Peer reviewed **ORIGINAL ARTICLE**

REDUCING UNNECESSARY BLOOD SMEAR EXAMINATIONS: CAN SYSMEX BLOOD CELL ANALYSERS HELP?

J Joubert[1] *MBChB, MMed (Haemat)* | **R Weyers**[2] *MBChB, MMed (Haemat)* | **J Raubenheimer**[3] *PhD*

[1]*National Health Laboratory Service (NHLS), Kimberley, and Department of Haematology and Cell Biology, Faculty of Health Sciences, University of the Free State (UFS), Bloemfontein, South Africa*
[2]*NHLS Pelonomi Hospital, and Department of Haematology and Cell Biology, Faculty of Health Sciences, UFS, Bloemfontein, South Africa*
[3]*Department of Biostatistics, Faculty of Health Sciences, UFS, Bloemfontein, South Africa*
Corresponding author: ***Dr. Jaco Joubert*** *| Tel: +27 51 405 3043 | Fax: +27 51 405 2923 | Email: jaco.joubert@nhls.ac.za*

ABSTRACT

Background The microscopic assessment of a peripheral blood smear is an essential diagnostic tool. Many haematology laboratories currently assess smears microscopically for every full blood count request, many of which may however be assessed unnecessarily – an important consideration in resource-constrained settings. Modern blood cell analysers are increasing in sophistication and can flag abnormal specimens that may require microscopy.

Objectives To evaluate the flagging efficiency of the Sysmex haematology analysers and to determine whether this potentially labour-saving technology could assist in safely reducing the number of unnecessary microscopic blood smear assessments.

Methods A total of 427 full blood count specimens collected consecutively over a 24-hour period at NHLS Pelonomi and NHLS Kimberley, were evaluated microscopically and compared with the instruments' abilities to flag potential morphological abnormalities.

Results The Sysmex blood cell analysers flagged 63.7% of specimens as "positive" and 36.3% as "negative". After microscopy, false positive flags were found to constitute 18.5% and false negative flags 5.4% of the total number of smears reviewed, giving a total of 23.9% incorrect assessments. No false negative flag was clinically critical.

Conclusion False negative results occurring with the Sysmex instruments' flagging systems in our settings are relevant, although not critical. The potential time and monetary savings of a flagging-based smear review policy may weigh heavier than occasional false negatives. In the African milieu, where laboratories are faced with the challenges posed by staff- and other shortages, relying on instrumentation flagging to guide smear review policy should be considered.

KEYWORDS

Sysmex, flagging, smear review

INTRODUCTION

Modern blood cell analysers are designed to detect unusual characteristics of blood specimens that may interfere with the blood cell counts or require a microscopic assessment of a peripheral blood smear. Both quantitative abnormalities of the blood cell counts and qualitative abnormalities in the specimen (for example, abnormal cell types such as blasts) can be detected, a process referred to as "flagging". The noted abnormalities are called "flags"[1,2]. The Sysmex XT-series blood cell analysers (Sysmex Corporation, Kobe, Japan) used in our laboratories employ this principle of flagging abnormal parameters on full blood count specimens[3].

Kimberley Hospital and Pelonomi Hospital in Bloemfontein are referral facilities with all secondary level and many tertiary level services available. Most specialities are represented and both hospitals are involved in training of clinical and laboratory staff. Kimberley Hospital is a 656-bed hospital in the city of Kimberley, serving the geographically vast Northern Cape province of South Africa. Pelonomi Hospital is a ±720 bed hospital in the city of Bloemfontein, and is the main secondary hospital serving the Free State province. The Kimberley National Health Laboratory Service (NHLS) laboratory processes between 90 000 and 110 000 clinical pathology tests per month,[4] which include 5 600 full blood counts (FBCs) per month, with ±3 000 requests for differential white cell counts (WCCs) on which peripheral blood smear morphology is then also performed[5]. NHLS Pelonomi processes 108 000 clinical pathology tests on average per month. Approximately 4 000 FBCs are received here monthly, with an almost equal number of requests for differential WCCs.

The microscopic assessment of a peripheral blood smear is an essential diagnostic tool in the haematology laboratory and can complement the numerical values on an FBC to provide a definite diagnosis or to guide further investigation of a patient. Peripheral smear assessments can also be used as a quality control tool for the FBC and provide excellent material for training purposes[6]. Therefore, it is currently the practice in our laboratories to microscopically assess a peripheral blood smear for every differential WCC request received by the laboratory. Unfortunately this process is very time consuming and requires the skill of an experienced morphologist, such as a haematopathologist or haematology laboratory technologist[2,6].

Internationally, laboratory budgets are being cut with a resultant reduction in staff numbers[7]. Consequently, the workload of laboratory personnel is increased, while they are constantly under pressure to deliver accurate test results to clinicians as soon as possible. When a patient's FBC is normal, a microscopic evaluation of the blood specimen may also not add any

value to the information provided by the blood cell analyser. Our objective was to evaluate the flagging efficiency of our Sysmex haematology analysers and to determine whether this potentially labour-saving technology could assist us in reducing the number of microscopic blood smear assessments undertaken in our laboratories, thereby reducing the workload of our haematology technologists.

MATERIALS AND METHODS

Study samples

In this study we included a total of 427 anticoagulated whole blood specimens collected in dipotassium ethylenediamine tetra-acetate (K_2EDTA) tubes and submitted for routine FBC analysis to NHLS Pelonomi and NHLS Kimberley. The FBCs were collected consecutively over a 24-hour period. Both laboratories received specimens from hospital in- and outpatient departments as well as peripheral clinics in the district. The study was approved by the Ethics Committee of the Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa (ECUFS number 212/2012).

Full blood counts

Automated FBCs and differential WCCs were performed with the Sysmex XT-series blood cell analysers (Sysmex Corporation, Kobe, Japan). NHLS Pelonomi routinely uses the Sysmex XT-1800i blood cell analyser and NHLS Kimberley the Sysmex XT-2000i instrument. Both these analysers employ the principle of flagging quantitative and qualitative abnormalities on FBC specimens. The instrument printouts of the FBCs are clearly labelled as "negative" (no quantitative or qualitative abnormalities present) or "positive" (abnormalities/flags detected). The difference between the XT-1800i and XT-2000i models is that the XT-2000i is capable of detecting immature granulocytes, generating automated reticulocyte counts and providing fluorescent optical platelet counts through the reticulocyte channel. None of these additional functions were utilised in this study. All other technologies employed by these two instruments (including blood cell enumeration and flagging) are identical[3].

The FBCs were performed according to the standard operating procedure (SOP) of each laboratory. At NHLS Pelonomi, three levels of internal quality control samples (low, normal, high) are run twice daily. At NHLS Kimberley controls are run every 8 hours with the shift change, alternating between manual (open) and closed modes[8].

Manual slide review

A peripheral blood film was evaluated microscopically for each specimen included in this study. All blood films were smeared manually and then Wright-stained using the Hematek automated slide stainer (Siemens Healthcare Diagnostics Inc., Tarrytown NY).

First, each laboratory's FBCs and peripheral smears were evaluated by one of the first two authors, being that particular laboratory's resident haematologist (primary assessment). The two haematologists noted their findings on the microscopic assessment of each peripheral smear and also indicated by means of "yes" or "no" whether the findings were clinically relevant, or added any additional information beyond that provided by the blood cell analyser on the FBC printout. Additional information collected on each specimen included the age and gender of the patient and from which clinical department or ward the blood specimen was submitted.

Thereafter a cross-over took place where the FBCs from one laboratory were evaluated by the haematologist from the other laboratory (secondary assessment). The haematologists were blinded to each other's primary evaluations. The recording of data during the secondary assessment followed the same format as that of the primary assessment.

After completion of the secondary assessment the two haematologists compared and discussed their results to formulate one consolidated opinion of whether a blood smear's findings were clinically relevant or not. These consolidated assessments were then used to evaluate the flagging efficiency of the Sysmex blood cell analysers. In one case they failed to come to an agreement on whether the microscopic findings were clinically relevant or not. Here the expert opinion of an independent third party (a randomly selected qualified haematologist from the referral laboratory, NHLS Universitas) was sought to arbitrate the case.

When the blood cell analyser flagged a specimen as "positive" and the blood smear analysis showed clinically relevant findings, the specimen was graded as "true positive". If the analyser flagged a specimen as "positive" and the smear showed no clinically relevant findings, the specimen was graded as "false positive". A "true negative" grading was awarded to "negative" flagged specimens with no clinically relevant findings on the smear. A specimen was graded as "false negative" when flagged "negative" by the analyser, but clinically relevant findings were observed on the smear.

Statistical analysis

Statistical analysis was done using the SAS/STAT software system (SAS, 2010) and the DAG-Stat program provided by Mackinnon[9,10]. Cross-tabulation with chi-square analysis was performed, and the sensitivity and specificity of the results provided by the Sysmex instruments were calculated. Inter-rater agreement between the assessments of the two haematologists was determined by calculating Cohen's Kappa.

RESULTS

Study population

Of the 427 FBCs included in the study, 212 (49.6%) were from NHLS Kimberley and 215 (50.3%) from NHLS Pelonomi. The patients' ages ranged from 0 days to 96 years. The majority of the FBCs were from adult patients (83.8%), while paediatric patients (age 1 month to 12 years) comprised 10.7% and neonatal patients (≤30 days old) 5.5% of the study population. Of all the patients, 61.7% were female.

At NHLS Kimberley, the greatest number of FBCs was submitted from outpatients (63.7%), while inpatient departments contributed 36.3% of FBCs. In contrast to this, NHLS Pelonomi received more FBCs from inpatient departments (55.8%) than from outpatients (44.2%). Table 1 summarises the specimen profiles of in- and outpatients at the two laboratories.

Flagging and smear review

As shown in Table 2, the Sysmex blood cell analysers flagged 272 (63.7%) of the FBCs as "positive" and 155 (36.3%) as "negative". After microscopic assessment of all 427 peripheral smears, it was found that the false positive and false negative flags respectively represented 18.5% and 5.4% of the smears reviewed, giving a total of 23.9% incorrect assessments. The sensitivity of the Sysmex flagging system to detect clinically significant findings (both quantitative and qualitative) was 89.4%

Table 1: Comparison of specimens for full blood counts obtained from in- and outpatients in different wards/settings analysed at NHLS Kimberley and NHLS Pelonomi.

Ward/setting	Laboratory			
	NHLS Kimberley (n=212)		NHLS Pelonomi (n=215)	
	n	%	n	%
Inpatients				
Intensive care units	4	1.9	25	11.6
Paediatric ward	11	5.2	15	7.0
Neonatal unit	11	5.2	8	3.7
Medical ward	6	2.8	10	4.7
Surgical ward	4	1.9	10	4.7
Oncology	38	17.9	0	0
Obstetrics & Gynaecology	0	0	37	17.2
Orthopaedics	0	0	11	5.1
Other	3	1.4	4	1.9
Total	77	36.3	120	55.8
Outpatients				
Peripheral clinic	126	59.4	43	20.0
Hospital outpatient departments	5	2.4	11	5.1
Casualty, emergency, trauma	3	1.4	40	18.6
Unknown	1	0.5	1	0.5
Total	135	63.7	95	44.2

(95% confidence interval (CI) 84.5-93.1%) and the specificity 62.6% (95% CI 55.7-69.1%). The negative predictive value of the instruments' flagging system was higher than the positive predictive value at 85.2% (95% CI 78.6-90.4%) and 71.0% (95% CI 65.2-76.3%), respectively. The false positive rate was 37.4% (95% CI 30.9-44.4%) and the false negative rate 10.7% (95% CI 6.9-15.6%). The efficiency (correct classification rate) of the Sysmex flagging system was 76.1% (95% CI 71.8-80.1%). The positive agreement between the analyser flags and the morphological assessment was 79.1% (95% CI 75.1-83.1%) and the negative agreement 72.1% (95% CI 66.9-77.3%).

Inter-rater agreement

Cohen's Kappa indicated fair agreement between the two observers (0.26, 95% CI 0.20-0.31) after the blind, independent assessment of the FBCs and peripheral smears. With 187 (43.8%) out of the 427 peripheral smear assessments, the two observers differed in their initial opinion on the clinical relevance of morphological findings. After review and consolidation of the initial results, the final consolidated assessment in 47.1% of cases agreed with observer A's initial assessment and in 52.9% of cases with the initial assessment of observer B.

DISCUSSION

In our laboratories, peripheral blood smear morphology is currently manually reviewed for every differential WCC request,[11] resulting in approximately 50-100 peripheral blood smears being reviewed per technologist per day at NHLS Kimberley and 100-140 at NHLS Pelonomi (depending on staffing), in addition to other routine haematology work. This approach undoubtedly acts as an additional quality control measure and reduces the risk of serious pathology being missed. However, with the availability of sophisticated modern instruments, the question invariably arises whether this is actually necessary. It has also been claimed that the presence of atypical cells among leukocytes will only be recognised with reasonable certainty (greater than 95%) in a 100-cell differential if such cells constitute >5% of the leukocyte count, casting doubt on the impact of manual smear review in general[12].

In 2005, the International Consensus Group for Hematology Review published guidelines in the form of "rules", intended to guide manual smear reviews following automated FBC and differential WCC results[13]. These guidelines have been validated and optimised,[14] although it is recommended that each labora-

Table 2: Sysmex flagging compared to haematologist microscopic assessment of peripheral blood smears.

Haematologist assessment	Sysmex flagging					
	Positive		Negative		Total	
	n	%	n	%	n	%
Clinically relevant	193	45.2	23	5.4*	216	50.6
Not clinically relevant	79	18.5#	132	30.9	211	49.4
Total	**272**	**63.7**	**155**	**36.3**	**427**	**100**

**False negative; #False positive*

tory has its own criteria for smear review[13-15]. When these rules are applied, various rates of smear review emerge, depending on the number of hospital beds,[2] patient composition (inpatient versus outpatient),[15] service level (primary versus tertiary care) [15] and choice of analyser[14,15]. A 2006 College of American Pathologists (CAP) Q-probe study found manual smear review rates of 16.2% among the 263 institutions surveyed, with the median institution reporting manual reviews on 26.7% of specimens[2]. Using the International Consensus Group guidelines, a South Korean group reported manual smear review rates of 28.6%, using the Sysmex XE 2100 (Sysmex, Kobe, Japan), and rates of 22.8% and 20.2% for the UnicelDxH 800 (Beckman Coulter Inc., Fullerton, CA) and ADVIA2120i (Siemens Diagnostics, Tarrytown, NY) analysers, respectively[15]. Pratumvinit *et al*.[14] reported slide review rates of 29.3% using these criteria and rates of 22.4% using their own locally developed laboratory criteria, on results obtained from the SysmexXE-5000 (Sysmex, Kobe, Japan) and CoulterLH750 analysers (Beckman Coulter Inc., Brea, CA)[14]. However, their study excluded haematology clinic specimens, which might have influenced review rates. Using review criteria adopted from their own laboratory's SOP, Leers *et al*.[16] 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)[16].

From the aforementioned publications it is clear that both NHLS Kimberley and NHLS Pelonomi are probably reviewing large numbers of smears unnecessarily. Moreover, although the 2006 CAP Q-probe reported that in 35.7% of the peripheral smears reviewed manually, participants claimed to have learned additional information beyond what was available on automated haematology analyser printouts alone,[2] it has been argued that in many clinical settings the proportion of specimens legitimately requiring a smear review may actually be as low as 10-15% or less[17].

As shown in Table 1, the bulk (63.7%; 135/212) of NHLS Kimberley's FBC workload is received from outpatients with the majority of these specimens (93.3%; 126/135), coming from outlying peripheral clinics, reaching the laboratory only at night. A high proportion of these smears (57.9%; 73/126) have storage-induced and even heat artefacts, probably as a consequence of the distances the specimens have to travel in the January, Northern Cape summer heat[18]. This has implications for peripheral blood smear review. Experienced morphologists may not always be available at night to assess these morphological changes as artefacts. Storage may also lead to artefactual changes in automated counts[18] and needs to be considered in this setting. Although relying solely on automated flagging in these cases may not be ideal, as this may potentially result in an increase in false positive results, relying on the morphological evaluations of an inexperienced morphologist may be even more misleading, with factitious nucleated red cells, neutropenia and lymphocytosis being typically recorded by inexperienced operators,[18] all of which may have clinical implications.

The Sysmex XT2000i has been shown to be able to perform a clinically valid FBC and differential WCC on K_2-EDTA anticoagulated blood 24 hours after collection when the specimen is stored at room temperature, although the mean platelet volume (MPV), mean cell haemoglobin concentration (MCHC), haematocrit (HCT), mean cell volume (MCV), and red cell distribution width (RDW) should be excluded from the report[19]. It has been concluded, however, that because of the greater variability of differential counts in specimens from HIV-infected patients, any results for an automated differential WCC on a specimen more than 24 hours old is best not reported[19]. This factor must be considered, as specimens from HIV-positive patients account for a large portion of our work. In contrast, as shown in Table 1, NHLS Pelonomi received more FBCs from inpatient departments (55.8%; 120/215) than from outpatients (44.2%; 95/215), of which only 45.3% (43/95) were from peripheral clinics. Storage artefacts, although still present, were encountered less in these smears and are probably not as big a factor in Bloemfontein as in the Northern Cape.

As shown in Table 1, paediatric and neonatal patients constituted the minority of specimens received in this study, comprising 10.7% and 5.5% of specimens, respectively. Although this accurately reflects practice at our respective institutions, tertiary centres may have more paediatric and neonatal patients, requiring alternative smear review policies, since these patients have different normal haematological reference ranges and peripheral blood smear morphology from adults,[18] making the blanket application of automated flagging systems in this context inap-

Table 3: Comparison of the Sysmex XT1800i and XT2000i morphological abnormality flagging efficiency with a selection of similar studies and performance evaluations.

Parameter (%)	Study			
	Current study	Leers *et al*.[16]	Langford *et al*.[21]	Kim *et al*.[15] (XE2100)
Sensitivity	89.4 (95% CI 84.5-93.1)	NR	95.8	60.3
Specificity	62.6 (95% CI 55.7-69.1)	NR	90.9	77.1
False negative rate	10.7 (95% CI 6.9-15.6)	NR	NR	9.7
False positive rate	37.4 (95% CI 30.9-44.4)	NR	NR	17.3
Positive predictive value	71.0 (95% CI 65.2-76.3)	NR	88.5	45.8
Negative predictive value	85.2 (95% CI 78.6-90.4)	NR	96.8	NR
Positive agreement	79.1 (95% CI 75.1-83.1)	47.4	NR	NR
Negative agreement	72.1 (95% CI 66.9-77.3)	93.6	NR	NR
Efficiency	76.1 (95% CI 71.8-80.1)	NR	92.9	NR
Cohen's Kappa	0.52* (95% CI 0.44-0.60)	0.438*	NR	NR

**Moderate agreement; NR = not reported*

propriate. Manual slide review rates may range from 38% to as high as 87% in paediatric patients,[20] and a similar scenario can be anticipated in haematology or oncology patients,[16] which constituted 49.4% (38/77) of NHLS Kimberley's inpatient specimens. Since our paediatric specimens were limited, direct comparison of our findings with a detailed performance evaluation of the Sysmex XT-2000i conducted at the Saint Louis Children's Hospital in St. Louis, Missouri, USA in 2002,[21] should be done with reserve. When developing a local smear review policy for settings such as our own, each laboratory should therefore take into consideration the probable number of old specimens reaching it from the periphery, the characteristics of the patients it serves (such as the number of paediatric and haematology patients) and specimens form HIV-positive patients.

The International Consensus Group Criteria consist of over 40 rules,[13] which although very useful, are intricate and possibly not ideal for use in laboratories that may not always have enough experienced and qualified personnel to apply them correctly. It is conceivable that serious pathology may slip through if these (actually quite liberal) criteria are not applied correctly. Furthermore, these rules do not take local population differences into account. Slide review is called for whenever a WCC is below 4.0 x 10^9/L,[13] although it is known that a value as low as 2.8 x 10^9/L may in fact be normal in an adult black male[22]. In order not to miss serious underlying pathology – the inherent danger of not reviewing a smear for every differential WCC – a simpler and perhaps more conservative method of screening is required, which would still decrease unnecessary slide reviews without compromising safety. It is in this context that the Sysmex instruments' flagging efficiency is evaluated here, essentially to answer the following question: Can it be used as a safe substitute for the complex Consensus Group Criteria to quickly guide a technologist in a busy, understaffed laboratory on whether a smear review is required or not? If so, smear review rates at NHLS Kimberley and NHLS Pelonomi would decrease from 100% to 63.7%, based on the positive versus negative flagging of our specimens.

Table 3 compares the statistical analysis of our instruments' flagging capabilities versus manual slide review with a selection of other studies. Although not all studies reported data on every parameter, it is clear that variation in reported performance exists. This may be due to differences in the study populations[21] and numbers of slides reviewed[15,16]. Although the specificity of our instruments' flagging systems are lower than reported in similar studies, their relatively good sensitivity may make them good screening tools, which would be their envisioned application in our laboratories. Moreover, since a large proportion of our specimens are from outpatients, higher sensitivity for screening may be more advantageous than higher specificity, which would be more useful in an inpatient setting where efficient patient care is more important[15].

In the screening setting, the number of false negatives would be the primary concern. Only 23 (5.4%) false negative smears were found on review, with a false negative rate of 10.7%, which is in line with Korean data for the XE 2100[15]. In these 23 false negative cases, manual slide reviews contributed clinically relevant information in all cases, but none were critical. The assessment as "clinically relevant" was unanimous in only eight of these 23 cases. In no instance would manual slide review have changed diagnosis, patient management or prognosis. This is in contrast to the findings of Leers *et al.*[16] where cases of hairy cell leukaemia and malignant lymphoma were not flagged by the XT-2000i, although serious pathology of this nature was not encountered in any of our 427 cases.

Although false negatives are particularly important in this setting, the number of false positive cases is not without consequences for laboratory efficiency. False positives constituted 18.5% (79/427) of all specimens, with a false positive rate of 37.4%, more than double the rate of 17.3% found by Kim *et al.* on the XE2100[15]. This means that more than a third of specimens without clinically relevant abnormalities on the peripheral blood smear would be flagged by the instrument. The positive predictive value of 71.0%, although much higher than the value of 45.8% reported for the XE2100,[15] implies that technologists may be manually reviewing slides unnecessarily in up to 29% of positively flagged cases. The efficiency (correct classification rate) of 76.1%, which is lower than the efficiency of 92.9% reported by Langford *et al.*,[21] implies that almost one out of every four specimens will not be flagged correctly, most of these being specimens flagged as abnormal but without clinically relevant morphological abnormalities on the blood smear. Although positive agreement between the flagging system and manual smear review was better than found by Leers *et al.*,[16] negative agreement was not. This finding should be kept in mind when developing a local screening policy, based on the exclusion of negatively flagged specimens.

If a laboratory decides to only review all positively flagged specimens instead of reviewing each and every specimen manually (or instead of utilising the International Consensus Group Criteria), it must, in addition to the other considerations already mentioned, weigh up the potentially unnecessary review of positively flagged normal cases, with the time-saving effects of omitting slide reviews on truly negative specimens. In our study 18.5% (false positives; 79/427) of specimens can be deemed as triggering unnecessary smear review. However, in 30.9% of cases (true negatives; 132/427) unnecessary smear review would have been avoided.

It is our estimate that an experienced technologist spends an average of approximately two minutes assessing a routine peripheral blood smear. It is probably a conservative estimate as manual smear review times of over four minutes per smear have been reported[23]. Since every smear is currently reviewed at our laboratories, this means that 854 minutes (2 minutes/smear x 427 smears) or 14 hours and 14 minutes, were spent by technologists on reviewing peripheral blood smears during the roughly 24-hour period of the study; 424 minutes (2 minutes/smear x 212 smears) or 7 hours and 4 minutes at NHLS Kimberley and 430 minutes (2 minutes/smear x 215 smears) or 7 hours and 10 minutes at NHLS Pelonomi. This includes 132 smears reviewed unnecessarily which means a total of approximately 264 minutes (132 true negative smears; 4 hours and 24 minutes) of technologist bench-time during this 24-hour period were essentially wasted at these two laboratories. In our local context, if a technologist's 40-hour work week is taken as one Full Time Equivalent (FTE), then a total of 96 360 minutes (264 minutes/day x 365 days) or 1 606 hours per year, which translates to 40.15 FTEs per year, would potentially be freed up, if our two laboratories were to abandon the current policy of reviewing each and every peripheral blood smear. Ultimately, each laboratory would be able to allocate roughly 20.08 FTEs of tech-

nologist bench-time per year to other tasks, greatly reducing the strain on understaffed laboratories, improving work-flow, result quality and efficiency.

The price for this improvement would be the impact of false negatives, which we believe is acceptable in our specific context. This can be managed by still preparing and staining a peripheral blood smear for every specimen, should the need for manual smear review subsequently arise. Should a clinician for instance request review at a later stage, the smear can still be retrieved. The impact of this on workload would probably be negligible, as a 2006 Q-probes study found that only 3.7% of smear reviews were performed following a request by a physician[2].

A Cohen's Kappa score of 0.26 (95% CI 0.20-0.31) indicates fair agreement between the two observers. Although literature on agreement of morphological assessments among histopathologists in many different settings is abundant, a paucity of data exists on inter-observer variability among haematologists, when reviewing peripheral blood smears. A 1985 study[24] found good correlation between technologists for neutrophil, normal lymphocyte and eosinophil counts, but more variability was noted in the estimation of band cells and abnormal lymphocytes and monocytes. The sensitivity of manual smear review for clinically important conditions in this study ranged from 100% to 34%, depending on the abnormality[24]. Haematologists' semi-quantitative assessment of anisocytosis has been shown to display inter-observer correlation as well as correlation with the quantitative RDW[25]. The same authors, however, still recommend that the RDW be used to quantify erythrocyte size variability, due to its extreme precision[25]. Significant inter-observer concordance (Kappa = 0.909) has also been shown among reviewers of bone marrow morphology in patients with previously diagnosed low-risk myelodysplastic syndrome[26]. Inter-observer agreement for the blast cell percentage in the peripheral blood of patients with myelodysplastic syndromes has also recently been shown to be very good among cytomorphologists,[27] although no such cases were encountered in our study. Agreement among haematologists when evaluating bone marrow trephine biopsies for essential thrombocythaemia was better for measures of general morphologic patterns than individual cellular features, and poor agreement was found when various parameters were to be synthesised into a diagnostic category[28]. The investigation of agreement between our two observers did not extend to diagnostic category or individual morphological features, but simply focused on the clinical relevance of a particular smear, since this would be the compelling issue in deciding on a screening policy.

Industry seems to also be taking cognisance of the need for refined flagging systems. Sysmex has developed new software for its XE-5000 with improved algorithms for flagging blast cells, abnormal lymphocytes or lymphoblasts, and atypical lymphocytes, which resulted in more specific flagging with less smears ultimately requiring manual review[7]. However, the utility of the blast flag reported by the XE-5000 as a sufficient criterion for performing a microscopic review was subsequently questioned[29].

Several automated digital imaging systems have also been introduced in recent years in an attempt to standardise smear review, increase laboratory efficiency and improve workflow and accuracy[30]. Three CellaVision instruments – the Diffmaster Octavia,[23] the CellaVision DM96,[23] and the CellaVision DM1200 (30),[30] – as well as the Nextslide Digital Review Network[30] have all been introduced for this purpose. Although these and other new, similar technologies such as the CellaVision Image Capture System (CICS) may be of great benefit to smaller, remote peripheral laboratories,[31] costs and availability may limit their widespread utilisation in medium-sized laboratories throughout South Africa at this time.

LIMITATIONS

Our study has several important limitations, most brought about by our local contexts. The selection of specimens did not include cases with serious pathology such as leukaemia or lymphoma. It also did not include cases where diagnosis relies heavily on morphology such as autoimmune haemolytic anaemia, fragmentation haemolysis or acute promyelocytic leukaemia. The effect of specimen age and heat artefacts on flagging and morphology also limits the broader applicability of our results. With the high prevalence of HIV in the South African context, our data may not be globally applicable and vice versa.

FUTURE RECOMMENDATIONS

In future, we would recommend that only positively flagged specimens' smears be reviewed. We do, however, recommend that these smears be checked thoroughly. Simple scanning at low magnification is not recommended. If a specimen is flagged, it is probably prudent to review the smear for any morphological abnormality, and not just for those abnormalities suggested by the flag. We also recommend that all flags be checked and not just certain flags, such as blast flags. Smears should probably still be prepared for every differential WCC request and retained according to the current local protocol for smear retention, which at NHLS Kimberley and NHLS Pelonomi is roughly one week. Better flagging software, especially to deal with the many false positives encountered in this and other similar studies, needs to be developed by industry. In light of staff shortages, automated digital imaging systems may offer a solution in future, especially for remote laboratories. Every laboratory should draft its own smear review and flagging utilisation policy, depending on the inpatient and outpatient profile, number of old specimens, paediatric patients and staff complement. Such a policy can be refined even further by encouraging qualified medical technologists to not only use the flagging system as a guide but also to evaluate the FBC in conjunction with previous counts and the available clinical details, in order to make a decision on the need for smear review. A qualified medical technologist should have the necessary knowledge, and if applied correctly, could assist in further improving the correct classification rate. Moreover, some specimens require urgent evaluation and these could also be identified by a qualified technologist.

One must not forget the philosophy behind peripheral blood smear review, so eloquently summarised by Briggs *et al*.[7]: "A smear is reviewed to provide information additional to or missing from the analyser report or to confirm results provided by the analyser. The challenge is to reduce the number of blood films examined without missing important diagnostic information. The objective for the laboratory is to reduce the number of specimens needing further action as far as possible without endangering patients by reporting false or misleading results, especially false-negative results. Any analyser that decreases the number of false-positive flags without losing sensitivity will increase laboratory efficiency."

CONCLUSIONS

False negative results do occur when using the Sysmex instruments' flagging systems, which are relevant but not critical. The potential time and monetary savings in a context of perpetual shortages may weigh heavier than occasional false negatives. The risk-benefit ratio appears in favour of using the flagging system. In the African milieu, where laboratories are faced with the numerous challenges posed by HIV, vast distances, staff shortages, and time and monetary constraints, relying on instrumentation flagging to guide smear review policy should be considered. Such an approach is, however, probably not appropriate for paediatric and neonatal specimens. Every laboratory should draft its own smear review and flagging utilisation policy depending on local circumstances. The potential liberation of many FTEs should be considered when laboratory planning and staff number management is done.

ACKNOWLEDGEMENTS

The authors thank the management and technologists of NHLS Kimberley and NHLS Pelonomi, as well as the management of Kimberley and Pelonomi Hospitals, for their assistance with this study. We also acknowledge Dr. Daleen Struwig, medical writer, Faculty of Health Sciences, UFS, for assistance with the technical and editorial preparation of the manuscript.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Bain BJ. Performing a blood count. In: *Blood Cells: A Practical Guide.* 4th ed. Malden, MA: Blackwell Publishing Inc., 2006: 20-60.
2. Novis DA, Walsh M, Wilkinson D, St. Louis M, Ben-Ezra J. Laboratory productivity and the rate of manual peripheral blood smear review: a College of American Pathologists Q-probes study of 95 141 complete blood count determinations performed in 263 institutions. *Arch Pathol Lab Med* 2006; 130: 596-601.
3. Sysmex Corporation. XT-2000i and XT-1800i automated haematology analyzers: mid-volume laboratory solution. Document number MKT-10-1136 05/2012 1M. Mundelein: Sysmex America Inc., 2012.
4. Brits D. 2012/13 volumes: NHLS Kimberley. 2013.
5. Joubert J. NHLS Kimberley Haematopathology Annual Report. 2013.
6. Gulati GL, Alomari M, Kocher W, Schwarting R. Criteria of blood smear review. *Lab Med* 2002; 33: 374-347.
7. Briggs CJ, Linssen J, Longair I, Machin SJ. 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* 2011; 136: 309-316.
8. Winter JD. NHLS Kimberley Standard Operating Procedure KHAEM – SOP-0017. Running of controls on the Sysmex XT2000i. 2007.
9. Mackinnon A. A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. *Comput Biol Med* 2000; 30: 127-134.
10. SAS Institute. SAS/STAT software, Version [9.3]. Cary, NC: SAS Institute Inc., 2010.
11. Khoza P. NHLS Standard Operating Procedure HAE0028: procedure for making slides and performing a differential count. 2013.
12. Rümke CL. Statistical reflections on finding atypical cells. *Blood Cells* 1985; 11: 141-144.
13. Barnes PW, McFadden SL, Machin SJ, Simson E. The International Consensus Group for Hematology Review: suggested criteria for action following automated CBC and WBC differential analysis. *Lab Hematol* 2005; 11: 83-90.
14. Pratumvinit B, Wongkrajang P, Reesukumal K, Klinbua C, Niamjoy P. Validation and optimization of criteria for manual smear review following automated blood cell analysis in a large university hospital. *Arch Pathol Lab Med* 2013; 137: 408-414.
15. Kim SJ, Kim Y, Shin S, Song J, Choi JR. Comparison study of the rates of manual peripheral blood smear review from three automated hematology analyzers, Unicel DxH 800, ADVIA 2120i, and XE 2100, using International Consensus Group Guidelines. *Arch Pathol Lab Med* 2012; 136: 1408-1413.
16. Leers MPG, Goertz H, Feller A, Hoffmann JJML. Performance evaluation of the Abbott CELL-DYN Ruby and the Sysmex XT-2000i haematology analysers. *Int J Lab Hematol* 2011; 33: 19-29.
17. Bain BJ. Diagnosis from the blood smear. *New Engl J Med* 2005; 353: 498-507.
18. Bain BJ. Morphology of blood cells. In: Bain BJ. *Blood Cells: A Practical Guide.* 4th ed. Malden, MA: Blackwell Publishing Inc., 2006: 61-174.
19. Hill VL, Simpson VZ, Higgins JM, et al. Evaluation of the performance of the Sysmex XT-2000i hematology analyzer with whole blood specimens stored at room temperature . *Lab Med* 2009; 40: 709-718.
20. Boer K, Deufel T. Workflow restrictions on hematology analyzers XE-2100 and XS-800 when assaying pediatric samples with limited volume. *Clin Chem Lab Med* 2009; 47: 607-611.
21. Langford K, Luchtman-Jones L, Miller R, Walck D. Performance evaluation of the Sysmex XT-2000i automated hematology analyzer. *Lab Hematol* 2003; 9: 29-37.
22. Bain BJ. Normal ranges. In: Bain BJ. *Blood Cells: A practical Guide.* 4th ed. Malden, MA: Blackwell Publishing Inc., 2009: 198–216.
23. Ceelie H, Dinkelaar RB, van Gelder W. Examination of peripheral blood films using automated microscopy; evaluation of Diffmaster Octavia and Cellavision DM96. *J Clin Pathol* 2007; 60: 72-79.
24. Koepke JA, Dotson MA, Shifman MA. A critical evaluation of the manual/visual differential leukocyte counting method. *Blood Cells* 1985; 11: 173-186.
25. Simel DL, DeLong ER, Feussner JR, Weinberg JB, Crawford J. Erythrocyte anisocytosis: visual inspection of blood films vs automated analysis of red blood cell distribution width. *Arch Int Med* 1988; 148: 822-824.
26. Howe RB, Porwitt-MacDonald A, Wanat R, Tehranchi R, Hellström-Lindberg E. The WHO classification of MDS does make a difference. *Blood* 2004; 103: 3265-3270.
27. Senent L, Arenillas L, Luño E, Ruiz JC, Sanz G, Florensa L. Reproducibility of the World Health Organization 2008 criteria for myelodysplastic syndromes. *Haematologica* 2013; 98: 568-575.
28. Wilkens BS, Erber WN, Bareford D, et al. Bone marrow pathology in essential thrombocythemia: interobserver reliability and utility for identifying disease subtypes. *Blood* 2008; 111: 60-70.
29. Eilertsen H, Vøllestad NK, Hagve T-A. The usefulness of blast flags on the Sysmex XE-5000 is questionable. *Am J Clin Pathol* 2013; 139: 633-640.
30. Yu H, Ok CY, Hesse A, et al. Evaluation of an automated digital imaging system, Nextslide Digital Review Network, for examination of peripheral blood smears. *Arch Pathol Lab Med* 2012; 136: 660-667.
31. Smits SM, Leyte A. Clinical performance evaluation of the CellaVision Image Capture System in the white blood cell differential on peripheral blood smears. *J Clin Pathol* 2013; 67: 168-172.