A Real-World Evaluation of a Semi-Supervised Artificial Intelligence Model (ASUS Blade) with the Cellavision DM9600

Bingwen Eugene FAN[1,2], David Tao Yi CHEN[3], Yi Xiong ONG[4], Kian Guan Eric LIM[4], Wei Yong Kevin Wong[4], Joshua ONG[3], Shu-Yu HSU[5], Pik Wan Erica CHIANG[2], Siti Thuraiya Binte Abdul LATIFF[2], Shu Ping LIM[4], Moh Sim WONG[2], Hemalatha SHANMUGAM[1], Stefan WINKLER[3], Ponnudurai KUPERAN[1,2]

[1]Department of Haematology, Tan Tock Seng Hospital, Singapore [2]Department of Laboratory Medicine, Khoo Teck Puat Hospital, Singapore
[3]ASUS Global Pte Ltd, Singapore [4]Department of Laboratory Medicine, Tan Tock Seng Hospital [5]ASUSTek Computer Inc., Taiwan

Introduction

Manual light microscopy (MM) of peripheral blood film (PBF) is tedious and time-consuming. A deep-learning Artificial Intelligence (AI) model, Blade, has been utilized to identify white blood cells (WBC) through a convolutional neural network (CNN) trained by semi-supervised deep learning technique using 185,412 peripheral blood film cells. We compared the performance of Blade with a commercial model, the Cellavision DM9600 (Cellavision) in a real world laboratory setting.

Objectives

To evaluate the performance of Blade in identifying common cell-types with reference from MM (current practice). To compare the performance of Blade in detecting abnormal differential count finding with that of Cellavision.

Methods

168 films were randomly selected from the routine laboratory bench to obtain a WBC differential count by evaluating 200 WBC per slide, using MM, Cellavision and Blade. 6 lab-certified medical technologists reviewed these slides on MM, followed by assessment by Cellavision, and then by Blade. The study assessed 13 categories: Neutrophil, Lymphocyte, Large Granular Lymphocyte, Reactive Lymphocyte, Monocyte, Eosinophil, Basophil, Metamyelocyte, Myelocyte, Blast, Smudge Cell, Artifact, Giant Platelet. Correlation (r, r^2) of differential count was plotted for both test methods against MM which was considered as the gold standard (Figure 1 and 2). Agreement between AI test methods and the reference method (MM) is calculated for various abnormalities (Figure 3). The performance (accuracy, sensitivity, specificity, and precision) of manual microscopy and AI based test methods (MM, Cellavision, Blade) for the detection of abnormal slides was evaluated.

Results

Blade's correlation to MM for correct identification of common cell-types is higher than that of Cellavision, ranging between 0.72 to 0.98 for Blade (Figure 1), and 0.62 to 0.97 for Cellavision (Figure 2). Agreement between test method and MM for characteristic differential count findings (abnormalities) show Blade with a higher agreement in detecting Granulocytosis, Lymphocytosis, Lymphopenia, Myelocytes, Metamyelocytes and Blasts. Conversely, the agreement is higher in Cellavision for Eosinophilia and Monocytosis (Figure 3). The performance of Blade in detecting abnormalities is 87.5%, 92.3%, 70.3%, 91.7% (accuracy, sensitivity, specificity, and precision). This is higher than that of Cellavision's DM9600, being 83.9%, 90.8%, 59.5%, 88.8% respectively. Statistical assessment proved non-inferiority of Blade to Cellavision, given a two-sided level of significance of 5%, 70% power and 5% limit. One possible confounding factor is that different regions of the same film are analysed on test and reference methods, which can lead to variation.

Results

Figure 1. Blade vs. Manual Microscopy

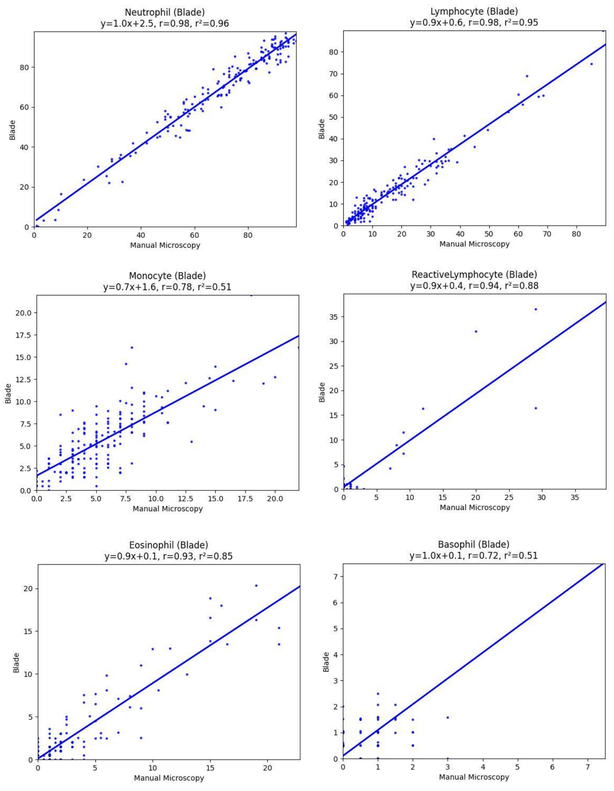

Figure 2. Cellavision vs. Manual Microscopy

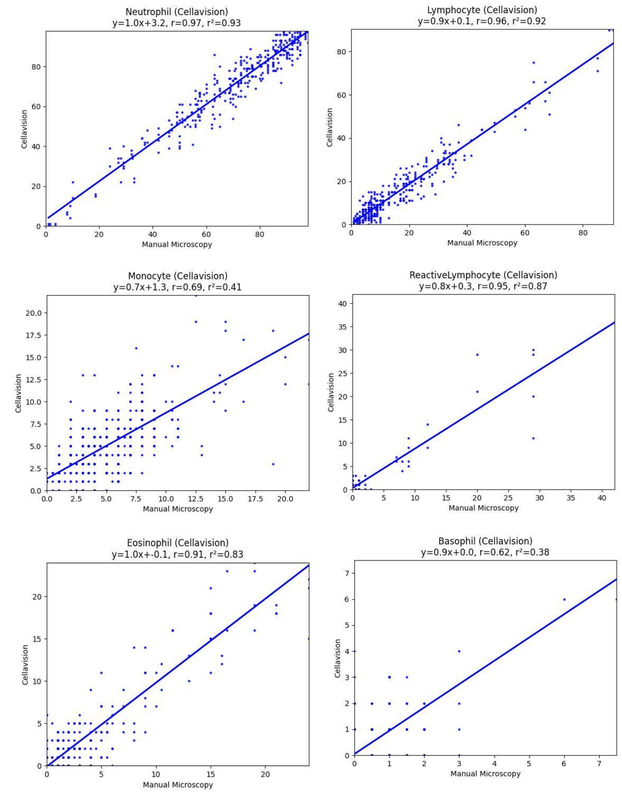

Figure 3. Agreement (Abnormalities)

Characteristic Differential Count Finding (Abnormal)	Proportional Cell Count (%)	n	Agreement with MM	
			Blade	Cellavision
Granulocytosis	> 80	80	92.9%	90.5%
Granulocytopenia	< 10	6	99.4%	99.4%
Eosinophilia	> 7	20	95.2%	96.4%
Lymphocytosis	> 50	13	99.4%	98.8%
Lymphopenia	< 7	46	91.7%	81.0%
Monocytosis	> 10	17	91.7%	92.3%
Myelocytes	> 2	13	94.6%	91.1%
Metamyelocytes	> 2	4	98.8%	96.4%
Blasts	> 2	7	100%	99.4%

Conclusion

The performance of Blade is higher than that of Cellavision. It is non-inferior and comparable in overall performance to Cellavision DM9600. Local laboratory data can possibly be used to augment existing algorithms and train existing commercially available solutions for higher accuracy in WBC identification.

References

Glassy, E.F. (Ed.) (2018) *Color Atlas of Hematology, An Illustrated Field Guide Based on Proficiency Testing*. 2nd Edition. Northfield, Illinois: College of American Pathologists.

Bain, B.J., Bates, I., Laffan, M. A., & Lewis, S.M. (2012) *Dacie and Lewis Practical Haematology*. 11th Edition. London: Churchill Livingstone.