

Article

Reagentless Vis-NIR Spectroscopy Point-of-Care for Feline Total White Blood Cell Counts

Teresa Guerra Barroso ¹, Carla Queirós ², Filipe Monteiro-Silva ³, Filipe Santos ³, António Hugo Gregório ⁴ and Rui Costa Martins ^{3,*}

¹ 1H-TOXRUN—One Health Toxicology Research Unit, University Institute of Health Sciences (IUCS), CESPU, CRL, 4585-116 Gandra, Portugal

² LAQV-REQUIMTE, Faculty of Sciences, University of Porto, R. Campo Alegre, 4169-007 Porto, Portugal

³ INESC TEC—Institute for Systems and Computer Engineering, Technology and Science—Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

⁴ Anicura CHV—Veterinary Hospital Center, R. Manuel Pinto de Azevedo 118, 4100-320 Porto, Portugal

* Correspondence: rui.c.martins@inesctec.pt

Abstract: Spectral point-of-care technology is reagentless with minimal sampling (<10 µL) and can be performed in real-time. White blood cells are non-dominant in blood and in spectral information, suffering significant interferences from dominant constituents such as red blood cells, hemoglobin and bilirubin. White blood cells of a bigger size can account for 0.5% to 22.5% of blood spectra information. Knowledge expansion was performed using data augmentation through the hybridization of 94 real-world blood samples into 300 synthetic data samples. Synthetic data samples are representative of real-world data, expanding the detailed spectral information through sample hybridization, allowing us to unscramble the spectral white blood cell information from spectra, with correlations of 0.7975 to 0.8397 and a mean absolute error of 32.25% to 34.13%; furthermore, we achieved a diagnostic efficiency between 83% and 100% inside the reference interval (5.5 to 19.5×10^9 cell/L), and 85.11% for cases with extreme high white blood cell counts. At the covariance mode level, white blood cells are quantified using orthogonal information on red blood cells, maximizing sensitivity and specificity towards white blood cells, and avoiding the use of non-specific natural correlations present in the dataset; thus, the specificity of white blood cells spectral information is increased. The presented research is a step towards high-specificity, reagentless, miniaturized spectral point-of-care hematology technology for Veterinary Medicine.

Keywords: point-of-care; spectroscopy; white blood cells; artificial intelligence



Citation: Barroso, T.G.; Queirós, C.; Monteiro-Silva, F.; Santos, F.; Gregório, A.H.; Martins, R.C. Reagentless Vis-NIR Spectroscopy Point-of-Care for Feline Total White Blood Cell Counts. *Biosensors* **2024**, *14*, 53. <https://doi.org/10.3390/bios14010053>

Received: 20 December 2023

Revised: 16 January 2024

Accepted: 18 January 2024

Published: 19 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Blood spectra are governed by multi-scale interference and matrix effects, superimposing physical (scattering) and chemical (absorbance) information. Unscrambling the information contained in spectral data, due to multi-scale interference, is the major challenge for point-of-care (POC) technologies [1,2] (Figure 1a). Multi-scale interference results from overlapping spectral bands of blood constituents, leading to interference at varying intensities and deviations from the Beer–Lambert law (BLL) [3]. Matrix effects are inherent to sample properties, such as pH or scattering effects, affecting the proportions of Mie and Rayleigh scattering and absorbance bands [4,5]. Clinical and analytical chemistry address spectral interference through sample composition simplification (e.g., separation using filtration, centrifugation, chromatography) or reaction specificity on biochips (e.g., immunological reactions, molecularly imprinted polymers, aptamers), not taking advantage of the rich information present in spectra [6–10].

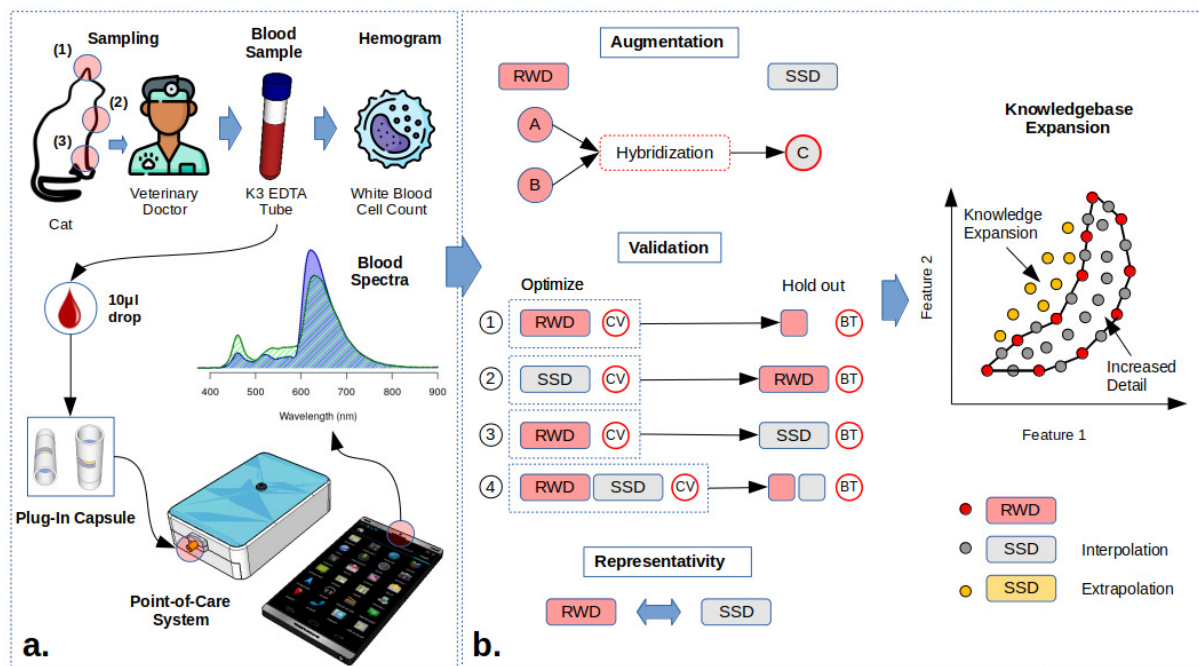


Figure 1. Total white blood cell counts: (a) sampling and spectroscopy procedures—venipuncture performed at the jugular (2) or cephalic (3) veins for hemogram analysis and single drop for POC spectral recording; a single drop can also be collected at the auricular vein (1). (b) Data augmentation through hybridization of real-world data (RWD) into synthetic spectroscopy data (SSD), validation procedures using RWD and SSD as optimization and validation datasets to test SSD representativity of RWD and knowledgebase expansion.

Unscrambling the multi-scaled interference and matrix effects, and unlocking constituent-specific information in each spectrum, lies at the heart of reagentless POC technology [1,11,12]. The use of artificial intelligence in biosensor development and operation is currently a reality [13,14]. Significant advances in searching for consistent covariance modes capable of isolating both inference and constituents information has allowed for the provision of information equivalence between spectra and hemograms, leading to the precise quantification and diagnosis of clinical conditions [1,11–16] and demonstrating the possibility of quantifying white blood cells (WBCs) in dog blood [2]. Herein, we explore the quantification of WBCs in cat blood through the isolation of specific WBC information in blood spectra, taking advantage of the feline physiology for spectral quantification, such as the use of scattering effects of WBCs vs. red blood cells (RBCs) and avoiding the use of inferential WBC quantification, such as the natural relationship between RBC and WBC levels.

1.1. White Blood Cells and Blood Spectroscopy

The measurement of WBCs is crucial in hematology due to its wide-ranging diagnostic significance, encompassing conditions like infections and leukemia. The main reasons for leucocytosis (increased WBC count) are inflammation, effects of stress, corticosteroids, physical activity, and presence of epinephrine [17]. High WBC counts are mostly associated with neutrophilia and inflammation, whereas neutropenia is mainly caused by increased peripheral tissue demand during intense inflammation and/or decreased bone marrow production [18]. Due to a substantial reservoir of neutrophils in the bone marrow, dogs and cats rarely experience leukopenia (low WBC count) [17].

Blood spectra information is dominated by the strong absorbance of hemoglobin (Hgb) [19] and bilirubin (Bil) [20]. RBCs dominate spectral fingerprinting and the presence of interferences must be taken into account, e.g., Bil interfering with Hgb-related features [1,19].

WBCs are traditionally measured by electrical impedance, laser light scattering, radiofrequency conductivity and flow cytometry [21]. Automated hematology analyzers rely on cell size, which impacts both impedance and scattering angle measurements. Therefore, development stage variations in blood cell sizes, and platelet (PLT) clumping [22], are the main reasons for inaccuracies in modern automated hematology systems [23–25]. To check the automated results, microscopy blood smear counting is applied as a gold standard quality control measure in clinical practice [25], which also allows for the analysis of nuclei appearance, cell size, cytoplasmic granules and toxic cytoplasm [26,27].

Spectroscopy is not limited by sample volume or reagents, making it well-suited for portable POC technology. Figure 1a showcases the prototype system, which integrates Internet of Things (IoT) electronics and software, controllable through a smartphone without the need for a dedicated app. A single drop of blood (10 μ L) is placed in a reusable capsule designed for insertion into the transmittance probe tip [1] (Figure S1). These capsules are constructed with opposing mirrors to optimize internal reflections and light is collected via a central pinhole fiber optic connected to the spectrometer that operates within the 300–800 nm wavelength range [1].

Ultraviolet–visible (UV-Vis) spectroscopy distinguishes the different organelles of leucocytes, including cytosol, nuclei, mitochondria and membranes [28]. Changes in absorbance (200–400 nm), fluorescence/phosphofluorescence and infrared patterns of WBCs reflect significant alterations in organelle composition, allowing for the diagnosis of leukemia [29–33] and its progression [34–37]. Infrared microscope spectroscopy of WBC components was used to diagnose hidden infections through support vector machines classification [38]. These studies show the possible existence of unique spectral features associated with WBCs, which allows us to overcome the problem of low quantities of WBCs in comparison to RBCs—a ratio of approximately 1:1000.

Feline Hgb has a lower affinity to oxygen (and is more readily uptaken by tissues) [26], resulting in a lower packed cell volume (PCV) and Hgb concentration than in canines [39]. Feline PLT are larger and exhibit greater size variation (15 to 18 μ m) in comparison to other domestic mammals [26]. In specific situations, there can be a strong correlation between RBCs and WBCs, and special care must be taken because optimal chemometrics/AI models can be dependent on intrinsic data correlations and a lack of causal connections to constituent absorbance [40].

Reagentless spectral POC demonstrated its effectiveness in quantifying hemogram parameters in dog and cat blood [1,2]: (i) Dog blood—RBCs (6.39%), Hgb (7.14%), hematocrit (HTC) (4.43%); (ii) cat blood—RBCs (5.67%), Hgb (4.08%), and HTC (1.69%). WBCs in dogs were quantified with a correlation of 0.8478, standard error of 6.92×10^9 cells/L and MAPE of 25.37%, allowing us to diagnose WBCs in the reference interval ($5.6\text{--}17.8 \times 10^9$ cells/L) [2].

In pure constituents, the signal varies uni-dimensionally along a single eigenvector, as it has only one source of variance—a covariance mode (CovM). CovMs isolate the interference between constituents by searching for a group of samples that maximizes the variance of the constituent being quantified [1,12]. In special cases, CovMs can be orthogonal, isolating specific and unique information [40]. Due to the significant biological variability, the covariance of large representative datasets becomes unstable and exhibits high dimensionality. It becomes necessary to unscramble various types of interference to accurately relate the quantitative information of a specific constituent within the context of its interferents through CovMs [12].

Each sample within a CovM exhibits a consistent covariance between spectral information (**X**) and constituents (**Y**). This implies the two blocks of information are similar, despite presenting with different bases (wavelengths vs. concentration). Consequently, the two information blocks demonstrate latent structural similarity ($\mathbf{t} \sim \mathbf{u}$), where **t** and **u** are derived independently via the singular value decomposition of **X** and **Y**: $\mathbf{X} = \mathbf{T}\mathbf{P}^t$ and $\mathbf{Y} = \mathbf{U}\mathbf{Q}^t$. Here, **P** and **Q** are the orthogonal bases of **T** and **U**, respectively. Ideally, within each CovM, the interference information is equivalent to the concentration ($\mathbf{t} \sim \mathbf{u}$), represented by a single eigenvector or one latent variable (1 LV). This offers a causal inter-

pretation of spectral interference by associating the absorbance bands of constituents [3] with the BLL law relationship [1,12].

1.2. Data Augmentation

Data augmentation can be used to enhance the diversity of the knowledgebase and improve the accuracy of model predictions [2,41,42]. Blood composition representativity is limited in real world data (RWD) due to the large biological variance. This is partially solved by self-learning models, which can assess if the measured spectra are predictable, suggesting an update to the knowledgebase [12]. Synthetic spectroscopy data (SSD) are generated by the random hybridization of the hemogram and spectral information from two samples (Figure 1b), creating new spectra and a corresponding hemogram with a 1:1 information mixture, which is comparable to the mixture of the two blood samples, resulting in the average composition of the two RWD samples and assuming no reaction between blood samples (e.g., antigen–antibody of blood type and Rhesus factor) [2]. SSD improves the knowledgebase by (i) increasing the level of local detail by the interpolation of the existing RWD samples and (ii) expanding the knowledgebase through extrapolation, once the boundary coordinates of the samples in the feature space are convex (Figure 1b).

The objective of this research is to demonstrate the feasibility of spectral WBC quantification in cat blood by (i) demonstrating the capacity of SSD to both represent and increase the level of detail of the RWD knowledgebase; (ii) benchmarking the global covariance through partial least squares (PLS) versus local covariance through Local-PLS (LocPLS) and Self-learning Artificial Intelligence (SLAI); and (iii) determining the WBC diagnostic potential of spectral POC through bias-variance analysis inside and outside the reference interval (RI).

2. Methods

2.1. Hemogram Analysis

Blood samples from cats, which are already employed in clinical diagnostic procedures, were obtained from the jugular vein through standardized venipuncture procedures conducted by qualified personnel at the Anicura CHV—Veterinary Hospital Center. The remaining blood from EDTA tubes, collected previously and still in a fresh state, was subsequently utilized for these tests. Hemogram parameters were determined using the Mindray BC-5800-vet auto-hematology analyzer (Mindray, Shenzhen, China) [43] (Figure 1a).

2.2. Spectroscopy

Figure 1a illustrates the Visible Short-Wave Near-Infrared (Vis-SWNIR) POC and IoT prototype platform [44]. It uses a Hamamatsu C12666MA spectrometer that manages light sources (LED or laser diodes). This version uses a power LED (4500 K) as a light source (Figure 1a). Optical configuration used a transmittance probe with six illuminations and a central collection fiber. Blood is analyzed inside a plug-in capsule with an opposing mirror configuration (5 mm path) (see Figure S1). The average of three blood spectra is recorded and spectra are corrected for scattering effects before further processing [45]. A total of 94 blood samples from different cats were used in this study for hemogram and spectral records.

The spectral POC operation is exemplified in Figure S1, with the following steps: (1) withdrawal of 10 μ L of blood from the ear lobe; (2) injection of blood sample into the capsule system; (3) plug-in of the capsule into the POC probe; and (4) use of IoT software to control POC measurement and record spectra for WBC calculation.

2.3. Spectral Data Augmentation

A synthetic dataset is created by the hybridization of two random RWD samples. Hybrid samples do not exist in the RWD, and may or not reflect existing blood samples in feline species (Figure 1a). However, the combination of information promotes knowledgebase expansion by spawning spectral gradients proportional to blood hemograms.

A total of 300 SSD samples were obtained by non-repeated random mixing of pairwise RWD spectral samples ($\mathbf{X}$) and hemogram composition ($\mathbf{Y}$), where $\mathbf{X}_{SSD} = \frac{1}{2}(\mathbf{X}_i + \mathbf{X}_j)$ and $\mathbf{Y}_{SSD} = \frac{1}{2}(\mathbf{Y}_i + \mathbf{Y}_j)$, with i and j random samples. Hybridization creates a synthetic sample with average spectral ($\mathbf{X}_{SSD}$) and blood composition ($\mathbf{Y}_{SSD}$) characteristics, where each pixel in the spectra is the average pixel intensity of $\mathbf{X}_i$ and $\mathbf{X}_j$ real-world spectra at a given wavelength, and the $\mathbf{Y}_{SSD}$ WBC value is the average value of two real-world hemogram WBC values, $\mathbf{Y}_i$ and $\mathbf{Y}_j$.

Figure 1b describes how SSD was tested for RWD representativeness (null hypothesis) by splitting model optimization by cross-validation (CV) and blind testing (BT): (1) RWD benchmark leave-one-out optimization and leave-one-out holdout blind testing in PLS models (RWD); (2) SSD optimized—leave-one-out CV optimization and blind testing with RWD (SSD>RWD); (3) optimized leave-one-out CV optimization and BT testing with SSD (RWD>SSD); and (4) joint RWD and SSD and random splitting into optimization using CV and BT datasets (90:10) (Figure 1b, RWD + SSD). Rotation to RWD and SSD datasets in schemes 1 (RWD) and 4 (RWD+SSD) was performed during optimization, so that all samples could be evaluated in BT. The increase in the information in a two-dimensional feature space via data augmentation is graphically represented in Figure 1b. RWD restricts the knowledgebase to a polygon area within the feature space (Figure 1b). Inside the polygon, the SSD is generated using the RWD samples at the edges, resulting in an increase in the representativeness of the local covariance.

Extrapolated samples in yellow in Figure 1b are obtained through the hybridization of the convex boundary samples of the RWD polygon (Figure 1). Extrapolated SSD retains the information on spectral gradients and interferences to be proportional to nutrient composition, extrapolating unknown characteristics.

2.4. Benchmarking

CovM search methodology was benchmarked against chemometrics methods:

- i. Partial least squares regression (PLS): PLS maximizes the covariance between the spectra matrix $\mathbf{X}$ and the blood hemogram matrix $\mathbf{Y}$, by determining the eigenvectors of $\mathbf{X}^t\mathbf{Y}$ [46,47]. The method forces the latent structures of spectra and hemogram ($\mathbf{t}$ and $\mathbf{u}$) to be equal (NIPALS or SVD algorithm), for the correspondent basis $\mathbf{P}^t$ and $\mathbf{Q}^t$ [46], where $\mathbf{X} = \mathbf{TP}^t$ and $\mathbf{X} = \mathbf{UQ}^t$. It further deflates sequential orthogonal eigenvectors of the remaining information in $\mathbf{X}^t\mathbf{Y}$. The number of latent variables (LV) is optimized by the cross-validation of the minimal predicted sum of squares. PLS uses an oblique projection to determine the b_{pls} coefficients in $\mathbf{Y} = \mathbf{X}b_{pls}$, which, albeit it linear combination of LVs, can model spectral information non-linearities [47,48].
- ii. Local PLS (LocPLS): breaks down the complexity of the feature space by KNN classification based on the similarity of spectral features given by the Euclidean distance between sample coordinates. An ensemble of PLS models is generated for each cluster. The dimensions of the feature space and number of the cluster was optimized by cross-validation/hold-out samples [49].
- iii. Self-learning Artificial Intelligence (SLAI): uses a two-step approach that (a) features space optimization for providing equivalence between the spectral information ($\mathbf{T}$) and blood hemogram ($\mathbf{U}$) latent spaces; and (b) performs a local covariance mode (CovM) search for unscrambling groups of samples that have the same type of interferences, providing a direct quantification between spectra and nutrients with a single dimension (or eigenvector) exhibiting stable co-variance ($\mathbf{X}^t\mathbf{Y}$) [1].

A two-step approach was used to build and validate the models: (i) model optimization through cross-validation using a training set; and (ii) prediction and corresponding metrics using blind testing samples. Local models (LocPLS and SLAI) use the leave-one-out CV with one hold-out sample in the model optimization step. The performance was evaluated according to the standard deviation (SE), mean absolute percentage error (MAPE), and Pearson correlation (R), as the linearity between WBC and predicted values. CRAN-

R software was used for all computations (PLSR package; LocPLS and SLAI using the author's code) [50].

2.5. Total Error and Bias Analysis

The American Society for Veterinary Clinical Pathology (ASVCP) establishes the total allowable error (TAE) of the clinical decision target values (TVa), which, in the case of hemograms, are the boundary values of the Reference Interval (RI, 5.5 to 19.5×10^9 cell/L) [25], with a maximum TV = 23.7×10^9 cell/L. The distribution of the total error (TE) results in ASVCP classes of optimal 7.15% (opt), desired 14.29% (des) and acceptable 21.45% (acp), which were used as bias performance indicators [51]. The percentage of correct diagnoses with increasing WBC count was used to assess the diagnostic capacity of the spectral POC.

3. Results and Discussion

3.1. WBC Blood Spectroscopy

Figure 2a presents two typical blood spectra with low and high WBC levels. WBC spectral information is spread along the wavelengths 400–800 nm. WBC increase leads to an increase in absorbance in the region of interest from 400 to 600 nm (ROI 1) and a decrease in absorbance from 600 to 800 nm (ROI 2). WBCs and RBCs/Hgb are interferents because several species of Hgb absorb in ROI 1 and 2, with higher intensity in the region of 530 to 576 nm, and with smaller contributions at 430 nm and 630 nm, with the following hemoglobin absorption bands: (i) HbA—433 nm and 558 nm; (ii) Oxy HbA—539 nm; (iii) Carboxy HbA—550; (iv) Ferrous HbA—541 nm and 576 nm; (v) Deoxy HbA—555 nm; (vi) Ferryl HbA—545 nm; and (vii) Ferric HbA—630 nm. [1,19,52]. Similar results were observed in dog blood, where ROI 1 and ROI 2 showed similar changes in absorbance for high and low levels of WBCs, with a more significant interference from RBCs/Hgb in dogs [2].

The scattering ratio of RBCs and WBCs can be defined as $S = 2\pi R/\lambda$, where R is the particle radius and λ the wavelength, Cat RBCs and WBCs have a diameter approximately of 5.5 μm and 18 μm , and are scattering dominated by geometric effects ($S \gg 1$), which partially explains the observed spectra (Figure 2c).

Despite of the larger number of RBCs in cats, these are smaller and contain less hemoglobin. The significant size discrepancy between WBCs and RBCs provides WBC information in blood spectra due to the increased scattering and absorption (Figure 2c). The scattering ratios of RBCs at 400 and 800 nm were 86 and 43, whereas those for WBCs were 282 and 143. WBC presence in the blood leads to a significant decrease in the observed spectral intensity, which is more relevant in ROI 1. As scattering is less intense at higher wavelengths, ROI 2 is significantly less affected by WBCs and RBCs, whereas with more WBCs, the signal increases in ROI 1 (Figure 2a,c).

The surface exposure to light of WBC is $\sim 213.71 \mu\text{m}^2$, whereas, in RBC is $\sim 23.75 \mu\text{m}^2$, the cell thickness of WBCs ($\sim 10 \mu\text{m}$) is significantly higher than anucleated RBCs ($\sim 2 \mu\text{m}$). The smaller area, thickness and lack of nuclei leads to more space between RBCs at low WBC levels (Figure 2c). High levels of WBCs significantly increase the absorbed light, as WBCs increase the exposed surface and light path length, as well as the cell packing in smaller spaces within the blood (Figure 2c). This results in light exposure ratios of 0.5% for combinations of low WBCs (5×10^9 cell/L) and high RBCs (9×10^{12} cell/L), and 22.5% for combinations of high WBCs (50×10^9 cell/L) and low RBCs (2×10^{12} cell/L). The absorbance ratio also increases at high WBC levels, as the light path length through cells can increase three to five times, supporting the reason why WBC information is present in the spectra.

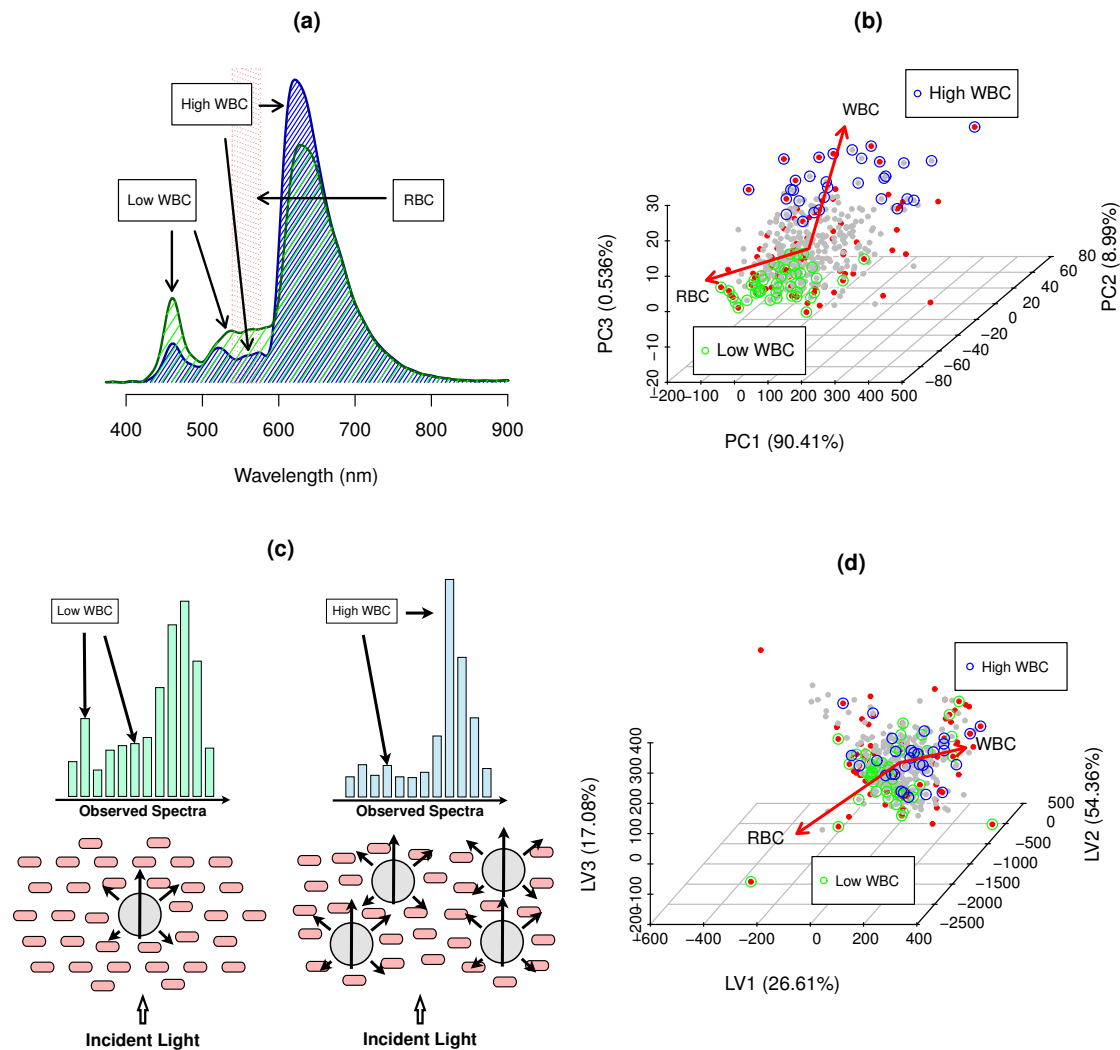


Figure 2. WBC spectral information: (a) cat blood spectra (■) low WBCs, (■) high WBC examples, and (■) RBC Hgb major interference bands (539–576 nm); (b) PCA scores of hemogram counts; and (c) Cell packing, scattering and WBC absorbance effects and impact on observed spectra; (d) PLS scores of blood spectra—translating maximum covariance to WBCs, where • SSD blood samples, • RWD blood samples, • low WBCs and • high WBCs and → hemogram PCA loadings and main gradient direction in PLS scores space.

RGB and WBC packing limitations impose a physical restriction, resulting in the intrinsic negative correlation between RBC and WBC latent spaces (Figure 2b,d). As WBC levels increase in abnormally high levels, the available free plasma volume for RBCs decreases, which imposes a natural negative correlation, which, despite being statistically valid, has no relation to specific WBC spectral information.

Figure 2b,d present the first three scores of the hemogram PCA analysis and spectral WBC PLS. PCA calculates the orthogonal eigenvectors that maximize the variance of the hemogram, whereas PLS maximizes the covariance between the spectra (X) and WBCs (Y) by extracting the orthogonal eigenvectors of the covariance matrix ($X^t Y$). The quantification of WBCs is only feasible if the spectral information has a similar eigenstructure to the hemogram variance space. If X is the spectra (samples $\times$ wavelength), then Y is the hemogram (samples $\times$ [RBC, Hgb, HTC, WBC, PLT]), being decomposed into: $X = T P^t$ and $Y = U Q^t$, where T and U are the coordinates (aka scores), P^t and Q^t (aka loadings). Once X and Y share significant common equivalent information, their eigenstructures must also reflect such similarities. Therefore, it is expected that T and U have a similar structure

and sample distribution along their space coordinates, despite being expressed by different bases $\mathbf{P}^t$ and $\mathbf{Q}^t$; this is because spectra carry hemogram information ($\mathbf{T} \simeq \mathbf{U}$).

Hemogram variance space $\mathbf{U}$ is presented in Figure 2b, where 99.94% of the variance is represented by three principal components, PC1 (90.41%), PC2 (8.99%) and PC3 (0.536%), where high levels of RBCs ($8\text{--}10 \times 10^{12}$ cell/L) and WBCs ($20\text{--}50 \times 10^9$ cell/L) dominate hemogram variance. RBCs and WBCs in the cat blood hemogram increase in opposite directions and are not orthogonal, exhibiting the natural negative correlation that is expected in the interaction of cell packing, scattering and absorbance effects, as shown by the main variance vectors (red arrows, Figure 2b).

Spectra have more information than blood hemograms, as they reflect both cellular and serum fractions. This affects the structure of $\mathbf{T}$ (Figure 2d). Only the WBC information should be used in the quantification provided by similarity ($\mathbf{T} \simeq \mathbf{U}$) at each CovM. $\mathbf{T}$ (Figure 2d) has a high degree of similarity with the hemogram space $\mathbf{U}$. Three LVs represent 98.77% of the covariance with WBCs (LV1 (26.94%), LV2 (54.36%) and LV3 (17.08%)). High WBC values also showed a negative correlation with high RBCs, where the main direction vectors were non-orthogonal (Figure 2d), maintaining a similar relationship to that shown in the hemogram space ($\mathbf{U}$) (Figure 2b). The structural similarity between $\mathbf{T}$ and $\mathbf{U}$ shows the hemogram and spectral information are similar at both global and local variance levels, allowing us to relate WBCs to spectral variance on both scales [1,2,12].

3.2. Significance and Representability

A significant challenge for the spectral quantification of non-dominant constituents, such as WBCs, is to establish a well-balanced representation of the knowledgebase, with a detail level allowing the isolation of consistent small-scale spectral variance. In this study, the RWD provides a low local representation of the global covariance ($\mathbf{X}^t\mathbf{Y}$), as the p -value of PLS-RWD is 0.0001. Blood spectra covariance is highly structured, with significantly high levels of locally stable co-variances, leading to the existence of multiple interference modes [1,2]. PLS-RWD can only relate the major spectral covariance with 2LVs to WBC. The PLS model averages the existing local covariances, but is not able to extract further LVs due to the lack of local representativeness, mostly capturing the natural negative correlation of WBCs and RBCs due to cell packing and scattering.

The results show that SSD provides a stable global covariance matrix ($\mathbf{X}_{SSD}^t\mathbf{Y}_{SSD}$), allowing the PLS algorithm (PLS-SSD>RWD) to extract 15 LVs with a p -value of 1.12×10^{-5} . The large number of LVs is consistent with a significant number of locally stable interferences. Another proof of SSD representativity of RWD is SSD prediction using the PLS-RWD>SSD model (p -value $< 2 \times 10^{-16}$), which leads to the conclusion that RWD and SSD represent the same information (Table 1, Figure S2).

Table 1. Results benchmarks for PLS, LocPLS and SLAI for real-world data, and combination of real-world data and synthetic samples data for WBC quantification in cat blood.

Method	nLV	R	Real World Data (RWD)				R	Synthetic Sample Data (SSD)			
			R ²	SE	MAPE (%)	p -Value		R ²	SE	MAPE (%)	p -Value
PLS ⁽¹⁾	2	0.1874	0.0350	5.498	55.45	0.001	–	–	–	–	–
PLS ⁽²⁾	15	0.5775	0.3336	6.109	44.56	1.12×10^{-9}	0.6802	0.4626	5.135	65.65	$< 2 \times 10^{-16}$
PLS ⁽³⁾	2	0.1874	0.0350	5.498	55.45	0.001	0.2634	0.0691	9.162	44.60	0.013
PLS ⁽⁴⁾	13	0.5553	0.3084	7.925	57.31	$< 2 \times 10^{-16}$	0.6320	0.3995	9.736	34.18	$< 2 \times 10^{-16}$
LocPLS ⁽²⁾	5	0.4546	0.2067	8.576	58.89	4.16×10^{-6}	0.6840	0.4679	4.480	31.19	$< 2 \times 10^{-16}$
LocPLS ⁽⁴⁾	7	0.7112	0.5059	6.673	46.78	$< 2 \times 10^{-16}$	0.6782	0.4599	9.188	32.16	$< 2 \times 10^{-16}$
SLAI ⁽²⁾	1–2	0.7975	0.6361	5.837	34.13	$< 2 \times 10^{-16}$	0.7369	0.5460	8.492	22.35	$< 2 \times 10^{-16}$
SLAI ⁽⁴⁾	1	0.8397	0.7051	5.192	32.25	$< 2 \times 10^{-16}$	0.7723	0.5965	7.947	25.86	$< 2 \times 10^{-16}$

⁽¹⁾ Cross-validation optimization and hold-out metrics calculation using RWD. ⁽²⁾ Cross-validation optimization using SSD and hold-out metrics calculation using RWD. ⁽³⁾ Cross-validation optimization using RWD and hold-out metrics calculation using SSD. ⁽⁴⁾ Cross-validation optimization and hold-out metrics calculation using RWD and SSD.

Local models (LocPLS, SLAI) with SSD>RWD attain significant p -values (4.16×10^{-6} and $<2 \times 10^{-16}$, Table 1). Local representativity is reflected in the lower number of LVs shown by LocPLS (5 LVs) and SLAI (1–2 LVs), resulting in a higher efficiency in isolating interference and accuracy in WBCs. Augmentation through hybridization increases local information by creating samples with mixed properties, complementing the information of the under-represented coordinates of the feature space (Figures 1b and 2c,d). The statistical significance of SSD representativity shows that hybridization is an efficient method for increasing knowledgebase information. Further evidence of the increased information provided by SSD can be seen when RWD and SSD are used together. PLS, LocPLS and SLAI show significant p -values ($<2 \times 10^{-16}$, Table 1). RWD has a significant diversity of information for SSD combinations that can improve local representativity.

3.3. WBC Quantification

The PLS-RWD has a low correlation ($R = 0.1874$) and a high MAPE (55.45%). The low correlation and number of LVs (2 LVs) shows that PLS-RWD lacks the local detailed information to enable the incorporation of WBC characteristics into more LVs. When PLS-RWD is used to predict SSD, a low correlation ($R = 0.2634$) and high error (MAPE = 44.60%) are obtained (Table 1).

R and MAPE improve significantly when SSD is used to predict RWD. PLS-SSD>RWD exhibits an $R = 0.5775$, MAPE = 44.56% and five LVs. A similar result is achieved when RWD and SSD are used in the PLS-RWD+SSD model ($R = 0.5573$, MAPE = 57.31%) using 13 LVs (Table 1). A higher number of LVs occurs due to the increase in the detail of the local relationship between T and U in the SSD dataset. WBC information is present in low-intensity variation in the local spectra, as WBC information is present in low-intensity variation in local groups.

The use of SSD in local models improves WBC quantification. LocPLS-SSD>RWD improves WBC quantification ($R = 0.4546$, MAPE = 58.89%), reducing the number of LV to five. LocPLS SSD prediction is better than RWD. RWD has abnormally high WBC values, which are under-represented in SSD. Hybridization is not very efficient for solving class imbalance at extreme values. Using SSD and RWD as the knowledgebase mitigates this problem. The LocPLS-RWD and SSD model has less extreme values of class imbalance, and, therefore, RWD prediction is significantly improved ($R = 0.7112$, MAPE = 46.78%), even when compared with SSD ($R = 0.6782$, MAPE = 32.16%), without a significant increase in LVs (seven LV).

SLAI-SSD>RWD has consistent results between SSD ($R = 0.7369$, MAPE = 22.35%) and RWD ($R = 0.7975$, MAPE = 34.13%), with only one to two LVs (Figure 3a, Table 1). SSD has sufficient details to be unscrambled into 40 CovMs.

Figure 3a shows that SLAI-SSD>RWD predictions inside the RI ($5.5\text{--}19.5 \times 10^9$ cell/L) exhibit significant dispersion, as well as in the ASVCP High WBC clinical decision interval ($SE = 5.837 \times 10^9$ cell/L; Figure 3a). Higher errors were observed at higher WBC levels ($>30 \times 10^9$ cell/L), due to the lack of better representativity at these levels and higher errors in the quantification of automated hematology. Similarly to LocPLS, joining RWD and SSD, the SLAI-RWD and SSD model significantly reduces the error at high WBC values (Figure 3b), with a significant improvement in WBC prediction of RWD ($R = 0.8397$, MAPE = 32.25%) and SSD ($R = 0.7723$, MAPE = 25.86%), making it possible to increase the detail to 60 CovMs with one LV.

3.4. Bias–Variance Analysis

Spectral POC linearity (Pearson R) and total error (MAPE) benchmarks are compared in Figure 4a,b. PLS is unable to provide qualitative or quantitative correlations, as R is below 0.65 for PLS-RWD, PLS-SSD>RWD and PLS-RWD>SSD. LocPLS has a significant increase in R , but WBC quantifications are qualitative only when RWD and SSD are used together. LocPLS also averages existing CovMs in a local group to build an ensemble of PLS models, but is still unable to unscramble the different types of interferences present in

the knowledgebase (Figure 4a). SLAI has a semi-quantitative correlation ($R \approx 0.75\text{--}0.85$). SLAI-SSD>RWD attains $R = 0.7975$, and, when using the full dataset, SLAI-RWD+SSD obtains $R = 0.8397$. SLAI unscrambles SSD and RWD into one LV CovMs.

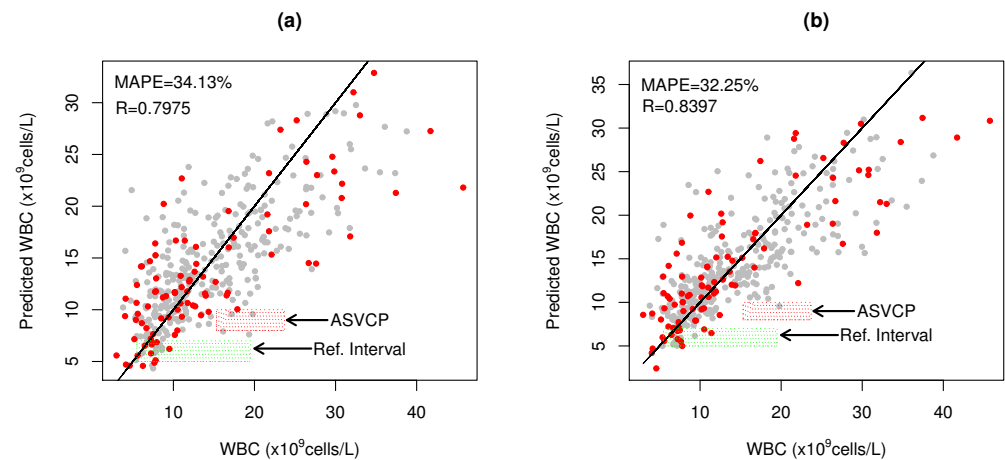


Figure 3. WBC prediction for (a) SLAI with SSD as CV optimization and RWD as HO blind test samples; (b) SLAI with RWD and SSD as CV and HO samples, where (*) represent the hybridized SSD samples and (•) the RWD blood samples, respectively. Green shaded rectangle (■) represents the WBC reference interval for cats ($5.5\text{--}19.5 \times 10^9$ cells/L) and red shaded rectangle (■) represents the ASVCP total allowable error tolerance for high WBC diagnosis.

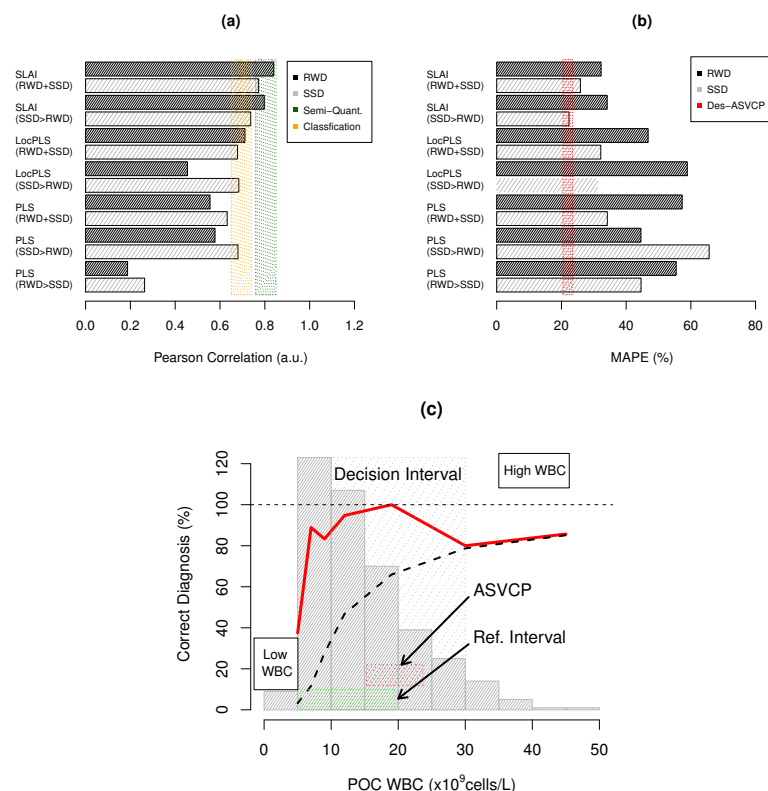


Figure 4. WBC quantification benchmarks: (a) Pearson correlation coefficient for RWD and SSD, where shaded rectangles represent semi-quantitative (■) and quantitative (■) results; (b) MAPE benchmarks where (■) represents the ASVCP TAE for WBCs (21.45%); and (c) POC WBC diagnostic capacity: (—) percentage of correct diagnoses within WBC value interval, (---) accumulated percentage of correct diagnosis, (■) WBC histogram distribution; and RI for WBC (■) and (■) TAE diagnosis error tolerance.

Figure 4b shows the total error in WBCs for RWD and SSD under the different modeling approaches. All models were unable to obtain a TE below the TAE desired by ASVCP (TE $\sim$ 21.45%). PLS and LocPLS models have a high TE: (i) PLS (MAPE $\sim$ 44.56–67.79%) and (ii) LocPLS (MAPE $\sim$ 46.78–58.89%). SLAI attains a semi-quantitative TE of 32.15–34.13% in RWD, nearer the TAE.

Table 2 presents the percentage of cases with optimal, desired and acceptable TE, inside and outside the RI. PLS-SSD>RWD attains 39.06% and PLS-RWD+SSD attains 42.19% when predicting the RWD. PLS has a similar performance in predicting RWD (42.19%) and SSD (47.77%) when used together (PLS-RWD+SSD), where the results inside the RI are significantly better than those outside the RI. LocPLS-RWD+SSD has 46.88% RWD acceptable results inside the RI and 36.67% outside the RI. LocPLS-SSD>RWD had poor results (29.69%) but presented 52.68% inside and 52.63% outside the RI for SSD (Table 2).

SLAI has the highest values of acceptable TE for RWD inside the RI, 50.00% using SLAI-SSD>RWD and 54.69% for SLAI-RWD+SSD. It also provides the lowest TE outside the RI for RWD (43.33% and 50.00%). The TE outside and inside RI are very consistent in SLAI when using only SSD or RWD with SSD (Table 2), decreasing from 54.69% to 50.00% in RWD and from 58.03% to 54.91% when not using the combination of RWD and SSD. Despite the SLAI performance, it cannot attain the TAE for WBCs of 21.45% [25]. This limit can be considered optimistic for today's state-of-the-art hematology, as state-of-the-art WBC and differentiated counts still have significant errors, and a wider decision interval must be considered in clinical practice [53]. Figure 4c presents the percentage of correctly diagnosed WBC cases and the percentage of accumulated diagnosed WBC cases for SLAI-SSD>RWD for each WBC level. The percentage of correctly diagnosed cases inside the RI ranges from 83.33% to 100% inside the RI, and 90% within the diagnostic error margin of high WBC levels. The diagnostic accuracy reaches 80% for extreme cases above 30×10^9 cell/L.

The study reveals high representativity within the RI ($5.5\text{--}19.5 \times 10^9$ cell/L), compared to extremely low or high levels above and below the RI. Besides the low representativity of abnormal values, extreme WBC values are associated with analytical errors in automated hemogram machines, due to the presence of abnormal cells or cell aggregates (e.g., platelet aggregates, parasitic inclusions, nuclear segmentation), which can be recognised as different types of WBCs, due to the similar scattering angles. In contrast, the high accuracy of the POC is found in the region of 7 to 30×10^9 cell/L, which incorporates both RI and decision tolerance for high WBC diagnosis.

Table 2. WBC bias analysis of WBCs in cat blood for spectroscopy POC—percentage of results in optimal, desired and acceptable categories.

Method	Real World Data						Synthetic Sample Data					
	% Inside RI			% Outside RI			% Inside RI			% Outside RI		
	Opt	Des	Accep	Opt	Des	Accep	Opt	Des	Accep	Opt	Des	Accep
PLS _{RWD} ⁽¹⁾	7.81	14.06	26.56	0.00	0.00	10.00	na	na	na	na	na	na
PLS _{SSD>RWD} ⁽²⁾	21.88	29.69	39.06	10.00	16.67	20.00	15.63	29.02	46.43	13.16	22.37	43.42
PLS _{RWD>SSD} ⁽³⁾	7.81	14.06	26.56	0.00	0.00	10.00	16.07	25.89	34.38	1.32	2.63	9.21
PLS _{SSD+RWD} ⁽⁴⁾	10.94	26.56	42.19	3.33	10.00	13.33	13.84	29.91	47.77	9.21	22.37	26.32
LocPLS _{SSD>RWD} ⁽²⁾	14.06	25.00	29.69	13.33	16.67	20.00	23.21	41.52	52.68	21.05	35.53	52.63
LocPLS _{RWD+SSD} ⁽⁴⁾	15.63	32.81	46.88	16.67	23.33	36.67	19.20	32.14	45.98	14.47	32.89	44.74
SLAI _{SSD>RWD} ⁽²⁾	20.31	34.38	50.00	13.33	30.00	43.33	23.66	38.39	54.91	17.11	26.32	43.42
SLAI _{RWD+SSD} ⁽⁴⁾	31.25	42.19	54.69	13.33	26.67	50.00	24.12	41.07	58.03	30.26	35.52	46.05

⁽¹⁾ Cross-validation optimization and hold-out metrics calculation using RWD. ⁽²⁾ Cross-validation optimization using SSD and hold-out metrics calculation using RWD. ⁽³⁾ Cross-validation optimization using RWD and hold-out metrics calculation using SSD. ⁽⁴⁾ Cross-validation optimization and hold-out metrics calculation using RWD and SSD. Highlighted bold values are considered are considered best in class.

3.5. CovMs Interpretation

Figure 5 presents two CovMs at low and high WBC levels. Figure 5a shows the PLS scores space (T), where low and high WBCs are represented with green and blue circles,

respectively. High WBC CovM spectra and quantification are presented in Figure 5c,d, with $R = 0.7718$ and $\text{MAPE} = 22.75\%$ within the range of $20\text{--}45.5 \times 10^9$ cell/L. Low levels of WBC CovM spectra and quantification are presented in Figure 5e,f, with $R = 0.7497$ and $\text{MAPE} = 0.7497$, in the range of $6\text{--}12 \times 10^9$ cell/L. High levels of WBCs are quantified using the regions of 430–470 nm and 580–680 nm, whereas low levels of WBC are quantified using 450–500 nm and 650–770 nm.

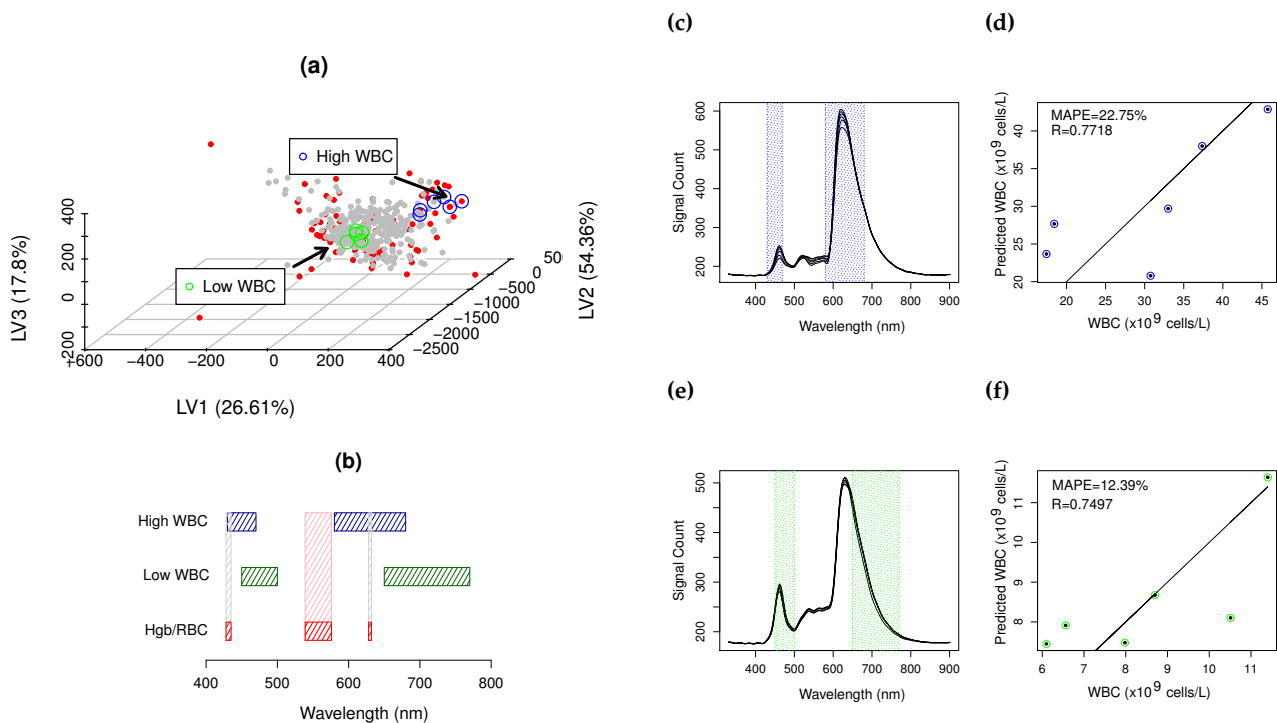


Figure 5. Representative low and high CovMs: (a) PLS scores space where (○) and (●) are low and high CovMs sample groups, ● SSD blood samples, ● RWD blood samples; (b) spectral ROIs used for the quantification of WBCs (■ low and ■ high CovM), (■) and (■) major and minor RBC/Hgb interference bands; (c,e) high and low CovM sample spectra and corresponding used ROIs to quantify WBCs; (d,f) prediction plot for high and low CovMs.

RBC/Hgb are dominant spectral features at low WBC values. WBCs absorb along all the UV-Vis spectra with more incidence in ROI 1 and 2, and RBC/Hgb information is mostly located in the ROI of 530–576 nm [1]. Low WBC CoMs use non-interferent RBC/Hgb spectral information, as presented in the interferogram of Figure 5b. A similar result is observed at high WBC levels, where minimal information overlapping occurs between WBCs and RBCs/Hgb (Figure 5b).

The apparent relationship between RBCs and WBCs, present in the global covariance (Figure 2), is not present at the CovM level, as it is an emergent property of the hemogram due to cell packing and scattering, only observable when the full knowledge-base is considered. At the CovM level, the interference mode is isolated, minimizing the major interferences, such as RBC/Hgb, by maximizing the WBC information into a single eigenvector.

Towards Information Specificity

The key principle of analytical methods is the use of specific information about a constituent, ensuring that unique information is used for its identification and quantification. Spectroscopy POC uses CovMs as a way to guarantee that constituents are quantified by specific interference or information. In a recent publication, we demonstrated that CovMs can exhibit orthogonal information to all other constituents, isolating specific information

on N, P and K in nutrient solutions for hydroponics [40]. A similar result is demonstrated here for WBCs, where CovMs exhibit independent properties to RBCs/Hgb.

AI and chemometric methods risk the use of natural correlation in data by not isolating the unique information of a constituent. Information isolation using CovMs has the following advantages:

- i. Using independent spectral information allows us to diagnose natural correlations, which are highly common in complex biological information. A natural correlation between WBCs and RBCs can lead to dataset specificity but not WBC specificity; that is, WBCs are inferred by the intrinsic correlation of WBCs and RBCs. The isolation of WBC information provides a cause–effect model, preserving both information equivalence and statistical significance;
- ii. The isolation of independent non-dominant information minimizes interference and maximizes WBC information. In this way, the sensitivity of WBC detection and quantification is increased (e.g., 85% at 7×10^9 cell/L, Figure 4c).

Spectroscopy POC works in the spectral information domain. It is possible to assess unique and specific information about an analyte at a local CovM level, once information similarity is verified. Furthermore, if there is enough representativeness in the local information, it becomes feasible to isolate the unique information, which is a step towards ensuring analytical specificity in spectroscopy POC for complex biological samples.

4. Conclusions

Despite the lower quantities of WBCs in relation to RBCs (1:1000), their spectral information is significant due to scattering and absorbance, accounting for 0.5 to 22.5% of spectral information; this results in the natural relationship with RBCs in hemogram and spectroscopy score spaces due to scattering and absorbance. Data augmentation by hybridization is very effective for reproducing real-world data. It further increases the level of detail at each CovM level, allowing for the unscrambling of specific information related to WBCs, with minimal interference from RBCs or Hgb, whereby the isolated CovMs exhibit high orthogonality towards RBCs. Isolating CovMs that contain only WBC information significantly increases the quantification sensitivity because the spectroscopy signal maximizes the WBC variance. This allows for achieving correlations of 0.7975 to 0.8397, a mean absolute error of 32.25% to 34.13% and a diagnostic efficiency between 83–100% inside the reference interval (5.5 to 19.5×10^9 cells/L) and 85.11% for cases with extremely high white blood cell counts. The obtained results are within the expected range of the existing state-of-the-art automated hematology analyzers in veterinary medicine.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/bios14010053/s1>: Figure S1: Spectral POC for Cat blood analysis: (1) withdrawal of 10 μ L of blood from ear lobe; (2) Injecting blood sample into capsule system; (3) plug-in the capsule into the POC probe; and (4) using IoT software to control POC measurement and record spectra for WBC calculation; Figure S2: Cat blood spectra: (a) real-world samples and (b) synthetic samples obtained by mixture of two samples spectra.

Author Contributions: Conceptualization, T.G.B., C.Q., F.M.-S., F.S. and R.C.M.; Methodology, T.G.B., F.S. and A.H.G.; Software, F.S.; Validation, T.G.B., F.M.-S. and A.H.G.; Investigation, F.S.; Data curation, T.G.B., C.Q., F.M.-S. and R.C.M.; Writing—original draft, R.C.M.; Visualization, C.Q.; Funding acquisition, R.C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the INESC TEC SEED project ‘MyVet—Reagent-less Point-of-Care for Veterinary Blood Analysis’. R.C.M. acknowledges their Fundação para a Ciência e Tecnologia (FCT) research contract grant (CEEIND/017801/2018). F.M.-S. acknowledges their Fundação para a Ciência e Tecnologia (FCT) PhD Thesis grant (SFRH/BD/09136/2020).

Acknowledgments: The authors acknowledge Anicura CHV—Veterinary Hospital Center, Porto, Portugal, for providing the necessary conditions for performing this research study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASVCP	American Society of Veterinary Clinical Pathology
BBL	Beer–Lambert Law
Bil	Bilirubin
BT	Blind Testing
CovM	Covariance Mode
CV	Cross-validation
EDTA	Ethylenediaminetetraacetic Acid
Hgb	Hemoglobin
HO	Hold out samples
HTC	Hematocrit
IoT	Internet of Things
LocPLS	Local Partial Least Squares
LV	Latent Variables
MAPE	Mean Absolute Percentage Error
PCV	Packed Cell Volume
PLS	Partial Least Squares
PLT	Platelets
POC	Point-of-care
R	Pearson Correlation
RBCs	Red Blood Cells
RI	Reference Interval
ROI	Region of Interest
RWD	Real World Data
SLAI	Self-learning Artificial Intelligence
SSD	Synthetic Spectroscopy Data
TAE	Total Allowable Error
TE	Total Error
TV	Target Values
UV-Vis	Ultraviolet–visible
Vis-SWNIR	Visible Short-Wave Infrared
WBCs	White Blood Cells

References

- Barroso, T.G.; Ribeiro, L.; Gregório, H.; Santos, F.; Martins, R.C. Point-of-care Vis-SWNIR spectroscopy towards reagent-less hemogram analysis. *Sens. Actuators B Chem.* **2021**, *343*, 130138. [\[CrossRef\]](#)
- Barroso, T.G.; Ribeiro, L.; Gregório, H.; Monteiro-Silva, F.; Neves dos Santos, F.; Martins, R.C. Point-of-Care Using Vis-NIR Spectroscopy for White Blood Cell Count Analysis. *Chemosensors* **2022**, *10*, 460. [\[CrossRef\]](#)
- Pasquini, C. Near infrared spectroscopy: A mature analytical technique with new perspectives—A review. *Anal. Chim. Acta* **2018**, *1026*, 8–36. [\[CrossRef\]](#) [\[PubMed\]](#)
- Olinger, J.M.; Griffiths, P.R. Quantitative effects of an absorbing matrix on near-infrared diffuse reflectance spectra. *Anal. Chem.* **1988**, *60*, 2427–2435. [\[CrossRef\]](#)
- Sparén, A.; Hartman, M.; Fransson, M.; Johansson, J.; Svensson, O. Matrix Effects in Quantitative Assessment of Pharmaceutical Tablets Using Transmission Raman and Near-Infrared (NIR) Spectroscopy. *Appl. Spectrosc.* **2015**, *69*, 580–589. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nishat, S.; Jafry, A.T.; Martinez, A.W.; Awan, F.R. Paper-based microfluidics: Simplified fabrication and assay methods. *Sens. Actuators B Chem.* **2021**, *336*, 129681. [\[CrossRef\]](#)
- Jiang, K.; Wu, J.; Qiu, Y.; Go, Y.Y.; Ban, K.; Park, H.J.; Lee, J.H. Plasmonic colorimetric PCR for Rapid molecular diagnostic assays. *Sens. Actuators B Chem.* **2021**, *337*, 129762. [\[CrossRef\]](#)
- Lewińska, I.; Speichert, M.; Granica, M.; Tymecki, Ł. Colorimetric point-of-care paper-based sensors for urinary creatinine with smartphone readout. *Sens. Actuators B Chem.* **2021**, *340*, 129915. [\[CrossRef\]](#)
- Zhou, L.; Vestri, A.; Marchesano, V.; Rippa, M.; Sagnelli, D.; Picazio, G.; Fusco, G.; Han, J.; Zhou, J.; Petti, L. The Label-Free Detection and Identification of SARS-CoV-2 Using Surface-Enhanced Raman Spectroscopy and Principal Component Analysis. *Biosensors* **2023**, *13*, 1014. [\[CrossRef\]](#)
- Popov, E.; Polishchuk, A.; Kovalev, A.; Vitkin, V. Raman Spectroscopy for Urea Breath Test. *Biosensors* **2023**, *13*, 609. [\[CrossRef\]](#)
- Monteiro-Silva, F.; Jorge, P.A.S.; Martins, R.C. Optical Sensing of Nitrogen, Phosphorus and Potassium: A Spectrophotometrical Approach toward Smart Nutrient Deployment. *Chemosensors* **2019**, *7*, 51. [\[CrossRef\]](#)

12. Martins, R.C.; Barroso, T.G.; Jorge, P.; Cunha, M.; Santos, F. Unscrambling spectral interference and matrix effects in Vitis vinifera Vis-NIR spectroscopy: Towards analytical grade ‘in vivo’ sugars and acids quantification. *Comput. Electron. Agric.* **2022**, *194*, 106710. [\[CrossRef\]](#)
13. Lin, K.; Cheng, D.L.P.; Huang, Z. Optical diagnosis of laryngeal cancer using high wavenumber Raman spectroscopy. *Biosens. Bioelectron.* **2012**, *35*, 213–217. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Coelho, B.F.O.; Nunes, S.L.P.; de França, C.A.; dos Santos Costa, D.; do Carmo, R.F.; Prates, R.M.; Filho, E.F.S.; Ramos, R.P. On the feasibility of Vis-NIR spectroscopy and machine learning for real time SARS-CoV-2 detection. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2024**, *308*, 123735. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Burns, D.; Ciurczak, E. *Handbook of Near Infrared Analysis*, 2nd ed.; Marcel Dekker, Inc.: New York, NY, USA, 2001.
16. Barroso, T.G.; Martins, R.C.; Fernandes, E.; Cardoso, S.; Rivas, J.; Freitas, P.P. Detection of BCG bacteria using a magnetoresistive biosensor: A step towards a fully electronic platform for tuberculosis point-of-care detection. *Biosens. Bioelectron.* **2018**, *100*, 259–265. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Tvedten, H.; Raskin, R.E. 4—Leukocyte Disorders. In *Small Animal Clinical Diagnosis by Laboratory Methods*, 5th ed.; Willard, M.D., Tvedten, H., Eds.; W.B. Saunders: Saint Louis, MO, USA, 2012; pp. 63–91. [\[CrossRef\]](#)
18. Burton, A.G.; Jandrey, K.E. Leukocytosis and Leukopenia. In *Textbook of Small Animal Emergency Medicine*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2018; Chapter 64, pp. 405–412. [\[CrossRef\]](#)
19. Philo, J.; Adams, M.; Schuster, T.M. Association-dependent absorption spectra of oxyhemoglobin A and its subunits. *J. Biol. Chem.* **1981**, *256*, 7917–7924. [\[CrossRef\]](#)
20. Tan, H.; Liao, S.; Pan, T.; Zhang, J.; Chen, J. Rapid and simultaneous analysis of direct and indirect bilirubin indicators in serum through reagent-free visible-near-infrared spectroscopy combined with chemometrics. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **2020**, *233*, 118215. [\[CrossRef\]](#)
21. Chabot-Richards, D.S.; George, T.I. White Blood Cell Counts: Reference Methodology. *Clin. Lab. Med.* **2015**, *35*, 11–24. [\[CrossRef\]](#)
22. Rishniw, M.; Pion, P.D. Evaluation of performance of veterinary in-clinic hematology analyzers. *Vet. Clin. Pathol.* **2016**, *45*, 604–614. [\[CrossRef\]](#)
23. Riond, B.; Waßmuth, A.K.; Hartnack, S.; Hofmann-Lehmann, R.; Lutz, H. Study on the kinetics and influence of feline platelet aggregation and deaggregation. *BMC Vet. Res.* **2015**, *11*, 276. [\[CrossRef\]](#)
24. Gulati, G.; Uppal, G.; Gong, J. Unreliable Automated Complete Blood Count Results: Causes, Recognition, and Resolution. *Ann. Lab. Med.* **2022**, *42*, 515–530. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Nabity, M.B.; Harr, K.E.; Camus, M.S.; Flatland, B.; Vap, L. ASVCP guidelines: Allowable total error hematology. *Vet. Clin. Pathol.* **2018**, *47*, 9–21. [\[CrossRef\]](#)
26. Harvey, J.W. The feline blood film: 2. Leukocyte and platelet morphology. *J. Feline Med. Surg.* **2017**, *19*, 747–757. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Meredith, A.M.; Bienzle, D.; Beeler-Marfisi, J. What is your diagnosis? A feline blood smear. *Vet. Clin. Pathol.* **2019**, *48*, 128–130. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Terent'yeva, Y.G.; Yashchuk, V.M.; Zaika, L.A.; Snitserova, O.M.; Losytsky, M.Y. The manifestation of optical centers in UV-Vis absorption and luminescence spectra of white blood human cells. *Methods Appl. Fluoresc.* **2016**, *4*, 044010. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Ramesh, J.; Kapelushnik, J.; Mordehai, J.; Moser, A.; Huleihel, M.; Erukhimovitch, V.; Levi, C.; Mordechai, S. Novel methodology for the follow-up of acute lymphoblastic leukemia using FTIR microspectroscopy. *J. Biochem. Biophys. Methods* **2002**, *51*, 251–261. [\[CrossRef\]](#)
30. Ramesh, J.; Huleihel, M.; Mordehai, J.; Moser, A.; Erukhimovich, V.; Levi, C.; Kapelushnik, J.; Mordechai, S. Preliminary results of evaluation of progress in chemotherapy for childhood leukemia patients employing Fourier-transform infrared microspectroscopy and cluster analysis. *J. Lab. Clin. Med.* **2003**, *141*, 385–394. [\[CrossRef\]](#)
31. Liu, W.; Howarth, M.; Greytak, A.B.; Zheng, Y.; Nocera, D.G.; Ting, A.Y.; Bawendi, M.G. Compact biocompatible quantum dots functionalized for cellular imaging. *J. Am. Chem. Soc.* **2007**, *130*, 1274–1284. [\[CrossRef\]](#)
32. Managò, S.; Valente, C.; Mirabelli, P.; Circolo, D.; Basile, F.; Corda, D.; De Luca, A.C. A reliable Raman-spectroscopy-based approach for diagnosis, classification and follow-up of B-cell acute lymphoblastic leukemia. *Sci. Rep.* **2016**, *6*, 24821. [\[CrossRef\]](#)
33. Lima, A.M.F.; Silva, J.D.P.d.; Daniel, C.R. Diagnosis, classification and monitoring of leukemia based on Raman spectroscopy. *Res. Soc. Dev.* **2021**, *10*, e67101421657. [\[CrossRef\]](#)
34. Sahu, R.K.; Zelig, U.; Huleihel, M.; Brosh, N.; Talyshinsky, M.; Ben-Harosh, M.; Mordechai, S.; Kapelushnik, J. Continuous monitoring of WBC (biochemistry) in an adult leukemia patient using advanced FTIR-spectroscopy. *Leuk. Res.* **2006**, *30*, 687–693. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Zelig, U.; Mordechai, S.; Shubinsky, G.; Sahu, R.K.; Huleihel, M.; Leibovitz, E.; Nathan, I.; Kapelushnik, J. Pre-screening and follow-up of childhood acute leukemia using biochemical infrared analysis of peripheral blood mononuclear cells. *Biochim. Biophys. Acta (BBA)-Gen. Subj.* **2011**, *1810*, 827–835. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Chaber, R.; Kowal, A.; Jakubczyk, P.; Arthur, C.; Łach, K.; Wojnarowska-Nowak, R.; Kusz, K.; Zawlik, I.; Paszek, S.; Cebulski, J. A Preliminary Study of FTIR Spectroscopy as a Potential Non-Invasive Screening Tool for Pediatric Precursor B Lymphoblastic Leukemia. *Molecules* **2021**, *26*, 1174. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Kochan, K.; Bedolla, D.E.; Perez-Guaita, D.; Adegoke, J.A.; Veettil, T.C.P.; Martin, M.; Roy, S.; Pebotuwa, S.; Heraud, P.; Wood, B.R. Infrared Spectroscopy of Blood. *Appl. Spectrosc.* **2021**, *75*, 611–646. [\[CrossRef\]](#) [\[PubMed\]](#)

38. Agbaria, A.H.; Beck, G.; Lapidot, I.; Rich, D.H.; Kapelushnik, J.; Mordechai, S.; Salman, A.; Huleihel, M. Diagnosis of inaccessible infections using infrared microscopy of white blood cells and machine learning algorithms. *Analyst* **2020**, *145*, 6955–6967. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Javinsky, E. Chapter 25—Hematology and Immune-Related Disorders. In *The Cat*; Little, S.E., Ed.; W.B. Saunders: Saint Louis, MO, USA, 2012; pp. 643–703. [\[CrossRef\]](#)
40. Silva, F.; Queirós, C.; Pinho, T.; Boaventura, J.; Santos, F.; Barroso, T.; Pereira, M.; Cunha, M.; Martins, R. Reagent-less spectroscopy towards NPK sensing for hydroponics nutrient solutions. *Sens. Actuators B Chem.* **2023**, *395*, 134442. [\[CrossRef\]](#)
41. Nalepa, J.; Marcinkiewicz, M.; Kawulok, M. Data Augmentation for Brain-Tumor Segmentation: A Review. *Front. Comput. Neurosci.* **2019**, *13*. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Shorten, C.; Khoshgoftaar, T. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [\[CrossRef\]](#)
43. Brown, M.; Wittwer, C. Flow Cytometry: Principles and Clinical Applications in Hematology. *Clin. Chem.* **2000**, *46*, 1221–1229. [\[CrossRef\]](#)
44. INESC TEC. *AgIoT—Modular Solution and Open Source IoT Solution for Agrofood Domain*; INESC TEC: Porto, Portugal, 2020.
45. Martens, H.; Stark, E. Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy. *J. Pharm. Biomed. Anal.* **1991**, *9*, 625–635. [\[CrossRef\]](#)
46. Geladi, P.; Kowalsky, B. Partial least squares regression: A tutorial. *Anal. Acta* **1988**, *185*, 1–17. [\[CrossRef\]](#)
47. Ergon, R. Re-interpretation of NIPALS results solves PLSR inconsistency problem. *J. Chemom.* **2009**, *23*, 72–75. [\[CrossRef\]](#)
48. Phatak, A.; Jong, S. The geometry of partial least squares. *J. Chemom.* **1997**, *11*, 311–338. [\[CrossRef\]](#)
49. Shen, G.; Lesnoff, M.; Baeten, V.; Dardenne, P.; Davrieux, F.; Ceballos, H.; Belalcazar, J.; Dufour, D.; Yang, Z.; Han, L.; Fernández Pierna, J.A. Local partial least squares based on global PLS scores. *J. Chemom.* **2019**, *33*, e3117. [\[CrossRef\]](#)
50. R-Core-Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: <http://www.r-project.org/> (accessed on 19 December 2023).
51. Harr, K.E.; Flatland, B.; Nabity, M.B.; Freeman, K.P. *ASVCP Guidelines—Allowable Total Error: Biochemistry*; American Society for Veterinary Clinical Pathology: Madison, WI, USA, 2013.
52. Meng, F.; Alayash, A.I. Determination of extinction coefficients of human hemoglobin in various redox states. *Anal. Biochem.* **2017**, *521*, 11–19. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Cook, A.M.; Moritz, A.; Freeman, K.P.; Bauer, N. Quality requirements for veterinary hematology analyzers in small animals—A survey about veterinary experts’ requirements and objective evaluation of analyzer performance based on a meta-analysis of method validation studies: Bench top hematology analyzer. *Vet. Clin. Pathol.* **2016**, *45*, 466–476. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.