

Validation of Criteria for Smear Review Following Automated Blood Cell Analysis in Ain Shams University Laboratory

Azza S Eldanasoury¹, Noha H Boshnak², Raghda E Abd El Monem³

¹MD, Professor of Clinical Pathology, Faculty of Medicine, Ain Shams University, Cairo, Egypt

²MD, Lecturer of Clinical Pathology, Faculty of Medicine, Ain Shams University, Cairo, Egypt

³M.Sc, Assistant Lecturer of Clinical Pathology, Faculty of Medicine, Ain Shams University, Cairo, Egypt

Abstract: *The rules suggested by the International Consensus Group for Hematology Reviews have been considered the international standard to indicate situations requiring a blood smear review; however no guidelines can be used universally without considering the actual demands and capabilities of each laboratory. This study aimed to validate and compare the performance of the criteria for smear review suggested by the Consensus Group with our laboratory criteria. 800 blood samples were selected randomly from the daily workload. Automated complete blood counts (CBC) and white blood cell (WBCs) differential counts were performed using Beckman Coulter LH750. Blood films were done for all samples and reviewed for positive smear findings as identified by the International Consensus. The 2 sets of criteria (the Consensus Group and our laboratory criteria) were applied on the samples. Compared with our laboratory criteria, the false negative rate of Consensus Group criteria was higher (9.25% versus 1.62%, $p < 0.05$), the review rate was lower (54.25% versus 71%, $p < 0.05$), the sensitivity was lower (82.13% versus 97.09%, $p > 0.05$), the specificity was higher (78.32% versus 62.04%, $p > 0.05$). In conclusion we found that peripheral smear review rate was significantly reduced by applying the consensus group criteria.*

Keywords: Beckman Coulter LH750, CBC, criteria, Smear review.

1. Introduction

Modern automated blood cell counters are the cornerstone of the hematology laboratory, providing quick, cost-effective, and accurate analysis of the blood cells. Peripheral blood smear review still play an important role to identify some morphologic abnormalities and to complete the definitive interpretation in some cases, however it is time consuming, laborious, and demanding the skill of experienced morphologist [1].

The number of samples requiring manual slide review greatly influences the laboratory costs, productivity and turnaround time (TAT) [2]. Therefore the laboratories have been left in a position of challenge to provide accurate test results as soon as possible without compromising the quality and affecting the patient care by reporting false results, leading to suboptimal diagnosis and treatment [3]. As a result each laboratory should have a set of criteria of action after automated analysis of blood cells [4]. In 2005, the International Society for Laboratory Hematology published a set of 41 rules applicable as review criteria for automated complete blood counts (CBC) and suggested that any laboratory adopting the criteria for action, following automated blood cell analysis, validate their operation before implementing them for use on patient samples [5].

The aim of the present work is to validate and compare the performance of the criteria for smear review suggested by the International Consensus Group for Hematology Review in 2005 (Consensus Group criteria) with the currently used criteria for smear review in the main laboratory of Ain Shams University hospital (our laboratory criteria).

2. Materials and Methods

The study was performed in the main laboratory of the Clinical Pathology Department, Ain Shams University, a tertiary care university hospital serving a large inpatient and outpatient population. The study samples were routine inpatient samples from all departments of Internal Medicine Hospital except for Hematology Oncology Department which has a dedicated laboratory.

Eight hundreds blood samples were selected over a 3 month period. 20 samples were selected at random from daily workload. As per the laboratory policy, 2mL venous blood samples were drawn using K3 EDTA (tripotassium ethylene diamine tetraacetate) vacutainers under complete aseptic conditions. Samples were analyzed within 3 hours of collection.

2.1 Automated CBC and white blood cell differential counts

Automated CBC and WBC differential counts were performed using Beckman Coulter LH750 (Beckman Coulter, Miami, Florida) that uses electric impedance technology to enumerate WBC, red blood cells (RBC), and platelet (PLT) counts, and to determine RBC and PLT volume. Hemoglobin (Hb) is determined by cyanmethemoglobin colorimetry.

Procedures for quality assurance and quality control were followed to ensure good performance using CBC-5D Hematology Controls. The hematology analyzer was regularly calibrated by standardized calibrators

(COULTER® S-CAL® Calibrator). During the course of the study there were no repair in the equipment, change in reagents, calibration standards or controls.

2.2 Manual differential and blood smear review

Blood films were smeared manually from all samples, and stained with Leishman stain, regardless of whether these blood smears would have been required or not according to our standard operating procedures. A 200 cell manual differential with a blood smear review was performed on all films in the study blindly (without knowing the results generated by the analyzer). The microscopic smear review focused on the morphology of blood cells. The positive smear results were reviewed by senior staff. If the film contained a positive finding, the sample was labeled positive. The criteria for a positive smear were applied as recommended by the International Consensus Group, which included abnormality of RBC morphology abnormality at 2+ or greater, malaria, giant Platelets at moderate or greater, Platelets clumps at greater than rare/occasional, Döhle bodies/toxic granulation/vacuoles at moderate or greater, blasts at 1 or greater, metamyelocytes at greater than 2, myelocytes/promyelocytes at 1 or greater, atypical lymphocytes at greater than 5, normoblast at 1 or greater, or plasma cell at 1 or greater [5].

2.3 Validation of Consensus Group rules and our laboratory criteria

The results of all 800 samples tested were reviewed according to the consensus rules as shown in Table 1. The main adaptation was exclusion of neonates, children and delta check rules. Actually, our laboratory does not receive samples from children or neonates for CBC and it was not possible to implement delta check rules because of limitations of the electronic hospital records.

The instrument results for the same 800 samples were also reviewed according to the criteria for smear review already in place in our laboratory, which are mainly based on the expert opinion of department senior professors (Table 1). According to each set of criteria (Consensus Group and our laboratory criteria) we decided on whether one or more rules were triggered and thus necessitate a blood smear review. A rule in the criteria would be triggered when the result was beyond the specified range and/or a specified flag appeared.

Sample classification

After defining which samples would have required smear review according to the criteria; the results were compared to the actual findings of the peripheral blood smear to classify samples into true positives, false positives, true negatives and false negatives. If a rule was triggered and the smear result was positive, the sample was graded as a "true positive". If a rule was triggered and the smear did not have any positive findings, the sample was graded as a "false positive". If a rule was not triggered and the smear result was negative, the sample was graded as a "true negative". If a rule was not triggered but the smear contained a positive finding, the sample was graded as a "false negative". This was done for both sets of criteria.

Additional positive smear cases

An additional 51 inpatient samples with positive smear findings were selected to enrich the tested samples and include findings that were not represented in the 800 samples. Both sets of criteria were applied to the samples.

2.4 Statistical Analysis

Statistical analysis was performed with Excel software (Microsoft, Remond, Washington). False positive, false negative, true positive and true negative rates were calculated for both Consensus Group criteria and our laboratory criteria. Efficiency, sensitivity, specificity, moreover positive predictive value (PPV) of flags, negative predictive value (NPV) of flags and the review rate for both Consensus Group and our laboratory criteria for smear review were calculated Chi-square tests with Yat's correction were used to compare the different performance specifications between both sets of criteria. A *p* value of <0.05 was considered statistically significant.

Table 1: Criteria for blood smear review applied in the study

Parameters	Consensus Group Criteria	Our Laboratory Criteria
CBC		
Hb (g/dL)	<7 or >2 above upper reference range	M: <10 or >19 F: <9 or >18
MCV (fL)	<75 or >105	<70 or >100
RDW (%)	>22	>20
WBC ($\times 10^9/L$)	<4 or >30	<2.5 or >20
PLT ($\times 10^9/L$)	<100 or >1000	<150, >610
Differential		
No differential	No differential	No differential
No. of neutrophils ($\times 10^9/L$)	<1 or >20	<1.5 or >15
No. of lymphocytes ($\times 10^9/L$)	>5	>3.5
No. of monocytes ($\times 10^9/L$)	>1.5	>1 with normal neutrophil count, >1.5 with absolute neutrophilia
No. of eosinophils ($\times 10^9/L$)	>2	>1
No. of basophils ($\times 10^9/L$)	>0.5	≥ 0.2
Suspect flags		
Nucleated red blood cell	Flag	Flag
Blast	Flag	Flag
Atypical lymphocyte	Flag	Flag
Dimorphic RBC	Flag	Flag
Immature granulocyte	Flag	Flag
PLT clump	Flag	Flag
Platelet (except PLT clump)	Flag	Flag
Verify differential	-	Flag
Cellular interference	-	Flag

3. Results

3.1 Analysis of smear review findings

The 800 blood films examined were related to 416 (52%) females (F) and 384 (48%) males (M) with a M/F ratio of 0.9. Their age ranged from 18 to 90 years with a median of 54 years. Following the Consensus Group definition of positive smears, 434 samples (54.25%) had positive smear results. Among the positive samples, 370 (52.63%) had RBC abnormalities, 98 (13.94%) had WBC abnormalities, 154 (21.90%) had PLT

abnormalities, 40 (5.68%) had both RBC and PLT abnormalities, 20 (2.84%) had both RBC and WBC abnormalities, 15 (2.13%) had both WBC and PLT abnormalities, and 6 (0.85%) had RBC, WBC, and PLT abnormalities.

3.2 Comparison of performance between the Consensus Group and our laboratory criteria for smear review

On applying the Consensus Group criteria for smear review and counting any sample requiring review as one sample

regardless of the number of criteria triggered, 434 samples of the 800 samples required review with review rate of 54.25%. While on applying our laboratory criteria the corresponding figures were 568 samples and 71.0%. The samples that required review were further analyzed according to the criteria triggered, i.e. a sample triggering more than one criterion was counted with each criterion (Table 2).

Table 2: Number of cases requiring smear review by each of the triggering criteria of the Consensus Group and our laboratory

	Consensus criteria		Our laboratory criteria	
	Rule triggered	Occurrence %	Rule triggered	Occurrence
WBC ($\times 10^9/L$)	<4	73	< 2.5	27
	>30	14	>20	41
PLT ($\times 10^9/L$)	<100	169	<150	266
	>1000	0	>610	12
Hb (g/dL)	<7	60	<10 (M), <9 (F)	235
MCV (fL)	<75	92	<70	49
	>105	8	>100	28
RDW (%)	>22	66	>20	106
Neutrophils ($\times 10^9/L$)	<1	18	<1.5	31
	>20	19	>15	53
Lymphocytes ($\times 10^9/L$)	>5	18	>3.5	53
Monocytes ($\times 10^9/L$)	>1.5	49	>1	75
Eosinophils ($\times 10^9/L$)	>2	2	>1	6
Basophils ($\times 10^9/L$)	>0.5	0	≥ 0.2	2
Suspect flags	Both consensus and our laboratory criteria			
	n		% of reviewed cases	
Dimorphic picture	47		11.1	
NRBC	61		14.4	
PLT clumps	60		14.2	
Giant PLT	76		17.9	
Immature granulocytes	62		14.6	
Variant lymphocytes	8		1.9	
Blasts	15		3.5	
No differential	30		7.1	
Cellular interference*	53		12.5	
Verify differential [#]	12		2.8	

* WBC histogram pattern consistent with interference at the 35 fL region. When the separation between the WBC populations is poorly defined on the histogram, WBC correction will be performed.

[#] When an unexpected data pattern is encountered. The message also is generated when WBC $> 1.5 \times 10^3$ cells/ μ L and MO% > 20 .

As shown in Table 2, the three most common flags encountered were giant platelet flags (76 cases; 17.9 %), immature granulocyte flags (62 cases; 14.6%) and normoblast flags (61 cases; 14.4%). The performance of the Consensus Group criteria for smear review was compared with our laboratory criteria (Table 3), it is worth noting that the review rate with the consensus group criteria was significantly different from that with our laboratory criteria with a p value < 0.001 , the false negative rate was higher with the Consensus Group criteria than our laboratory criteria ($p < 0.001$), however, the efficiency was not different between the 2 sets of criteria ($p = 0.5$).

Table 3: Comparison of performance between the Consensus Group and our laboratory criteria

Parameter	Consensus criteria	Our laboratory criteria	P Value
False negative, n(%)	77(9.25%)	13(1.62%)	< 0.001
False positive, n(%)	80(10%)	134(16.75%)	< 0.001
True negative, n(%)	289(36.12%)	219(27.37%)	< 0.001
True positive, n(%)	354(44.25%)	434(54.25%)	0.00
Efficiency	80.37%	81.62%	0.566
Review rate	54.25%	71%	< 0.001
Positive predictive value	81.56%	76.40%	0.526
Negative predictive value	78.96%	94.39%	0.163
Sensitivity	82.13%	97.09%	0.098
Specificity	78.32%	62.04%	0.051

Table 4: Analysis of samples triggered by only one of the two sets of criteria

Parameter	Cutoffs	Cases Between 2 Cutoff Values, n	False negative n (%)	False positive n (%)
Low MCV (fL)	<75 * <70 #	43	8 (18.6)	4 (9.3)
High MCV (fL)	>105 * >100 #	20	0 (0.0)	2 (10)
RDW (%)	>22 * >20 #	40	0 (0.0)	5 (12.5)
Low hemoglobin, (g/dL)	<7 * <10 #	175	0 (0.0)	29 (16.5)
Low WBC ($\times 10^9/L$)	<4 * <2.5 #	46	1 (2.2)	8 (17)
High WBC ($\times 10^9/L$)	>30 * >20 #	27	0 (0.0)	5 (18.5)
Low neutrophil count ($\times 10^9/L$)	<1 * <1.5 #	13	0 (0.0)	4 (30.7)
High neutrophil count ($\times 10^9/L$)	>20 * >15 #	34	0 (0.0)	10 (29.4)
High lymphocyte count ($\times 10^9/L$)	>5 * >3.5 #	35	1 (2.8)	15 (42.8)
High monocyte count ($\times 10^9/L$)	>1.5 * >1 #	26	1 (3.8)	9 (34.6)
Low platelet count ($\times 10^9/L$)	<100 * <150 #	97	0 (0.0)	28 (27.8)

* According to consensus criteria, # According to our laboratory criteria

Table 4 shows analysis of the samples that required smear review by one of the 2 sets (either the Consensus Group or our laboratory) but not the other. For example, 43 cases had MCV between 70 and <75 fL. According to the consensus group criteria, all 43 cases would require smear review. Four of them were actually false positives. On the other hand, these 43 cases would not have their smear reviewed according to our laboratory criteria, yet 8 of them were actually false negatives. Another example, 20 cases had MCV between 100 and <105 fL. According to our laboratory criteria, all 20 cases would require smear review and 2 were false positives. On the other hand, these 20 cases would not call for smear review according to Consensus Group criteria. Actually, none of them fell into the false negative group (they did not show any positive finding).

3.3 Analysis of false negatives and false positives

We conducted analysis of false negatives encountered on applying either set of criteria and found that RBC morphology was the most frequent cause with both Consensus Group criteria and our laboratory criteria. Within the red cell morphology, microcytosis and hypochromia were the most commonly missed, being 29 cases by the consensus and 6 cases by our laboratory criteria (Table 5). With our laboratory criteria, the false negative results that had abnormal PLT morphology were just two cases of giant platelets, while application of Consensus Group criteria missed PLT clumps in 21 cases and giant platelets in 16 cases. Regarding abnormal WBC morphology 1 case of toxic granulation and 2 cases of immature granulocytes were missed by our laboratory criteria. On the other hand with Consensus Group criteria, fifteen cases were missed for immature granulocyte, 7 cases with toxic granulation and 5

cases were missed for atypical lymphocytes. No cases of blasts were missed by applying either set of criteria (Table 5).

Table 5: Analysis of false negative findings

Smear findings	Consensus criteria		Laboratory criteria	
	n	%	n	%
WBC abnormalities	27	18.48	3	12.49
Metamyelocytes, myelocytes, promyelocytes	15	10.27	2	8.33
Blasts	0	0	0	0
Atypical lymphocytes	5	3.42	0	0
Toxic granulation	7	4.79	1	4.16
RBC abnormalities	82	56.16	19	79.17
NRBC	2	1.37	0	0
Microcytes	29	19.86	6	25
Hypochromia	29	19.86	6	25
Anisocytosis	17	11.64	4	16.17
Target cells	2	1.37	2	8.33
Spherocytes	2	1.37	1	4.17
Rouleaux	1	0.68	0	0
PLT abnormalities	37	25.34	2	8.33
Platelet clumps	21	14.38	0	0
Giant platelets	16	10.96	2	8.33
Total false negative occurrences	146	100	24	100
Total false negative cases	77	9.25	13	1.62

Table 6 shows the top six causes of the false positive results with the application of each set of criteria. Platelet flags (32 cases) were the most common cause of false positive smear reviews with the consensus group criteria. On the other hand, with our laboratory criteria, platelet flags (32 cases) came third as a cause of false positive smear review after platelet count criteria and low Hb levels (46 and 35 cases, respectively).

Table 6: Top 6 rules that cause false positive results with the Consensus Group and our laboratory criteria

Criteria	Rate of occurrence	
	n	%
Consensus group criteria		
1-PLT clump and giant PLT flags	32	34.40
2-PLT<100 or >1000 ($\times 10^9/L$)	14	15.05
3-TLC < 4 or > 30 ($\times 10^9/L$)	13	13.97
4-NRBC flag	12	12.90
5-Immature granulocyte flag	12	12.90
6-Monocytes >1.5 ($\times 10^9/L$)	10	10.75
Total false-positive occurrences	93	100
Total false positive cases	80	10
Laboratory criteria		
1-PLT < 150 or > 610 ($\times 10^9/L$)	46	26.14
2-Hb < 9 (F) or < 10 (M) gm/dL	35	19.88
3-PLT clump and giant PLT flags	32	18.18
4-Lymphocytosis >3.5 ($\times 10^9/L$)	25	14.20
5-Neutrophil < 1.5 or >15 ($\times 10^9/L$)	20	11.36
6-TLC >20 or < 2.5 ($\times 10^9/L$)	18	10.22
Total false-positive occurrences	176	100
Total false positive cases	134	16.75

3.4 Additional positive smear cases

Table 7 shows the positive findings seen in the additional 51 samples were analyzed. Regarding WBCs abnormalities, left shift (promyelocytes, myelocytes, metamyelocytes) occurred 11 times, blast cells 8 times and atypical lymphocytes 6

times. The RBCs abnormalities included normoblasts (19 cases), schistocytes (4 cases), sickle cells (3 cases), ovalocytes (2 cases), target (2 cases), tear drops (2 cases) and malaria (1 case). As to PLT abnormalities, giant platelet occurred 10 times and platelet clumps 4 times.

Applying both sets of criteria, all 51 cases required smear review as they triggered one or more of the rules. For example the cases with immature granulocytes triggered the low platelet rule of the Consensus Group in 5 cases, generated immature granulocytes flag in 4 cases and triggered the RDW rule in 3 cases. On the other hand, the same positive finding (the presence of immature granulocytes) triggered the low PLT rule of our laboratory criteria in 6 cases, the RDW rule in 5 cases and low Hb rule in 5 cases.

Table 7: Description of positive findings in the additional 51 samples and the rules they most commonly triggered

Findings	n	Consensus Group rules, n (%)	Our laboratory rules, n (%)
Metamyelocytes, myelocytes, promyelocytes	11	Low PLT rule 5/11 (45%) IGs flag 4/11 (36%) RDW rule 3/11 (27%)	Low PLT rule 6/11 (54%) RDW rule 5/11 (45%) IGs flag 4/11 (36%)
Blast cells	8	Blast flag 6/8 (75%) Low PLT rule 6/8 (75%) Giant PLT flag 3/8 (37%)	Blast flag 6/8 (75%) Low PLT rule 6/8 (75%) Low Hb criteria 6/8 (75%)
Atypical lymphocytes	6	Variant lymph flag 3/6 (50%) RDW rule 2/6 (33%) Low PLT rule 1/6 (16%)	Variant lymph flag 3/6 (50%) RDW rule 3/6 (50%) Low PLT and high TLC 1/6 (16%)
Normoblasts	19	Giant PLT flag 11/19 (57%) RDW rule 11/19 (57%) Low PLT rule 9/19 (47%)	Low Hb rule 14/19 (73%) PLT rule 15/19 (78%) Giant PLT flag 11/19 (57%)
Schistocytes	4	Giant PLT flag 4/14 (28%)	RDW rule 7/14 (50%)
Sickle cells	3	RDW rule 3/14 (21%)	Low Hb rule 6/14 (42%)
Ovalocytosis	2	Dimorphic flag 3/14 (21%)	Giant PLT flag 4/14 (28%)
Target cells	2		
Tear drops	2		
Malaria	1		
Giant platelets	10	Giant PLT flag 5/10 (50%) Low PLT rule 4/10 (40%) Low Hb rule 2/10 (20%)	Low Hb rule 6/10 (60%) Giant PLT flag 5/10 (50%) Low PLT rule 5/10 (50%)
Platelet clumps	4	Giant PLT flag 4/4 (100%) Cellular interference 4/4 (100%) Low PLT rule 3/4 (75%)	Giant PLT flag 4/4 (100%) Cellular interference 4/4 (100%) Low PLT rule 4/4 (100%)

3.5 Positive predictive value of suspect flags

We analyzed the suspect flags obtained for all 851 samples to evaluate the positive predictive value of each flag for its

particular abnormality. For example, immature granulocytes flag appeared 67 times, it was associated with the presence of immature granulocytes in the smear in 33 of these 67 times. Then the abnormality-specific PPV of the IG flag was calculated (49.25%). The PPV for blast flag, atypical lymphocyte flag, NRBC flag, giant platelets and platelet clumps were 41.66%, 30.76%, 21.91%, 33.70% and 22.22%, respectively. Table 8 also shows the number of occurrences of these flags without the corresponding abnormality on smear review (false positive flag).

Table 8: Positive predictive value of each suspect flags for its specific abnormal finding

Suspect flags	Specific abnormal finding in smear		Abnormality-specific PPV of flags
	Present	Absent	
Immature granulocytes			
Present	33		49.25%
Absent	39 [#]	34 [*]	
Blasts			
Present	10		41.66%
Absent	3 [#]	14 [*]	
Atypical lymphocytes			
Present	4		30.76%
Absent	13 [#]	9 [*]	
NRBC			
Present	16		21.91%
Absent	27 [#]	57 [*]	
Giant platelets			
Present	30		33.70%
Absent	30 [#]	59 [*]	
Platelet clumps			
Present	14		22.22%
Absent	16 [#]	49 [*]	

*False positive flagging # False negative flagging

4. Discussion

The commission on laboratory accreditation of the College of American Pathologists (CAP) requires each laboratory to have criteria about when to perform manual smear review following automated blood count analysis [2].

In the present study, the use of the Consensus Group criteria generated quite high review criteria (54.25%). This review rate is, however, comparable to the 46.06% overall review rate reported by Comar et al., (2014) [6], applying Consensus Group criteria on results obtained by 2 analyzers. It is even closer to the 54.45% microscopic review rate they reported on using the XT-2000i than to the 43.86% of the XE-2100D.

On the other hand, it is much higher than reported by other investigators who also applied the Consensus Group criteria. Wei et al., (2010) [7] review rates were 37.94%, 35.56%, 33.44% and 37.94% respectively on Siemens Advia 2120 (Siemens Diagnostics, Tarrytown, NY), Sysmex XE-2100, Sysmex XT-1800i and Sysmex XS-800i (Sysmex, Kobe, Japan). Pratumvinit et al., (2013) [8] reported a rate of 29.33% on applying Consensus Group on results obtained from the SysmexXE-5000 and Coulter LH750 analyzers. Using the same guidelines, a South Korean group [9] reported manual smear review rates of 28.6%, using the Sysmex XE 2100, and rates of 22.8% and 20.2% for the

UnicelDxH 800 (Beckman Coulter Inc., Fullerton, CA) and Advia 2120 analyzers, respectively. In the study of Froom et al., (2009) [10] the smear review rate was 13.9% on using Consensus Group criteria.

The high review rate in the present study can be explained by a number of factors that were confirmed in other studies. First, our laboratory serves a large university hospital, a tertiary care center. Patients who seek medical treatment at large hospitals tend to have illnesses of greater severity which may generate greater number of abnormal RBC results that are mostly positive and easy to trigger the rules. This issue was raised by the CAP Q- probe study program (2006) and emphasized by Wei et al., (2010) [7]. The latter stressed that not only the type of instrument, but also the type of hospital can lead to different performance if the smear review criteria were introduced without any adoption.

Second, the present study was conducted on inpatients. Patient composition (inpatients versus outpatients) has been recognized as a factor affecting review rate [9]. This probably explains the significantly lower review rate (13.9%) reported by Froom et al., (2009) [10] on analyzing outpatient samples.

A third factor which could have helped lower the review rate is the application of delta check. Unfortunately, this could not be used because of limitations of LIS currently in use (upgrading is in place). All patients' samples were thus considered as first time samples. At least a proportion of the samples that required smear review would not have required it if delta check criteria were applied and the difference was found within acceptable limits. This was the same for Comar et al., (2014) who did not apply delta check rules and reported a comparably high review rate even though their samples were from both inpatients and outpatients.

The review rate when our laboratory criteria were applied was even higher than the Consensus Group criteria (71.0% compared to 54.26%). It is worth noting that these criteria were in use in our laboratory but have not been validated. That is why they can't be considered as optimized criteria and compared to the rates reported by different authors applying their own optimized criteria.

In fact, by using their own laboratory optimized criteria, Pratumvinit et al., (2013) [8] and Comar et al., (2014) [6] lowered their review rate to 23.37%, 37.3% respectively. Similarly Wei et al., (2010) [7] optimized the rules and finalized the criteria suitable for their laboratory. The review rates were lowered to 31.07%, 30.00%, 30.01% and 30.09% with their 4 analyzers, and there was no hematological malignant tumor missed. Using their own review criteria, Leers et al. (2011) [11] found 25.4% of specimens to require microscopic review after initial processing with the XT-2000i and 22.7% after using the CELL-DYNRuby (Abbott Laboratories, Diagnostics Division, Santa Clara, CA). The 13.9% review rate reported by Froom et al., (2009) [10] using the Consensus Group criteria was lowered to 7.2% by using their stringent criteria and to 2.6% by using delta check.

In the Q-probes study, designed by the CAP to benchmark performance in laboratory medicine (2006) [2], each laboratory used the criteria he had in place. These could be Consensus Group criteria, adapted Consensus Group criteria or optimized own laboratory criteria. The rate of manual review varied considerably among participants. The median institution (50th percentile) performed manual reviews of peripheral smears on about one fourth (26.7%) of all specimens submitted for CBC determination. In the 10% of institutions in which the lowest percentage of peripheral blood smears were reviewed manually, participants examined less than 10% of specimens. Among the 10% of institutions in which the largest percentage of smears were reviewed, the review rate was 50% or more.

The false negative rate is of utmost importance in the evaluation of review criteria. It is related to their capability to filter samples with relevant morphological abnormalities (i.e. positive smear findings). In the present study, Consensus Group criteria resulted in 9.25% false negative rates whereas our laboratory criteria were associated with a significantly lower, false negative rate (1.62%), ($p < 0.05$), probably at the expense of a review rate of 71.0%.

Similarly, Comar et al., (2014) [6] reported a 6.73% false negative rate with screening criteria adapted from Consensus Group. On the other hand, a much lower false negative rate (2.22%) was verified by Pratumvinit et al., (2013) [8] with Consensus Group criteria than with their own laboratory criteria (8.09%). After optimization of their laboratory review criteria, they were able to achieve a false negative rate of 2.98%. A false negative rate of <3% was reported by Wei et al., (2010) [7] by applying both the Consensus Group and their optimized criteria to results obtained from 4 different analyzers (Siemens Advia 2120, Sysmex XE-2100, Sysmex XT-1800i and Sysmex XS-800i).

The consensus group considered 5% as the maximum acceptable false negative rate to ensure patient safety [5]. However, we believe, in agreement with Wei et al., (2010) [7] that this 5% is rather high and that false negatives should not exceed 3% with no leukemia or lymphoma missed.

Fortunately, none of the cases with blasts would have been missed by either the Consensus Group criteria or our laboratory criteria. However, since there were only 5 cases with blasts in the 800 samples, 8 additional samples with blasts were analyzed to check our criteria. Still, no cases with blasts were missed, although the blast flag was only detected in 10 of 13 cases. Each of the 3 cases that did not show the blast flag triggered more than one rule in both sets of criteria including low Hb level, high lymphocyte count, low PLT, low neutrophil count, normoblast flag and cellular interference flag.

The same observation was reported by Pratumvinit et al., (2013) [8]. No cases of blasts would have been missed by the Consensus Group criteria or their own laboratory criteria. The blast flag was only detected in 11 of 12 cases. One additional case was triggered by low WBC counts, a low neutrophil count, and low PLT count criteria. Froom et al., (2009) [10] reported that a blast flag was found in 22 cases of 24 acute leukemia patient studied. They added that 19 of

the 24 cases were also triggered by the criteria of basic parameters alone (low Hb, low WBC or high WBC or low PLT).

On the other hand Comar et al., (2014) [6] found that one false negative sample contained blasts. They stressed that it is unacceptable to fail to detect cases of undiagnosed hematological malignancies; and advised each institution to evaluate the need to perform smear reviews in all patients in the hematology unit even at the expense of an increased microscopic review rate.

Inversely related to the false negative rate is the negative predictive value. In the present study, the negative predictive value showed that in 78.96% of the times in which the Consensus Group criteria did not indicate the need for blood smear review, the sample analyzed really did not contain any positive smear findings.

In the present study, the observed sensitivity was 81.13%. Out of 434 samples with positive smear findings; 354 were correctly screened by applying the Consensus Group criteria. An even higher sensitivity (97.09%) was achieved using our laboratory criteria. The specificity of Consensus Group criteria (78.32%) was higher than that of our laboratory criteria (62.04%). However, neither the sensitivity nor specificity showed statistically significant difference.

It is to be noted that with both criteria sensitivity was higher than specificity. This was particularly evident with our laboratory criteria and in line with the higher review rate observed with our laboratory criteria. This is in agreement with Comar et al., (2014) [6] who reported that sensitivity (77.19%) was higher than specificity (67%), again in accordance with their high review rate.

Analysis of false negatives showed that they were mainly related to RBCs abnormalities in both Consensus Group and our laboratory criteria. In our laboratory criteria, they were - by far- more prevalent than WBCs or PLT abnormalities.

The contribution of the individual rules to false positives was different on applying Consensus Group and our laboratory criteria. This can simply be a reflection of the fact that the threshold to trigger some of rules is not the same for both sets of criteria. In the current study, our laboratory criteria were the same as those of the Consensus Group for no or incomplete differential counts, high hemoglobin values and the presence of suspect flags. We used more sensitive criteria leading to more peripheral blood smear reviews for high WBC count, all WBC differential counts, RDW, low hemoglobin value and PLT counts. However, our criteria were stricter for low MCV and low WBC count. Actually, with our laboratory criteria the MCV has to be <70 fL and the low WBC count <2.5 x 10⁹/L to trigger the rule and call for smear review.

Theoretically, the use of our laboratory stricter criteria has the potential to avoid some of false positives encountered with the Consensus Group criteria while it carries the risk of missing some true positives. In our study, of the 43 cases that fell between the two cutoffs for low MCV (to be smear reviewed according to Consensus Group but not according to our laboratory criteria), 8 were false negatives with the <70

fL criterion while 4 were false positives by the <75 fL criterion.

This was the case in Pratumvinit et al., (2013) [8] study where the false negatives increased after they adjusted the threshold to the stricter <70fL. However, they did not worry about this increase as they believe that in patients with microcytosis, the detailed blood smear review may not be useful to discriminate between iron deficiency anemia, thalassemia minor, and anemia of chronic disease. Froom et al., (2009) [10] went even further and did not perform slide review in patients with microcytosis. They rather perform reflex testing using hemoglobin electrophoresis if not previously done in samples with MCV <75 fL and RBCs $\geq 4.5 \times 10^{12}/L$. This led to a new diagnosis of β -thalassemia minor of 67.8% the patients they tested.

The lower cutoff value defined by our laboratory criteria for low TLC criteria resulted in missing 1 case as false negative while avoiding 8 cases that were false positive with the Consensus Group criteria.

On the other hand, our criteria were more sensitive than the Consensus Group for low Hb value, high MCV value, RDW, high WBC count, low and high neutrophil count, high lymphocyte count, high monocyte count and low platelet count. Altering the threshold in that direction has the potential to duly request a smear review for some of the cases that were considered as false negatives with the less sensitive Consensus Group criteria. However, it carries the risk of unduly smear review cases that have no positive smear findings i.e. increase the false positives.

For all parameters, the use of more sensitive laboratory criteria resulted in more false positive cases. This was not accompanied by correction of the status of any of the false negatives by the Consensus Group criteria except for high lymphocyte and monocyte count parameters (one case for each parameter).

Although it seems that our more sensitive laboratory criteria just increased false positives without substantially ameliorating the false negatives encountered with the less sensitive Consensus Group criteria, yet, verifying such assumption requires appropriate further statistical analysis which was not attempted in the present study.

Fortunately all 51 additional samples that had positive smear reviews triggered the rules in both Consensus Group and our laboratory criteria. Cases were further analyzed to define which rule(s) were triggered. Most cases triggered more than one rule. However, it is worth noting that one of the 2 cases with ovalocytosis was not triggered by any of RBC parameter criteria as hemoglobin, MCV and RDW but it was reviewed because of the presence of absolute lymphocytosis $>3.5 \times 10^3 /\mu L$ and giant platelet flag. This case would have been missed if the unrelated accompanying lymphocytosis or giant platelet flag were not present.

One of the main rules for smear review, irrespective of the criteria a laboratory uses, relates to instrument suspect flags. The instruments use flags to notify the user that the automated differential WBC count may not be correct and

requires review. The International Consensus Group for Laboratory Hematology Review noted that the false positive rate is largely due to instrument suspect flagging. This was seen across all instruments used in the study which represented the majority of multiparameter hematology analyzers used in hematology laboratories at the time of the study. They explained that these analyzers are intended to be used as screening devices and to flag suspect abnormal samples for further review [5].

Most automated cell counters are installed with factory-set or factory-recommended settings for the thresholds. It is then up to the individual laboratories to adjust the thresholds to the clinical needs of their patients and clinical staff [14]. As is the case with many laboratories, we use the factory-set defaults.

In the present study, instrument flags were responsible for 60.2% of the false positive results on applying Consensus Group criteria but only 18% on applying our laboratory criteria. The difference was probably because of the more sensitive threshold set for many of the rules in our laboratory criteria. This increased the percentage of false positives due to rules other than instrument flags.

Because flags are supposed to have some level of specificity for the presence of certain abnormal findings, we calculated the abnormality-specific positive predictive value of some flags. The positive predictive values varied between 21.9% (PPV of NRBC flag for the presence of nucleated red blood cells and 49.3% (PPV of IG flag for the presence of promyelocytes, myelocytes or metamyelocytes). This is, however, different from results obtained by Sireci et al., (2010) [14] on Sysmex XE-2100 and XE-5000 for WBC specific flags, where PPV% varied between 5.4% for blast flags to 33% for IG flag.

Comparisons of the efficiency rates of flagging reported in the literature on automated cell counters are inherently difficult because of the use of different types and models of automated cells counters in different patient populations [14]. In general, unless very well tested, a lower positive predictive value of serious flags (as the blast flag) is safer, given the clinical need to detect all cases of blasts. It is important to stress again that -in the present study- none of the cases with blasts were missed (if the blast flag was not triggered another rule was triggered).

Whereas the PPV% of the atypical lymphocyte flag was 30.8% meaning that when the flag is generated there is a 30.8% chance of finding atypical lymphocytes in the smear, yet this flag was not generated in 13/17 (76.47%) cases where atypical lymphocytes were actually present in the smear.

The difficulty in correctly classifying lymphocytes findings either as within reference range or as atypical was pointed out by Koepke (1977) [15] who reported a coefficient of variation of 88% for the atypical lymphocyte count. More recently van der Meer et al., 2004, 2007 [16], [17] found significant interobserver variability in the classification of lymphocytes as atypical or within reference range.

Furthermore, when the same cell was shown twice, it was classified by 34% of the observers as a different subtype.

Because of the limited reproducibility of the atypical lymphocyte count, we agree with Pratumvinit et al., (2013) [8] that missing some cases with increased atypical lymphocytes was acceptable. However, analysis of false negatives showed that 4 of the 9 cases with atypical lymphocytes that were not triggered by atypical lymphocyte flags, were triggered by other rules in the Consensus Group criteria and all were triggered in our laboratory criteria.

In our study, we did not include band cell counts as positive smear findings. The International Society of Laboratory Hematology recommends that the use of band cell counts should be in accordance to laboratory standard operating procedures. We opted not to consider increased bands as positive smear findings in view of the questionable clinical utility besides the differences in defining what is to be considered a band. In fact, studies indicate that "the band count is a nonspecific, inaccurate, and imprecise laboratory test" with a review of the literature providing little support for clinical utility of the band count in patients over 3 months of age [12].

As our laboratory does not test samples from newborn, we were not concerned that underreporting of band forms would have an adverse clinical effect. This is in agreement with Ward et al., (2012) [13] and Sireci et al., (2010) [14]. On the other hand, Comar et al., (2014) [6] considered a band count of >8% as a positive smear finding. They explained that a substantial proportion of doctors in Brazil claimed that there is an association of a nuclear shift to the left of neutrophils with infectious and inflammatory conditions. In their study, band count represented the commonest cause of false negatives.

Although in the present study, we compared the performance of the two sets of criteria, the Consensus Group criteria and our laboratory criteria, and tried to describe the consequences of the different triggering thresholds in some rules, yet, solid conclusion could not be made in this regard as it required more sophisticated statistical analysis. In addition, optimization of the criteria was beyond the scope of the study.

Moreover we can safely conclude that the use of Consensus Group criteria for smear review yielded a significantly lower review rate than our laboratory criteria. These rates, though relatively high, are acceptable as long as it is within the capability of the staff and does not affect the TAT. Even with the higher false negative rate while using the Consensus Group criteria we did not miss any serious finding.

There are a number of limitations in the current study. The study did not include outpatients and it was conducted on samples from the Department of Internal Medicine that are particularly prone to have more abnormal CBCs and consequently a higher review rate. The findings of the study may not be valid for samples from inpatients in the Surgical Departments (general surgery, orthopedics, etc). These patients, apart from the cause of admission to hospital, are generally healthier and have less abnormal CBCs.

Samples were collected from routine workload over a period of 3 months, so some uncommon positive findings were not adequately represented in the study such as malaria, RBC autoagglutination and rouleaux formation. The study did not either include cases where diagnosis relies heavily on morphology such as autoimmune hemolytic anemia or microangiopathic hemolytic anemia.

In the view of the present study, we would recommend the extension of the study on inpatient samples to include the abnormalities under-represented or not represented in the present study. Validation of the criteria for smear review on outpatient samples as they are expected to give a lower review rate and ameliorate the TAT. Optimization of the review criteria to suit the population served and the analyzers used in the laboratory with inclusion of the Delta check should be attempted, aiming at reducing the number of blood films examined without missing important diagnostic information. This will improve the efficiency and reliability of the CBC results directly released without a smear review. When a specimen is flagged and smear review is indicated, it is prudent to review the smear thoroughly for any morphological abnormality and not only for the abnormality suggested by the flag.

5. Disclosure

The authors state that they have no conflict of interest.

References

- [1] D. Ryan, "Examination of the blood cells," in Williams Hematology, K. Kaushansky, M. Lichtman, E. Beutler, T. Kipps, J. Prchal, U. Seligsohn, (eds) Eighth Edition. The McGraw-Hill Companies, 2010.
- [2] A. Novis, M. Walsh, D. Wilkinson, et al, "Laboratory productivity and the rate of manual peripheral blood smear review. A College of American Pathologists Q-probes study of 95141 Complete blood count determinations performed in 263 Institutions," Archives of Pathology & Laboratory Medicine, (130), pp. 596-601, 2006.
- [3] S. Roohi, "Implementation of international slide review criteria for improving the efficiency of the hematology laboratory," Apollo Medicine, (7), pp. 286-289, 2010.
- [4] A. La Porta, A. Bowden, S. Barr, "Workflow improvement and impact of the new Beckman Coulter LH 1500 high throughput automated hematology workcell," Laboratory Hematology, (10), pp. 95-101, 2004.
- [5] P. Barnes, S. Mcfadden, J. Machin, E. Simson, "The International Consensus Group for Hematology Review: Suggested Criteria for Action Following Automated CBC and WBC Differential Analysis," Laboratory Hematology, (11), PP. 83-90, 2005.
- [6] S. Comar, M. Malvezzi, R Pasquini, "Are the review criteria for automated complete blood counts of the International Society of Laboratory Hematology suitable for all hematology laboratories?," Revista Brasileira de Hematologia e Hemoterapia (36), pp. 219-225, 2014.
- [7] C. Wei, W. Wei, W Xin, "Development of the personalized criteria for microscopic review following four different series of hematology analyzer in a Chinese large scale hospital," Chinese Medical Journal, (123), pp. 3231-3237, 2010.
- [8] B. Pratumvinit, P. Wongkrajang, K. Reesukumal, et al, "Validation and Optimization of Criteria for Manual Smear Review Following Automated Blood Cell Analysis in a Large University Hospital," Archives of Pathology & Laboratory Medicine, (137), pp. 408-414, 2013.
- [9] S. Kim, Y. Kim, S. Shin, et al, "Comparison Study of the Rates of Manual Peripheral Blood Smear Review From 3 Automated Hematology Analyzers, Unicel DxH 800, ADVIA 2120i, and XE 2100, Using International Consensus Group Guidelines," Archives of Pathology & Laboratory Medicine, (136), pp. 1408-1413, 2012.
- [10] P. Froom, R. Havis, M. Barak, "The rate of manual peripheral blood smear reviews in outpatients," Clinical Chemistry and Laboratory Medicine, (47), pp. 1401-1405, 2009.
- [11] M. Leers, H. Goertz, A. Feller, "Performance evaluation of the Abbott CELL-DYN Ruby and the Sysmex XT-2000i hematology analyzers," International Journal of Laboratory Hematology, (33), pp. 19-29, 2011.
- [12] P. Cornbleet, "Clinical utility of the band count," Clinics in Laboratory Medicine, (22), pp. 101-31, 2002.
- [13] J. Ward, S. Fertel, B. Bonomo, et al, "The degree of bandemia in septic ED patients does not Predict impatient mortality," American Journal of Emergency Medicine, (30), pp. 181-182, 2012.
- [14] A. Sireci, R. Schlaberg, A. Kratz, "A Method for Optimizing and Validating Institution-Specific Flagging Criteria for Automated Cell Counters," Archives of Pathology & Laboratory Medicine, (134), pp. 1528-1533, 2010.
- [15] JAA. Koepke, "Delineation of performance criteria for the differentiation of leukocytes," American Journal of Clinical Pathology, (68), pp. 202-206, 1977.
- [16] W. van der Meer, C. S. Scott, M. H. de Keijzer, "Automated flagging influences the inconsistency and bias of band cell and atypical lymphocyte morphological differentials," Clinical Chemistry and Laboratory Medicine, (42), pp. 371-377, 2004.
- [17] W. van der Meer, W. van Gelder, R. de Keijzer, H. Willems, "The divergent morphological classification of variant lymphocytes in blood smears," Journal of Clinical Pathology, (60), pp. 838-839, 2007.

Author Profile



Azza S Eldanasoury received her M.B B.Ch degree in Medicine from Ain Shams University in Cairo, Egypt in 1981. She got her MS and MD degrees in Clinical Pathology in 1985 and 1991, respectively. Currently, she is a professor of Laboratory Hematology in the same university. Besides teaching postgraduate students and supervising MS and MD theses, she works in the Ain Shams University Hospital main laboratory. She is an expert in flow cytometry in addition to extensive experience in blood banking. She is a technical assessor with the Egyptian Accreditation Council (EGAC). For three years (2010 - 2012), she worked as undersecretary of the Egyptian Ministry of Health as Director General of the Central Public Health Laboratories.



Noha H Boshnak received her M.B B.Ch degree in Medicine from Ain Shams University in Cairo, Egypt in 2003. She Worked at Ain Shams university Hospital as a Resident Doctor in the Clinical Pathology Department from 2005 till 2007. She got her MS degree in Clinical Pathology in 2007. She worked as an assistant lecturer of Clinical Pathology from 2008 till 2011. She got the MD degree in 2011. Currently, she is a lecturer of Laboratory Hematology in the same university. She worked at bone marrow transplantation unit in Ain shams university hospitals from 2012 till 2014.



Raghda E Abd El Monem received her M.B B.Ch degree in Medicine from Ain Shams University in Cairo, Egypt in 2010. From 2010 till 2011 she stayed as house officer in Ain Shams university Hospital. She Worked at Ain Shams university Hospital as a Resident Doctor in the Clinical Pathology Department from 2012 till 2015. She got her MS degree in Clinical Pathology in 2015. Currently, she is an assistant lecturer of Laboratory Hematology in the same university.

