hospital, Madison, for their support. They also acknowledge Bruce Rotkowitz and Fisher Diagnostics for allowing them to publish portions of the Fisher Hematology Group Data Reports, and Marge Wilson of St. Mary's Hospital, Madison, for her continuing advice and provision of the neonatal hematology data.

References

1. Bockelman HW, Cembrowski GS, Kurtycz DFI, Garber CC, Westgard JO, Weisberg HF: Quality control of electrolyte analyzers: evaluation of anion gap average. Am J Clin Pathol 1984; 81:219–223
2. Bull BS, Elashoff RM, Heilbron DC, et al: A study of various estimators for the deviation of quality control procedures from patient erythrocyte indices. Am J Clin Pathol 1974; 61:473–481
3. Bull BS, Korpman RA: Intralaboratory quality control using patients' data, Methods in hematology, vol 4, Quality control. Edited by I Cavill. New York, Churchill Livingstone, 1982, pp 121–150
4. Cavill I, Jacobs A: The interpretation and significance of laboratory results, Methods in Hematology, vol 4, Quality control. Edited by I Cavill, New York, Churchill Livingstone, 1982, pp 173–181
5. Cavill I, Ricketts C, Fisher J, Walpole B: An evaluation of two methods of laboratory quality control. Am J Clin Pathol 1979; 72:624–627
6. Cembrowski GS, Chandler EP, Westgard JO: Assessment of "average of normals" quality control: procedures and guidelines for implementation. Am J Clin Pathol 1984; 81:492–499
7. Cembrowski GS, Westgard JO, Kurtycz DFI: Use of anion gap for the quality control of electrolyte analyzers. Am J Clin Pathol 1983; 79:688–696
8. Dorsey DB: Quality control in hematology. Am J Clin Pathol 1963; 40:457–464
9. Koepke JA, Protextor TJ: Quality assurance for multichannel hematology instruments, four years experience with patient mean erythrocytic indices. Am J Clin Pathol 1981; 75:28–33
10. Lappin TRJ, Farrington CL, Nelson MG, et al: Intralaboratory quality control in hematology—comparison of two systems. Am J Clin Pathol 1979; 72:426–431
11. Matoth Y, Zaizor R, Varsano I: Postnatal changes in some red cell parameters. Acta Paediatr Scand 1971; 60:317–323
12. Westgard JO, Groth T: Power functions for statistical control rules. Clin Chem 1979; 25:863–869
13. Westgard JO, Barry PL, Hunt MR, Groth T: A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem 1981; 27:493–501