The Egyptian Journal of Hospital Medicine (January 2023) Vol. 90, Page 1579-1587

Evaluation of Thresholds for White Blood Cell Specific Flags on Sysmex XN 1000 Hematology Analyzer

Azza S Eldanasoury, Christine H Hakim, Rasha A El-Gamal, Noha H Boshnak*
Clinical Pathology, Faculty of Medicine, Ain Shams University, Cairo, Egypt
Corresponding author: Noha H Boshnak, **Tel.:** +20 122 293 2829, **ORCID:** 0000-0003-2787-8908,
E-mail: nhboshnak@hotmail.com

ABSTRACT

Introduction: The Sysmex XN-1000 autoanalyzer is designed to improve the analytical performance for counting blood cell and abnormal cell flags specificity. Each laboratory has its own decision to change the trigger thresholds for many of these flags. This would safely decrease the number of unnecessary blood film reviews, reduce the workload, and improve the turnaround time. **Objective:** The aim of the current study was to assess the performance of the thresholds at which Sysmex XN-1000 white blood cell specific flags ("Blast/Abnormal lymphocytes", "Atypical lymphocytes" "Left Shift", and "Immature granulocytes") are triggered as well as to optimize those trigger thresholds in order to raise the positive predictive value (PPV) of our system of flags for those specific abnormalities.

Materials and Methods: Microscopic slide review was done for 400 blood samples when one or more of the interest flags were triggered using the factory default settings.

Results: For all flags ("Blast/Abnormal lymphocytes", "Atypical lymphocyes" "Left Shift", and "Immature granulocytes"), the sensitivity was excellent (100%, 92.8%, 91%, and 95.8%, respectively) but the specificity was much less (14.2%, 19.3%, 25% and 1.9%, respectively). A statistical method (Youden Index) was applied for optimizing the thresholds of the 4 flags aiming at improving their specificity. Statistically speaking, the optimal thresholds for best efficiency were 290 for the "Blast/Abnormal lymphocytes" flag, 150 for the "Atypical lymphocytes" flag, 190 for the "Left Shift" flag and $0.42\times10^3/\mu L$ for the "Immature granulocytes" flag. **Conclusion:** Considering the clinical impact of the abnormalities that are suspected when a flag is activated, it was prudent to privilege sensitivity over specificity and keep both the "Blast/Abnormal lymphocytes" and the "Immature granulocytes" flagging thresholds at the factory default settings in order not to miss any cases of clinical importance. On the other hand, the thresholds of "Atypical lymphocytes" flag can probably be safely raised to 150 and the "Left Shift" flag to 190, thus reducing the number of unduly triggered samples while maintaining the sensitivity clinically acceptable.

Keywords: CBC, Flags, Sysmex XN, WBC differential count, Sensitivity, Specificity, Ain Shams University.

INTRODUCTION

One of the most commonly requested blood tests in clinical practice is the complete blood count (CBC) [1]. Although it is ideal for each blood count to include a stained blood film examination, unnecessary blood film reviews expand workload and substantially lower laboratory productivity; thus, time- and cost-effective rationalization is required [2].

A major benefit of improved capabilities and performance of automated hematology analyzers recording differential white blood cell (WBC) counts is to decrease the number of blood films requiring microscopic review [3]. Instrument flags are the factor affecting the decision on when to make, stain and examine a blood film. The majority of automated hematology analyzers are equipped with factory-set or factory-recommended values for the flagging thresholds. However, laboratories can adjust those thresholds based on their patients' clinical needs and clinical staff variables [4]. To balance the threat of missing pathological cells with laboratory efficiency, cutoffs must be carefully optimized [5].

The objective of the current study was to assess the performance of the factory-set thresholds at which Sysmex XN -1000 WBC specific flags are triggered in relation to the findings of manual WBC differential count and -if possible- to optimize those thresholds in our laboratory to improve their positive predictive value (PPV) for specific abnormalities, consequently, the number of unnecessary manual differential WBC counts is reduced.

PATIENTS AND METHODS

This study was conducted in the main laboratory of the Tertiary Care Ain Shams University Hospitals that serve a significant number of both inpatients and outpatients.

The study included 400 blood samples from adult patients. The differential WBC count on the samples was requested by the clinician, and one or more of the WBC flags were triggered using the factory default settings, then, a blood film was examined. Over the course of five months, the samples were collected (approximately 20 samples from daily work routine).

According to laboratory procedure, two mL venous blood samples were obtained using K3 EDTA (tripotassium ethylenediaminetetraacetic acid) and processed within two hours of collection.

Automated CBC and differential WBCs count were performed using the XN-1000 automated analyzer (Sysmex, Kobe, Japan). Only the 4 WBC-specific flags with user-definable thresholds were studied: (1) Blasts/Abnormal lymphocytes (Blast/Abn lymph), (2) Atypical lymphocytes (Atypical lymph),

Received:22/08/2022
Accepted:25/10/2022

(3) Left shift and. (4) Immature granulocytes (IG). These flags are produced by scattergram patterns that are characteristic of specific abnormalities. We did not examine WBC count flags that had unadjustable triggering thresholds or flags that suggested potential platelet or red blood cell abnormalities.

Whether a flag is triggered or not is based on a preset threshold called "Q-flag". Sysmex hematology analyzers use flow cytometry grade light scattergram events in a predetermined flagging zone on a scale of 0-300, in increments of 10 and when the grading exceeds the default threshold which is equal to 100, a flag is generated. The Q-flag value concept is applied to the studied flags except for the "IG" flag. The device can count the number of IG cells and the sample is flagged if the count is $\geq 0.10 \times 10^3/\mu L$.

To ensure the proper performance of the hematology analyzer, procedures for quality assurance and control were followed. Also, the analyzer was calibrated on a regular basis and daily whole blood controls were used to monitor performance. During the study there were no reagent changes, instrument repairs, calibration standards or controls changes. Patient samples were never run unless a successful quality control was obtained.

Manual WBC differential and smear review: From each sample a blood film was manually smeared and stained with Leishman stain. A 200-cell differential WBC count was performed for all 400 flagged samples by one of the authors. The slides were examined to detect any abnormal cell populations. The International Consensus Group's guidelines for a positive smear were followed: blasts at 1 or greater, myelocytes/ promyelocytes at 1 or greater, metamyelocytes at 2 or greater, atypical lymphocytes/ abnormal lymphocytes at 5 or greater [3]. The slides that showed no abnormalities at all were scanned again by a senior hematologist to ensure the absence of any of the abnormal findings of interest.

It is to be noted that atypical lymphocytes were defined as having reactive or plasmacytoid morphology, abnormal nuclear shape or nucleoli. The so-called atypical lymphocyte is a non-neoplastic lymphocyte that is detected in peripheral blood and is considered to be a nonspecific reaction to stress caused by a variety of conditions, while, abnormal lymphocytes correspond to lymphoblasts.

Sample classification: Samples were categorized as being "true positive" (TP) for specific abnormality, TP for any abnormality, false positive (FP) for specific abnormality and FP for any abnormality as follows:

1. If a flag was triggered and positive smear result was found for the abnormality related to the flag i.e., specific abnormality, the sample was considered as a "TP for specific abnormality" e.g., Blast/Abn lymph flag was triggered and the smear review showed blast cells.
2. If the flag was triggered and positive smear result was found for any other abnormality, whether specific or not, the sample was considered as "TP for any abnormality" i.e., Blast/Abn lymph flag was triggered and the smear review showed blast cells and/or any other abnormality (atypical lymphocytes for example).
3. The sample was considered as "FP for specific abnormality if a flag was triggered and the smear did not show the specific abnormality for the triggered flag.
4. The sample was considered as "FP for any abnormality" if a flag was triggered and neither specific nor any of the other abnormalities of interest was detected in the smear.

Optimization of the flag threshold: Optimization of the thresholds for triggering each of the 4 flags was attempted using the same 400 samples analyzed. For each sample, the exact numeric value of each flag was retrieved from the autoanalyzer and inserted into an Excel spreadsheet. Optimization started by virtual raise in each flag's cutoff from the factory default setting in increments of 10 units. We began by assuming that the threshold was raised to 110 and decided which of the flags would remain triggered and which would not. Then, we calculated for each flag the number of TP, FP, true negative (TN) and false negative (FN) for the specific abnormality at this hypothetical threshold (110). This was repeated raising the threshold in increments of 10. The optimized threshold for each flag was selected as the threshold with the highest Youden index (YI) where the best sensitivity and specificity values meet. This is further detailed in the statistical analysis section.

After maximizing the YI, we used clinical judgment to decide whether the suggested thresholds resulted in cases with true abnormalities being missed and whether their miss could be tolerated.

Ethical Consideration:

The Ethical Committee at the Faculty of Medicine, Ain Shams University approved the study. This study was conducted in compliance with the code of ethics of the world medical association (Declaration of Helsinki) for human subjects.

Statistical Analysis

MedCalc© version 15.8 (MedCalc© Software bvba, Ostend, Belgium) and XLSTAT© version 2014.5.03 (Addinsoft, Inc., Brooklyn, NY, USA) were used for the analysis of the data. To assess the accuracy of the Q-flag values for each flag for diagnosis of specific abnormalities a receiver-operating characteristic (ROC) curve was generated. A YI optimized method was used whereby the criterion associated with the highest YI was identified as the best cut-off value. The YI is a measure of both sensitivity and specificity, and it is used to summarize an assay's diagnostic efficiency at different cutoffs. When the YI is maximized, the threshold for an assay indicates the optimum performance profile of a test, which is the maximum vertical distance from the

diagonal to the ROC curve and is calculated with the equation: YI = Sensitivity + (Specificity – 1) [4].
The statistical study also included the following calculations for sensitivity, specificity, PPV, and efficiency:

- Sensitivity (%) = TP/ (TP + FN) × 100
- Specificity (%) = TN/ (TN + FP) × 100
- PPV (%) = TP/ (TP + FP) × 100
- Efficiency (%) = (TP + TN)/Total samples evaluated × 100

RESULTS

The 400 blood samples were obtained from 215 (54%) females (F) and 185 (46%) males (M) with M/F ratio of 1:1.2. Their ages ranged from 18 to 90 years with a median of 54 years.

I. Occurrence of the different flags among the studied samples

Among the 400 flagged samples, the flag" Blast/Abn lymph" was seen in 223 cases (55.7%), the flag "Atypical lymph" was seen in 311 cases (77.7%), while 97 cases (24.2%) showed the "Left Shift" flag, and 150 cases (37.5%) showed the flag "IG".

II. Analysis of the smear review findings

Out of the 400 samples, 319 (79.7%) showed positive findings upon smear review: 156 samples (48.9%) showed ≥ 5 atypical lymphocytes; 88 samples (27.5%) had ≥ 1 myeloid precursor; 55 samples (17.3%) had ≥ 5 band cells and 20 samples (6.3%) had ≥ 1 blast cell or abnormal lymphocyte.

III. Performance of the 4 Flags at the Factory Default Settings

Table 1 shows the performance of the different flags at factory default settings. The sensitivity of the flags to detect the presence of the flag-specific abnormality was quite high, ranging from 100% for the "Blast/Abn lymph" flag to 91% for the "Left Shift" flag. On the other hand, the best specificity was obtained for the "Left Shift" flag (25%), followed by the "Atypical lymph" (19.3%), then "Blast/Abn lymph" flag (14.2%) and lastly the "IG" flag (1.9%).

The abnormality-specific PPV of each flag ranged from 68.5% for the" Left Shift" flag to 10.3% for the "Blast/ Abn Lymph" flag. The efficiency of the "Blast/Abn lymph" was 21.9% while that of "IG", "Atypical lymph" and "Left Shift" flags was 62%, 62.3% and 64.9%, respectively. The table also shows the PPV of each flag to detect any abnormality (overall PPV) as well as the overall efficiency. Here, the flagging is considered true when any abnormality is present in the smear whether specific for the flag or not.

Table 1: Performance of the 4 flags at the factory default settings

Abnormality	Sensitivity %	Specificity %	PPV%	Efficiency %	YI %	Overall PPV %	Overall Efficiency %
Blast/Abn lymph	100	14.2	10.3	21.9	1.14	77.8	70.4
Atypical Lymph	92.8	19.3	61.9	62.3	1.9	60.4	57.9
Left Shift	91	25	68.5	64.9	1.3	100	33
IG	95.8	1.9	63.4	62	0.4	80	32

PPV: Positive predictive value, YI: Youden-index, IG: Immature granulocytes.

IV. Optimization of the flags' thresholds

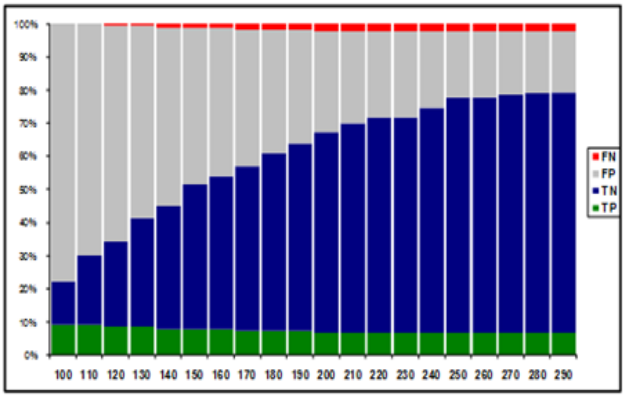
The optimal threshold for each of the flags for the specific abnormality was selected as the threshold with the highest YI among all thresholds including the factory default settings. These corresponded to Q-flag values of 290, 150, and 190 for the "Blast / Abn lymph" flag, the "Atypical Lymph" flag and the "Left Shift" flag respectively. For the "IG" flag the optimized point was at a count of 0.42 $x10^3/\mu L$.

The performance of the different flags at their optimized thresholds is displayed in Table 2 and Figure 1, 2. As expected, the abnormality-specific PPV of all flags was improved at optimized thresholds. However, the overall PPVs were improved by optimization only for the "Blast/Abn lymph" and the "Atypical lymph" flags.

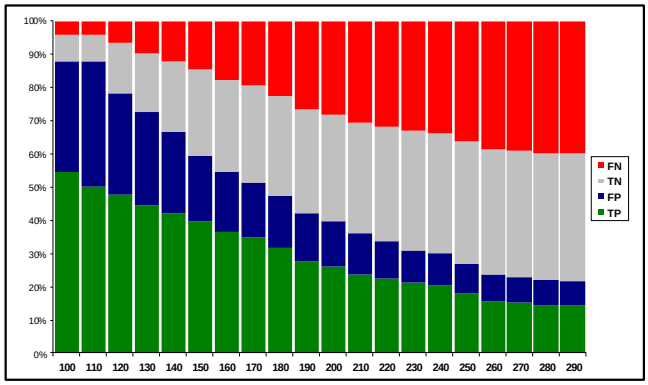
Table 2: Performance of the 4 flags at the optimized settings

Flag	Q	Sensitivity %	Specificity %	PPV %	YI	Efficiency %	Overall PPV %	Overall Efficiency %
Blast/Abn lymph	**290**	**75.0**	**79.8**	**26.8**	**1.55**	**79.4**	**78.6**	**36.3**
Atypical Lymph	**150**	**72.8**	**56.3**	**66.5**	**1.29**	**65.3**	**63.8**	**55.9**
Left Shift	**190**	**52.5**	**61.1**	**69.6**	**1.14**	**55.7**	**81.3**	**55.7**
IG	**0.42**	**55.4**	**77.6**	**79.7**	**1.33**	**64.0**	**67.5**	**49.3**

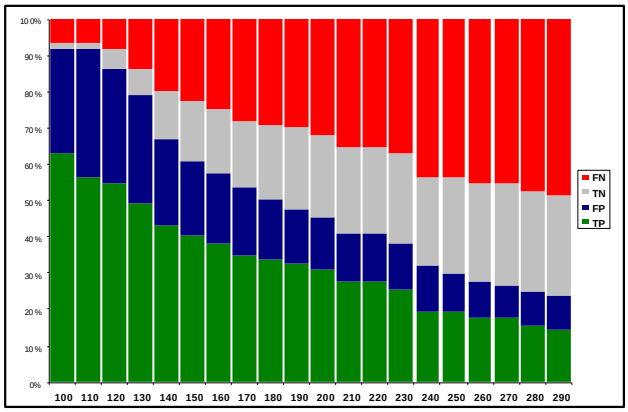
PPV: Positive predictive value, YI: Youden-index, IG: Immature granulocytes



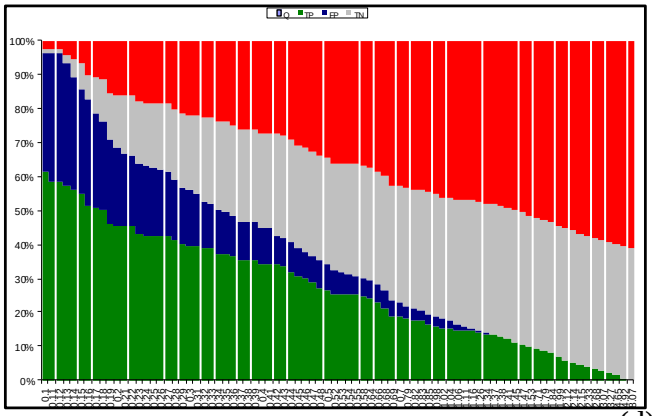
(a)



(b)



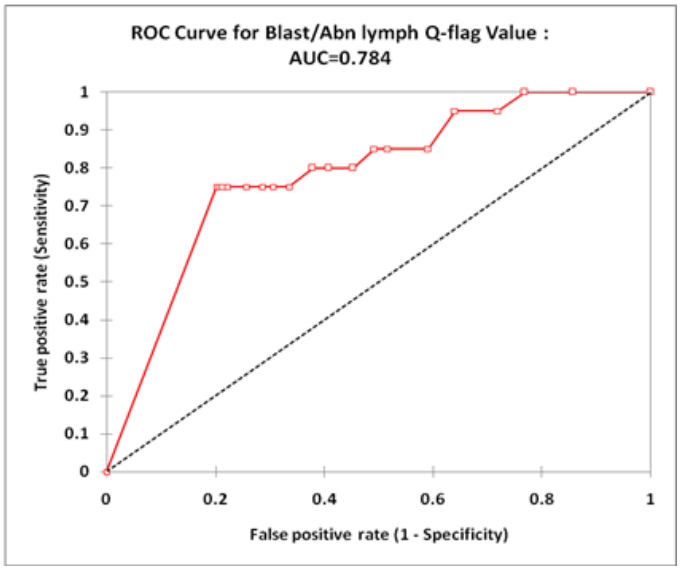
(c)



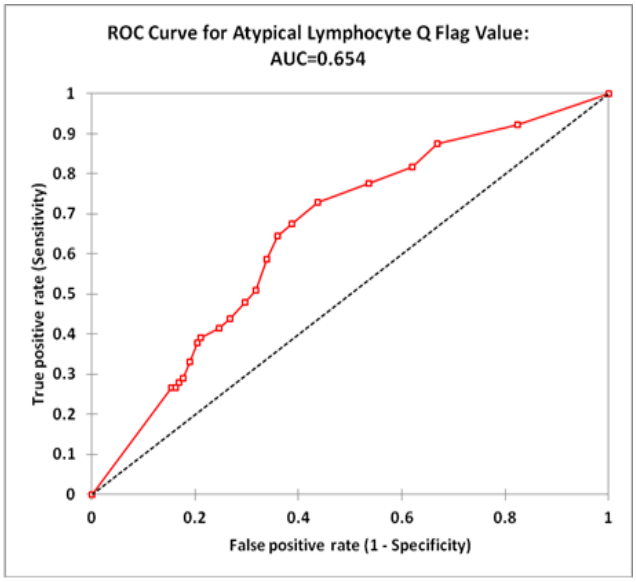
(d)

(d)

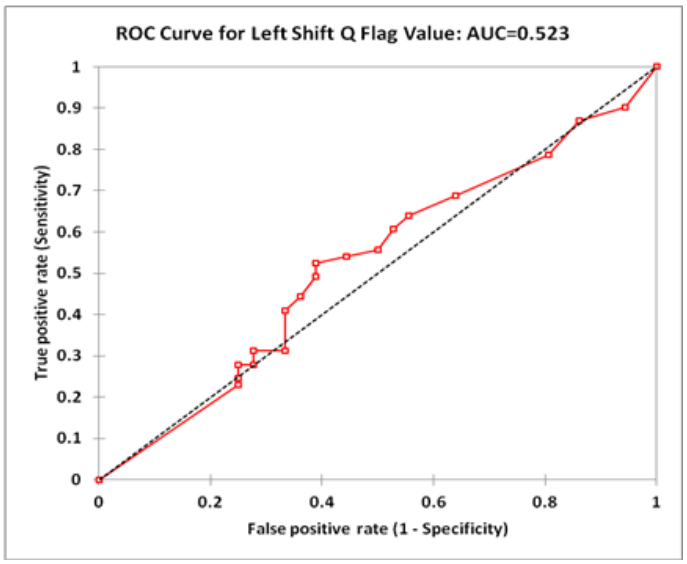
Figure (1): True positive (TP), true negative (TN), false positive (FP) and false negative (FN) rates for:
(a) “Blast/Abn lymph” Q-flag values for its specific abnormality.
(b) Atypical lymphocyte” Q-flag values for its specific abnormality.
(c) “Left Shift” Q-flag values for its specific abnormality.
(d) Immature granulocyte count for its specific abnormality.



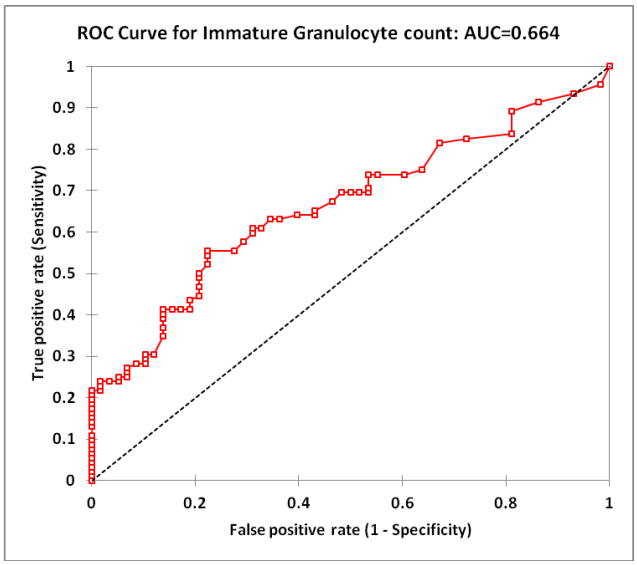
(a)



(b)



(c)



(d)

Figure (2): Receiver operating characteristic (ROC) curve for the diagnostic accuracy of "Blast/Abn lymph" Q-flag value (a), "Atypical lymp" Q-flag value (b), "Left Shift" Q-flag value (c) and IG count (d)
IG: Immature granulocytes

VI. Missed cases after optimization

To further investigate the effect of optimization from a clinical point of view, we analyzed the cases that were missed as a result of the new cutoff values. We found that we missed 5 cases with blast cells, the 1st case generated a Q-flag value of 120, and the others generated Q-flag values of 140, 170, 200 and 210. We also missed 46 cases with atypical lymphocytes, 34 of them had 5 atypical lymphocytes, 8 of them had 6 atypical lymphocytes and 4 cases had 7 atypical lymphocytes. For the "Left Shift" flag we missed 29 cases, all of them contained 5% band cells except for 6 cases which had from 6-10% band cells. We also missed 41 cases that contained ≥ 1 myeloid precursor. This is shown in Figure 3.

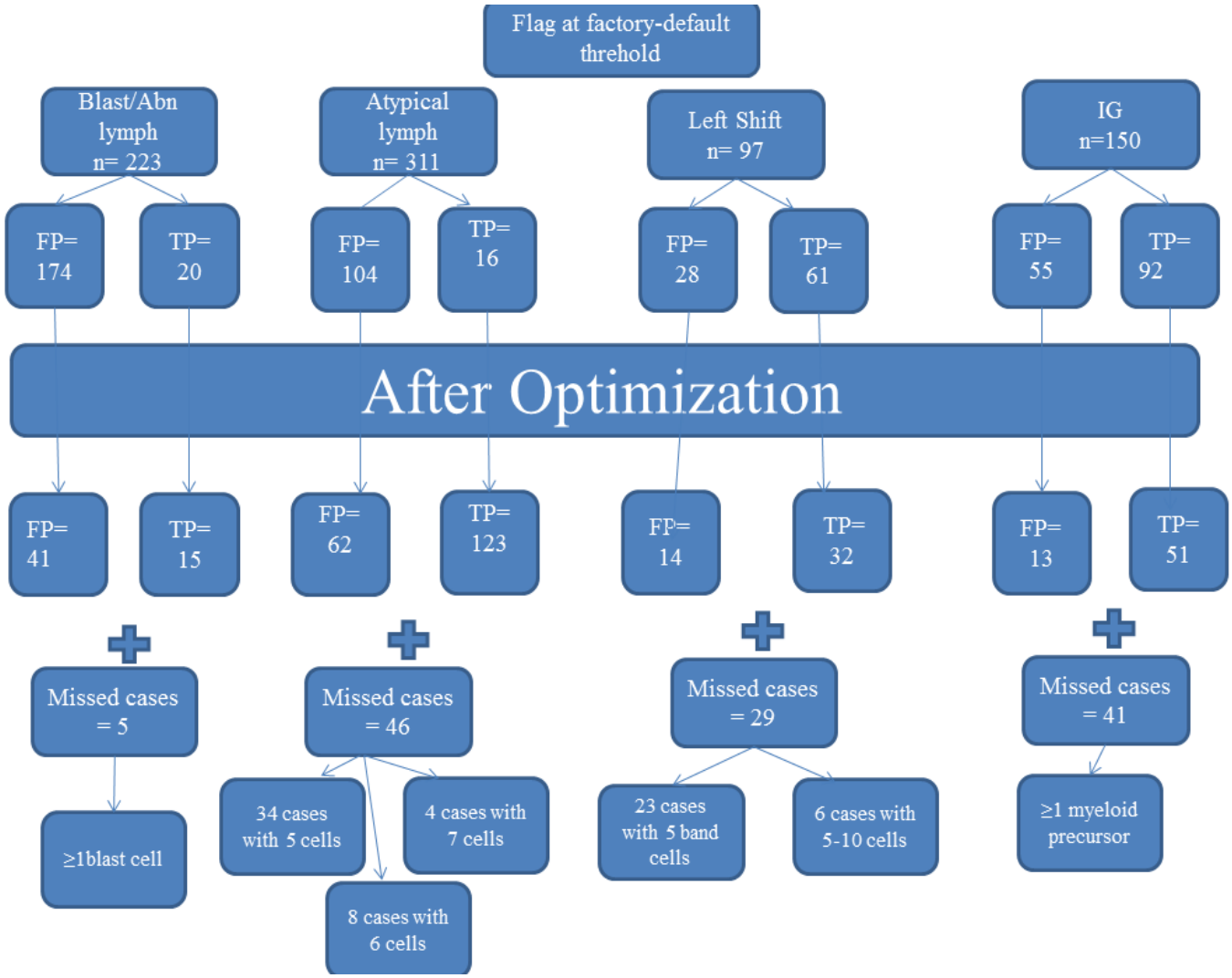


Figure (3): Number of FP and TP cases at the factory-default settings and number of FP, TP and missed cases after YI optimization. Note that the optimized thresholds for "Blas/Abn Lymph" and "IG" flags were not accepted when clinical relevance was considered.

FP: false positive, TP: true positive , YI: Youden-index, IG: Immature granulocytes.

DISCUSSION

Reporting prompt and accurate blood cell counts with clearly defined turnaround time is essential because these results often have an impact on medical decision making. As clinical hematology laboratories in tertiary hospital settings frequently encounter large number of blood samples, improving the workflow efficiency without compromising the diagnostic sensitivity has always been a challenge [6,7]. In the current study, we first assessed the performance of the factory defined thresholds at which Sysmex XN-1000 WBC specific flags are triggered in relation to manual differential count findings and optimized those thresholds to raise the PPV of our flag system for their specific abnormalities, thus, the number of unnecessary manual differential WBC counts decreased.

The flag "Blast/Abn lymph", which is triggered when there is a possibility that blasts or immature cells are present, was obtained in 55.7% of samples studied. Upon comparison with smear review findings, the sensitivity for the flag was excellent (100%) but its abnormality specific PPV was weak (10.3%) due to lack of specificity (14.2%). The sensitivity revealed in our study was close to that found by **Becker *et al.*** [8] who

studied 765 pediatric cases and showed that the XN-1000 analyzer had a high sensitivity for blast identification (99%), as it missed only one case that was proved to contain blasts by microscopic examination. However, their results showed a relatively higher specificity (46%). A lower sensitivity than ours (74.4%) was also revealed in another study [9] of flagging performance of Sysmex XN-1000 analyzers in onco-hematologic patients. The specificity was on the other hand higher than ours (94.8%).

The "Blast" flag reported by other models [three Sysmex analyzers (Sysmex XE-5000)] were also evaluated by **Eilersten *et al.*** [5] who studied the utility of the "Blast" flag as a sufficient indicator for a smear review and found a sensitivity ranging from 71% to 75% at the factory default settings which was lower than that obtained by our XN-1000 analyzer but with a higher specificity ranging from 57% to 69%.

The 14.2% specificity for the "Blast/Abn lymph" flag in the current study is much lower than that for Sysmex XN-1000 in above mentioned studies and is also lower than that reported for the "Blast" flag from other Sysmex models (XE-5000 and XE-2100). Values ranging from 80 to 97.5% were reported in different studies [7,10-14].

The differences in between our study's findings and the published papers may be due to more than one factor. The samples involved in the current study were not selected and represented the usual mix that we receive in our laboratory, while many other studies were conducted on oncohematologic patients. This might have influenced the efficiency of flagging. Another important cause is the difference in the flagging system between the different analyzers where the XN series has a flag called the "Blast/Abn Lymph" flag. The appearance of this flag could be followed by reflex testing of the samples using the WPC (white cell precursor channel), which separates this flag to either "Blast" flag or "Abnormal lymphocyte/Lymphoblast" flag, whereas the XE series gives those two separate flags from the start. Unfortunately, this reflex testing is not done in our laboratory.

It is worth noting that the high specificity mentioned in other studies came at the expense of a lower sensitivity. In view of clinical relevance of blast detection, it is clear that prioritizing sensitivity above specificity is more crucial.

Our results showed that at the factory default settings the sensitivity for the "Atypical lymph" flag was 92.8%, the specificity was 19.3% and the abnormality specific PPV was 61.9%.

The sensitivity for the "Atypical lymph" flag in our study is higher than that for the 5 blood cell analyzers studied by **Depoorter *et al.*** 15] but the specificity was much lower. The 5 analyzers were Abbott Cell-Dyn Sapphire, Sysmex XE-2100, XN-2000, Beckman Coulter DXH-800 and Siemens ADVIA-2120. With 86% sensitivity and 90% specificity, the DXH-800 appeared to be the most efficient analyzer with its 'variant Ly' flag followed by Sysmex XN-2000 with 72% sensitivity and 71% specificity.

Similarly, a study [11] evaluated two analyzers using 43 samples with atypical lymphocytes in peripheral blood. For the XE-5000 analyzer, the "Atypical lymph" flag's sensitivity, specificity, and efficiency were, 51.2%, 95.3%, and 93.4%, compared to 53.5%, 88.1%, and 86.6% for the XE-2100 analyzer. Another study [9] found that the XE analyzer's Atypical lymph flag had a sensitivity of 70% and a specificity of 82.3%, while the XN analyzer's sensitivity and specificity to detect atypical lymphocytes were 78.3% and 95.2%, respectively.

The lower specificity of the "Atypical lymph" flag in our study might reflect the difficulty in interpreting lymphocyte morphology and classifying the lymphocytes as within reference range or atypical as reported by **Koepke** [16]. He observed 88% coefficient of variation for the atypical lymphocyte count using results from proficiency-sample microscopic slides that were distributed to more than 4000 laboratories [16]. Moreover, the assessment of lymphocytes as atypical or within reference range varied significantly among observers, according to a study by **van der Meer *et al.*** [17] when 671 technicians at 114 hospital laboratories received PowerPoint presentations of WBCs. Additionally, when the same cell was displayed twice in the PowerPoint presentation, 34% of the viewers identified it as a different subtype. The heterogeneity of atypical and abnormal lymphocytes and the difficulty to distinguish them microscopically was emphasized by **Jones *et al.*** [18] who admitted that even skilled laboratory technicians are not always sure to which type they belong, leading to a significant number of false positive "Atypical lymph" flags. This may account for the variable results obtained in different studies including ours.

The abnormal lymphocytes flag, based on the manufacturer, particularly denotes lymphocytes that are malignant. However, the term "abnormal lymphocytes" was avoided in the European consensus report on blood cell identification, which instead suggested using the terms "atypical lymphocytes-suspect neoplastic" or "atypical lymphocytes-suspect reactive" to characterize morphological variations from the normal [18].

At the factory default settings, we found that the "Left Shift" flag had a good sensitivity of 91% but with a low specificity of 25%. The abnormality specific PPV was 68.5% and the abnormality specific efficiency was 64.9%.

The sensitivity of the "Left Shift" flag in our study was higher than that stated by **Stramminger *et al.*** [19] (53%), but the specificity and abnormality specific efficiency were lower (92 and 86% respectively). May be the inter-individual variation of the microscopic counts is significant factor in the difference between our study and those obtained in other studies. Moreover, the variations in the band cells' morphological definitions

with the consequent difference in the reference range for band cell count can also be a factor, as in our study we used a reference range of ≥ 5% band cells, while other studies as **Ryan** [20] suggested a reference range of 3%.

We could not compare our results to other studies performed on other Sysmex generations as those generations have the flag "IG" triggered by left shift, band cells and myeloid precursors whereas in XN series the two flags are separate.

The last flag evaluated in this study was the "IG" flag. The percentage of immature granulocytes is a reproducible measure that is helpful in the diagnosis of several diseases [4]. The results of our study showed a high sensitivity of 95.8% for the "IG" flag at the factory default settings but with a very poor specificity of 1.9%. Compared with XN-2000 evaluated in other studies [13,15], we had a better sensitivity (ranging from 28% to 88% in their study) but worse specificity (ranging from 84% to 95% in their study).

The abnormality specific PPV of the ''IG'' flag in our study was 63.4% which was near to that of the XN-2000 analyzer (68%) but higher than Cell Dyn Saphire (45%) and DXH-800 (48%) [13], whereas the abnormality specific efficiency (62%) was lower than the 81% stated in another study [15].

After evaluating the performance of the 4 flags of interest at the factory default settings, we moved to study if we could optimize their thresholds for better specificity without compromising the sensitivity using the YI.

As expected, the optimized thresholds resulted in higher specificity of each flag for its particular abnormality with an obvious reduction in the number of false positive samples; however, a variable decrease in the sensitivity occurred.

Statistically speaking, by raising the threshold of the "Blast/Abn lymph" from 100 to 290, the specificity increases from 14.2% to 79.8% (the FP decreases from 174 cases to 41 cases) and the abnormality specific PPV raised from 10.3% to 26.8%. Unfortunately, this was accompanied by missing 5 cases with blast cells, which were flagged at factory-default settings. This led to keeping our threshold at the factory-default setting for the" Blast/Abn lymph" flag or at most raising it to 110 where we would miss no cases. It is conceivable that a relatively nonspecific flagging is needed to satisfy the clinical requirement to identify all cases of blasts and avoid creating additional false-negative cases. A similar decision to keep the cutoff value for the blast flag at the factory setting (Q=100) was reached by other investigators. **Gossens *et al.*** [21] did not raise the threshold for blast flagging though such elevation would decrease the false-positive rate of their flagging performance by 37% resulting in reduction of smear review by 14%. **Eilertsen *et al.*** [5] found that changing the flagging threshold from 100 to 300 improved the specificity of the blast flag, and therefore decreased the review rate by 12%. However, the number of false-negative cases raised to 19%, which they did not accept. On the other hand, **Sireci *et al.*** [4] raised the flagging threshold of the "Blast" flag from 99 at the factory default settings to 200 and found that there were no cases of more than 1% blasts missed by their optimized settings.

The rise of the threshold of the "Atypical lymph" flag from 100 to 150 increased the specificity from 19.3% to 56.3% (FP decreases from 104 cases to 62 cases). We thought that missing some cases with higher numbers of atypical lymphocytes (n=46) was acceptable due to the inherent limited reproducibility of the atypical lymphocyte count.

Raising the threshold of the "Left Shift" flag from 100 to 190 increased the specificity from 25% to 61.1% (number of FP cases decreased from 28 to 14) and the abnormality specific PPV of this flag from 68.5% to 69.6%. There is still debate on the clinical significance of left shift (increased band cell count). This became less significant once more precise diagnostic indicators for acute inflammation, such as cytokines and procalcitonin, were introduced [19]. Moreover, the review of the literature reveals limited support for the clinical utility of the band count in patients above 3 months of age [4]. Since our laboratory does not receive neonatal blood samples, we were not worried that underreporting of band forms due to changes in our flagging thresholds would have a negative clinical impact. The observation that the overall PPV of the "Left Shift" flag decreased from 100% to 81.3% after optimization raised concern about the possibility of missing some important abnormalities after optimization. Further investigation showed that we would only miss 51 cases with ≥5 band cells, 4 cases with ≥5 atypical lymphocytes and 6 cases with ≥1 myeloid precursor as the rest of the cases would be flagged by their specific flag and would not be missed.

For the "IG" flag, raising the threshold from $\geq 0.10\times10^3/\mu L$ to $0.42\times10^3/\mu L$ decreased the number of FP cases from 55 to 13. The abnormality specific PPV of this particular flag also increased from 63.4% to 79.7%, confirming the conclusion of **Rosenthal *et al.*** [22] that using a threshold of $0.5\times 10^3/\mu L$ could reduce the number of unnecessary reviews for IG. However, though raising the flagging cutoff spared reviewing 40 cases, it missed 41 true positive cases. That is why we opted for keeping the threshold at the factory default settings.

Previous studies reported that flagging sensitivity is influenced by the total WBC count, with reduced sensitivity in leukopenic specimens and lower specificity in specimens with WBC counts higher than $10\times 10^3/\mu L$. Other studies, however, have found only a mild impact of WBC count on overall efficiency [4].

As stated by **Ruzicka *et al.*** [23] the identification of IG by Sysmex XE-2100 was least sensitive in specimens with low WBC counts, whereas the flagging sensitivity for blasts was great in specimens with normal and increased WBC counts. This potential

confounder was not, however, addressed in the current study.

In conclusion, we have assessed the performance of the 4 user definable WBCs specific flags for their specific abnormality at the factory default settings. We have also applied a method for optimizing the thresholds of these flags to improve the specificity of these flags and decrease the number of unnecessary film reviews. We have chosen to keep both the "Blast/Abn lymph" and the "IG" flagging thresholds at the factory default settings to privilege sensitivity over specificity in order not to miss any cases of clinical importance, whereas we chose to raise the threshold of "Atypical lymph" flag to 150 and the "Left Shift" flag to 190 improving their specificity.

We strongly recommend testing the new thresholds for "Atypical lymph" and the "Left Shift" flags using verification set of samples that are independent from the optimization set to confirm their performance. It is to be noted that although this study was conducted on the Sysmex XN-1000 analyzer, the overall optimization approach can be applied to any hematology analyzer that employs quantitative flagging criteria.

Funding sources: No funding received for this study.
Conflicts of interest: The authors declare that they have no conflict of interest.

REFERENCES

1. **Buttarello M , Plebani M (2008):** Automated blood cell counts. State of the art. Am J Clin Pathol.,130:104-16.
2. **Briggs C , Bain B (2017):** Basic hematological techniques. In: Dacie and Lewis Practical Hematology, Barbara J Bain, Bates I, Laffan M and Lewis M, (eds) Twelfth Edition. Elsevier Churchill Living Stone. Chapter 3: 18-49.
3. **Barnes P, McFadden S , Machin S (2005):** The international consensus group for hematology review: suggested criteria for action following automated following CBC and WBC automation analysis. Lab Hematol., 11:83-90.
4. **Sireci A, Schlaberg R , Kratz A (2010):** A Method for Optimizing and Validating Institution-Specific Flagging Criteria for Automated Cell Counters. Arch Pathol Lab Med., 134:1528-33.
5. **Eilersten H, Nina K ,Hagve T (2013):** The usefulness of Blast Flags on the Sysmex XE-5000 is questionable. Am J Clin Pathol., 139:633-40.
6. **Shabnam R (2010):** Implementation of International Slide Review Criteria for Improving the Efficiency of the Haematology Laboratory. Apollo Medicine, 7:286-8.
7. **Seo J, Lee S , Kim S (2014):** Performance evaluation of the new hematology analyzer Sysmex XN- series. Int J Lab Hematol., 37:155-64.
8. **Becker P, Fenneteau O ,Costa L (2016):** Performance evaluation of the Sysmex XN1000 hematology analyzer in assessment of the white blood cell count differential in pediatric specimens. Int J Lab Hematol., 38:54-63.
9. **Furundarena J, Sainz M, Uranga L, Cuevas L, Lopez I, Zubicaray J, Bizjak N ,Robado M (2016):** Comparison of abnormal cell flagging of the hematology analyzers Sysmex XN and Sysmex XE-5000 in oncohematologic patients. Int J Lab Hematol., 39:58-67.
10. **Kang S, Kim H, Ham C, Lee D , Cho H (2008):** Comparison of four hematology analyzers, Cell-Dyn Sapphire, Advia 120, Coulter LH-750 and Sysmex XE-2100, in terms of clinical usefulness. Int J Lab Hematol., 30:480-6.
11. **Briggs C, Linssen J, Longair I , Machin S (2011):** Improved flagging rates on the Sysmex XE-5000 compared with the XE-2100 reduce the number of manual film reviews and increase laboratory productivity. Am J Clin Pathol., 136:309-16.
12. **Briggs C, Longair I, Kumar P, Singh D , Machin S (2012):** Performance evaluation of the Sysmex haematology XN modular system. J Clin Pathol., 65:1024-30.
13. **Hotton J, Broothaers J, Swaelens C , Cantinieaux B (2013):** Performance and abnormal cell flagging comparisons of three automated blood cell counters: Cell-Dyn Sapphire, DxH-800, and XN-2000. Am J Clin Pathol., 140:845-52.
14. **Meintker L, Ringwald J, Rauh M , Krause S (2015):** Comparison of automated differential blood cell counts from Abbott Sapphire, Siemens Advia 120, Beckman Coulter DxH800, and Sysmex XE-2100 in normal and pathologic samples. Am J Clin Pathol., 139:641-50.
15. **Depoorter M, Goletti S, Latinne D , Defour J (2015):** Optimal flagging combinations for best performance of five blood cell analyzers. Int J Lab Hematol., 37:63-70.
16. **Koepke J (1977):** A delineation of performance criteria for the differentiation of leukocytes. Am J Clin Pathol., 68:202-26.
17. **van der Meer W, Scott C , de Keijzer M (2004):** Automated flagging influences the inconsistency and bias of band cell and atypical lymphocyte morphological differentials. Clin Chem Lab., 42:371-7.
18. **Jones A, Tailor H, Liesner R, Machin S , Briggs C (2015):** The value of the white precursor cell channel (WPC) on the S ysmex XN1000 analyzer in a specialist pediatric hospital. Am J Clin Pathol., 68:161-5.
19. **Stramminger G, Auch D, Diem H , Sinhas P (2002):** Performance of the XE-2100 leucocyte differential. Clin Lab Haem., 24:271-80.
20. **Ryan D (2006):** Examination of the blood. In: Williams Hematology, Lichtman MABE, Kipps TJ,Seligsohn U, Kaushansky K, Prchal JT, (Eds) Seventh Edition. Chapter 7: 11-9.
21. **Gossens W, van den Driessche M , Brusselmans C (2000):** Optimization of the flagging criteria on Sysmex SE-9500. Sysmex J int., 10:18-20.
22. **Rosenthal N, Connell B , Brown B (2012):** Evaluation of a new method for the enumeration of Nucleated Red Blood Cells on the new Sysmex XN Automated Hematology Analyzer. Int J Lab Hematol., 34:1-18.
23. **Ruzicka K, Veitl M, Thalhammer-Scherrer R , Schwarzinger I (2001):** The new hematology analyzer Sysmex XE-2100: performance evaluation of a novel white blood cell differential technology. Arch Pathol Lab Med.,125:391-6.