A COMPARATIVE STUDY OF SEVEN IN-HOUSE AND TWO LABORATORY HEMATOLOGY INSTRUMENTS

MARTINA BECKER

INAUGURAL-DISSERTATION
zur Erlangung des Grades eines
Dr. med. vet.
beim Fachbereich Veterinärmedizin
der Justus-Liebig-Universität Gießen

édition scientifique
VVB LAUFERSWEILER VERLAG

1. Auflage 2008

1st Edition 2008

Printed in Germany

VVB LAUFERSWEILER VERLAG
édition scientifique

STAUFENBERGRING 15, D-35396 GIESSEN
Tel: 0641-5599888 Fax: 0641-5599890
email: redaktion@doktorverlag.de

www.doktorverlag.de

Aus der Klinik für Kleintiere, klinische Pathophysiologie und klinische Laboratoriumsdiagnostik

Betreuer: Prof. Dr. Andreas Moritz

A Comparative Study of Seven In‑House and Two Laboratory Hematology Instruments

INAUGURAL-ISSERTATION
zur Erlangung des Grades eines
Dr. med. vet.
beim Fachbereich Veterinärmedizin
der Justus-Liebig-Universität Gießen

eingereicht von

Martina Becker

Tierärztin aus Madrid/Spanien

Gießen, 2007

Mit Genehmigung des Fachbereichs Veterinärmedizin
der Justus-Liebig-Universität Gießen.

Dekan: Prof. Dr. Dr. G. Baljer

Gutachter:

Prof. Dr. A. Moritz

Prof. Dr. A. Wehrend

Tag der Disputation: 21.12.2007

Content

1. Introduction and Literature Overview

The evaluation of blood has become an integral part in the routine clinical assessment of healthy and diseased companion animals. The packed cell volume (PCV), or as in humans the total hemoglobin concentration, and the microscopic examination of a blood smear have been used since the beginning of veterinary care and still form the cornerstones in clinical hematology. While manual counting of the various forms of blood cells by microscopic counting chambers has been practiced, it was found to be tedious work and the cell counts are not very accurate. The first automated blood cell counter was introduced by Coulter a half century ago. Since then they have evolved into advanced hematology instruments utilizing impedance and laser flow cytometry. While in the past these technologies were limited for application in large clinical pathology laboratories, several smaller and affordable automated hematology analyzers have recently been developed for in-clinic usage. Their actual means of measurements and analyses vary and the practical usefulness and test result accuracy of these different point-of-care and laboratory instruments have not been comprehensively compared to each other and against reference methods. There are several reports from the university and industry comparing one laboratory instrument with one in-clinic instrument. For instance the QBC VetAutoread (Idexx Laboratories) was first introduced and is still a commonly used hematology analyzer, but clinical studies revealed some limitations compared to the reference laboratory methods (1-6). Bienzle et al. (1) precluded reliance on differential cell counts, due to the inability to identify abnormal cells, the high percentage of error messages and the wide confidence intervals. There are some actual studies available, comparing other in-clinic hematology analyzers to reference laboratory methods (7-14, 50). Dewhurst et al. (12) assessed the accuracy of the VetScan HMT (Abaxis) in comparison to the Cell-Dyn 3700 (Abbott Laboratories) and manual methods, by interpreting results of Mann-Whitney U test, Bland-Altman difference plots and Deming regression. The instrument performed well on canine and feline

samples. Except the platelet counts in feline samples should be interpreted with caution, as they can be unreliable. Papasouliotis et al. (13) compared the white blood cell differential in percentages determined by the LaserCyte hematology analyzer (Idexx Laboratories) to a 100-cell manual differential. Statistical analysis included Wilcoxon signed rank test, Deming regression and Bland-Altman difference plots. Good correlation was only achieved for feline eosinophils. The authors remark that the reference method might be a limitation of their study and that more studies are needed to assess the clinical significance of the obtained differences. Although Bland-Altman analysis is performed beside correlation analysis in newer studies, assessment of the accuracy is mostly based on correlation coefficients and a previously published classification scheme (1, 4, 12 and 45). But according to different authors and journal editorial boards, the assessment of methods, based on correlation coefficients is not appropriate (15-17). Therefore the following chapter will highlight some important considerations regarding method validation.

1.1. Method validation

Central to current laboratory quality theory is definition of medical decision limits, clinical quality requirements and acceptable levels of total allowable error. The ISO 15189 recommends that quality requirements for each test within the laboratory should be established (18). The definition of medical decision limits is based on the clinical interpretation of various levels of an analyte. Databases recommending clinical decision limits for human laboratory tests are available (19), but there are no comprehensive summaries available for veterinary testing. However common usage and interpretative guidelines for veterinary medical testing can be used to determine one or more levels of medical significance for laboratory analytes. The clinical quality requirements at each medical decision limit are also based on the way that clinical veterinary medical laboratory data is interpreted based on empirical usage and within the literature (20). If actual error observed in a test exceeds the clinical quality requirement than the method or instrument may not be suitable for use or clinical quality requirements and interpretation may need to be modified in order to reflect the capability of the instrument/method.

In order to determine if an instrument truly performs according to the specifications, described by the manufacturer and to determine if the instrument/method has performance capability that meets the needs of a veterinary testing facility, based on the medical decision limit and clinical quality requirements used for veterinary patients in that facility a series of experiments should be conducted as part of the instrument/method validation process.

The meaning of method validation is error assessment (15). That means method validation is used to estimate how much error might be in a test result produced by the method. Within the method validation process different experiments are used to assess the different types of errors. A random error is an error which can be either positive or negative and whose direction and magnitude cannot be predicted (15). Random error is assessed by replication experiments. A systematic error is an error that is always in one direction. If the systematic error stays the same over a range of concentrations, it is called constant systematic error. If it changes as concentration changes it is called proportional systematic error (15). Constant systematic error is assessed by interference studies, proportional systematic error by recovery experiments. Furthermore both systematic errors can be estimated by method comparison experiments. Finally the total error is the combined effect of random and systematic error, and can be estimated by the bias from method comparison plus three times the standard deviation from replication experiment (15). The total error is used for the judgment of acceptability of a method. In human medicine several regulatory requirements for method validation are available, defining total errors (21-27). These kinds of requirements are lacking in veterinary medicine. As no recommendations for veterinary medicine are available right now, the assessment should be based on clinical quality requirements and data within the literature. Each clinical pathologist might have a different opinion about clinical requirements. Hence, the goal should be to define some recommendations for veterinary medicine in the future. The following table shows suggestions for clinical requirements published by Freeman et al. (1999) for the most important hematological parameters (20). The CLIA (Clinical Laboratory Improvement Amendments) requirements for human medicine are included (24).

Parameter	Unit	Total allowable errors (%)	
		Freeman et al.	CLIA
WBC	10^9/l	20	15
RBC	10^{12}/l	10	6
Hb	g/dl	7	7
HTC	%	10	6
PLT	10^9/l	n/a	25

Table 1: Recommendations for clinical requirements for veterinary hematology, according to Freeman et al. (20) and clinical requirements for human medicine, according to CLIA (24)

It needs to be considered that for different concentration levels different clinical requirements might be necessary. In the different recommendations only one total allowable error is defined for the whole concentration range.

1.1.1. Linearity experiment

The linearity experiment assesses the useful analytical range of a method, i.e. the lowest and highest test results that are reliable. The National Committee for Clinical Laboratory Standards (NCCLS) recommends a minimum of four, preferably five different levels (15). NCCLS recommends analyzing each level four times; Westgard indicates that three replicates are generally sufficient (15). The means of the measured results are plotted on the y-axis; the expected results are plotted on the x-axis. The reportable range is then assessed by linear regression statistics or by manually drawing the best straight line that fits the lowest point in the series. Furthermore, errors can be quantified by comparing the observed means with the theoretically expected values (15).

1.1.2. Replication experiment

The replication experiment is performed to estimate the imprecision or the random error of a method. The imprecision of a method greatly depends on the time period the experiment is performed. For example, when an experiment is performed within an analytical run the effects due to day-to-day changes in operating conditions will not be observed. This kind of imprecision is known as »within-run« imprecision.

Whereas an experiment conducted over several days and several runs is known as »between-run« imprecision or total imprecision. This estimate includes the long term variations and is normally a more realistic estimate (15).
The number of materials and concentrations to be tested depends on the clinical decision levels of the test. A minimum of two or three different levels is normally required (15). Each material should be analyzed at least 20 times (15). Mean, standard deviation (SD) and coefficient of variation (CV) should be calculated. The mean describes the central location of the measurements; the SD describes the expected distribution of the results (66% of the measurements are expected to be within plus/minus 1SD; 95% within plus/minus 2SD and 99.7% within plus/minus 3SD (if the data follows a bell-shaped Gaussian distribution)). The CV equals the SD divided by the mean and expressed in percentage. The judgment on acceptability depends on the total allowable error: for within-run imprecision SD should be less than one quarter of the total allowable error, for between-run imprecision SD should be less than one third of the total allowable error (15).

1.2.3. Method comparison

A comparison of method is performed to assess inaccuracy or systematic error. The comparative method must be carefully selected and should ideally be a reference method (15). The International Council for Standardization in Hematology (ICSH) defines accuracy as 'a measure of agreement between the estimate of value and the true value' (28). The true value must be obtained by a reference method and the only hematological parameters that can be estimated accurately are hemoglobin concentration, PCV, red and white cell counts and the differential leukocyte count (28). Alternatively the comparability can be assessed. This is defined as the ability of the instrument to produce results which agree satisfactorily with those obtained by procedures in routine use (28). The comparability is an accepted method to evaluate hematology instruments (28). But in case of differences between the methods, it is necessary to identify which method is inaccurate (15).
Concerning the number of patient specimens the quality of the experiment depends more on the concentration range than on the actual number of analyzed specimens. The samples should cover the whole reportable range (15).
Several different statistical and interpretative approaches have been used in the past to assess the results of method comparison. The ordinary linear regression is used to

predict values for a response variable (Y) based on one independent variable (X). The relationship between the variables must be linear. The linear relation is then expressed as y=a+bx, whereas least-squares estimation method is used to minimize the vertical distance between points and the fitted regression line. Therefore imprecision only occurs in the Y variable. The correlation coefficient r, the slope and intercept with 95% confidence limits and the standard deviation of the residuals $S_{y|x}$ are calculated. The correlation coefficient r has traditionally been used to judge the acceptability of method performance, but the correlation coefficient should more properly be used to judge the acceptability of the concentration range of the data being used to calculate the regression statistics (15). The following example underlines this statement. Figure 1 shows the regression analysis for MCV in canine samples, comparing a newly introduced hematology analyzer to a reference method. In figure 2 the comparison for the same instruments is shown for feline samples. Finally figure 3 includes canine and feline samples.

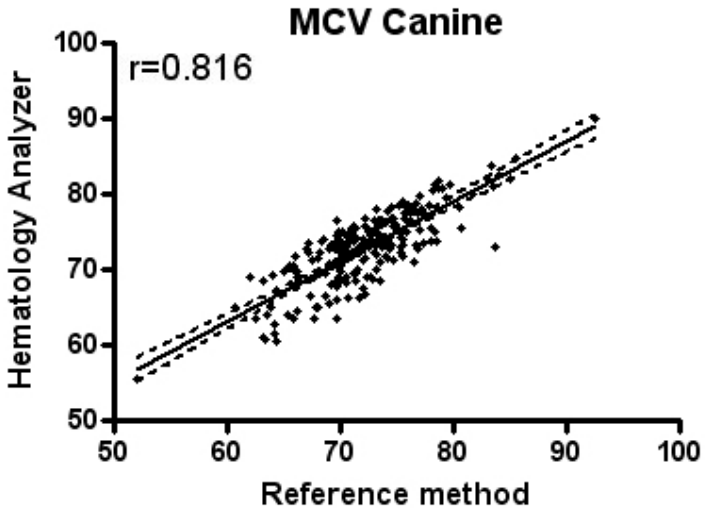


Figure 1: Linear regression for MCV in canine samples

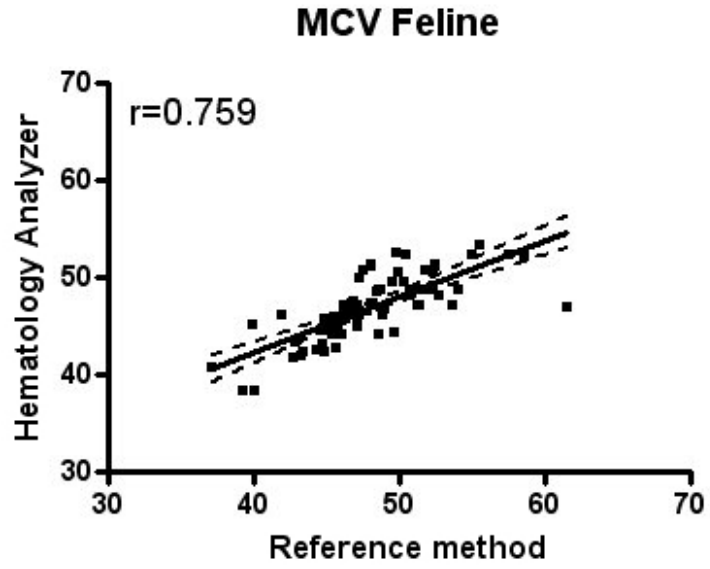


Figure 2: Linear regression for MCV in feline samples

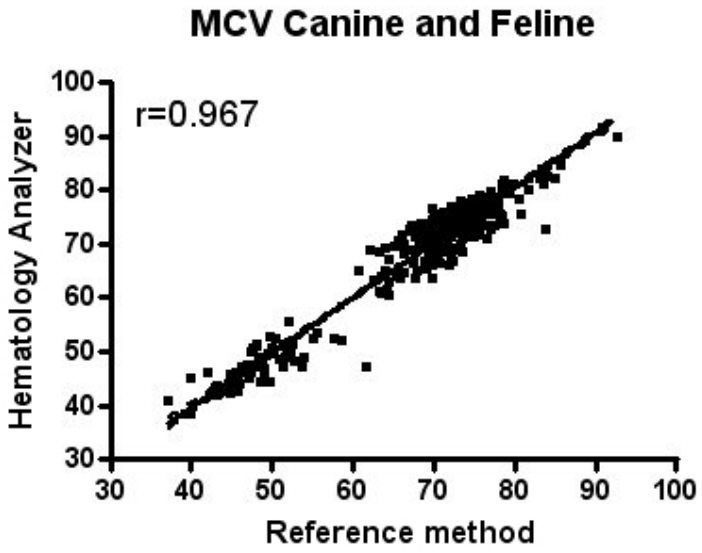


Figure 3: Linear regression for MCV in canine and feline samples

Including canine and feline samples in the statistics clearly elevates the correlation coefficient to a value of 0.967, just because of the wider concentration range. Therefore if the correlation coefficient is less than 0.975, ordinary linear regression may not be reliable and data improvement or alternate statistics are appropriate (15; 29). Alternate statistics are Deming regression, Passing & Bablok regression and Bland-Altman plots (see below).

With regression analysis the calculated slope reflects the proportional systematic error, whereas the intercept reflects the constant systematic error and $S_{y|x}$ the random error. $S_{y|x}$ is influenced by two effects, namely by the total analytical imprecision ($S_{a,tot}$) and by sample-related effects (29). When $S_{y|x} >> S_{a,tot}$ sample-related effects are present.
In contrast to ordinary linear regression, Deming regression allows imprecision in both methods (X and Y). Therefore it is the better choice when the comparable method is not a true reference method, and some kind of imprecision is also expected in this method. Like ordinary linear regression, slope and intercept with 95% confidence limits and $S_{y|x}$-values are calculated. However the comparison of $S_{y|x}$ and $S_{a,tot}$ can only be performed with $S_{y|x}$-values obtained from ordinary linear regression (29). »Passing & Bablok« regression also allows imprecision in both methods (X and Y), but the imprecision need not have constant variance across the sampling range (necessary for ordinary linear and Deming regression).
A major disadvantage of »Passing & Bablok« regression is the absence of calculation of the $S_{y|x}$-value (29).

A Bland-Altman plot is created by plotting for each sample the difference between the two measurements as a function of the average of the two measurements. The presented bias is computed as the value determined by one method minus the value determined by the other method, and describes the systematic error. If one method is sometimes higher and sometimes lower, the bias will be close to zero. If it is not close to zero, this indicates that the methods are producing different results. Furthermore, 95% limits of agreement, which basically describe the precision or random error, can be calculated. If a relationship between the difference and the average is visible (any kind of trend), then the 95% limits of agreement are not appropriate since the observed range of differences depends on the average value (16). By assessing the relationship between the difference and average a statement can be made if a bias is mainly caused by constant or proportional systematic error.

A paired t-test can be used to test whether a difference between the means exists or not. The t-value is a ratio of the bias and the standard deviation of the differences (S_d). Bias represents the systematic error and S_d the random error. That means t expresses the magnitude of the systematic error in multiples of random error (15). Statistic tables provide critical t-values, depending on selected probabilities and degrees of freedom. If the observed t-value is greater than the critical t-value, found in the table, there is a difference between the two means. This means a systematic error is present. But the t-value itself says nothing about the amount of error and therefore about the acceptability of a method performance. The only conclusion which can be drawn from a t-value is whether a systematic error is present or not (15).

Summarizing the descriptions above the following statistics should be performed to calculate the different types of errors, including the total allowable error: The regression analysis is used to calculate r-values and $S_{y|x}$-values. Deming regression should be performed to get appropriate estimates for intercept and slope. Finally Bland-Altman plots are generated and the calculated bias is used to verify systematic error. In case of high r-values ($r \geq 0.975$) the type of systematic error can be characterized based on intercept and slope. In case of low r-values ($r < 0.975$), the estimates might not be accurate, and the characterization should be based on the plot. Total error can then be calculated in the following way: bias of the method comparison experiment plus three times standard deviation of the replication study (15). Bias represents the systematic error and standard deviation the random error. As indicated above $S_{y|x}$, as well as 95% limits of agreement of the bias also delineate the random error. But in contrast to the standard deviation of the replication experiment, these values are calculated from all data of the method comparison study which should ideally include very high and very low values. The comparison of $S_{y|x}$ and $S_{a,tot}$ is used to verify if there are sample-related effects present. Based on the total error a judgment of the method performance can be made.

1.2.4. Judgment of method performance

The decision on acceptability depends on the size of the observed errors relative to the total allowable error. The total error of the evaluated method is calculated from the bias of the method comparison experiment and the standard deviation obtained

in the replication experiment. Different recommendations about calculation are available (15):

- Bias plus 2 times SD < TE_a
- Bias plus 3 times SD < TE_a
- Bias plus 4 times SD < TE_a

All three calculations can be utilized in a graphical decision tool, so called Method Decision Chart (30). They are constructed as follows:

1. On the y-axis, the allowable inaccuracy (bias, %) is plotted. The scale should go from 0 to the total allowable error (TE_a)
2. On the x-axis, the allowable imprecision (CV, %) is plotted. The scale should go from 0 to 0.5xTE_a
3. A line for bias+2SD is drawn from TE_a on the y-axis to 0.5TE_a on the x-axis
4. A line for bias+3SD is drawn from TE_a on the y-axis to 0.33TE_a on the x-axis
5. A line for bias+4SD is drawn from TE_a on the y-axis to 0.25TE_a on the x-axis
6. Label the region beneath the line for bias+4SD as excellent performance, between bias+4SD and bias+3SD as good, between bias+3SD and bias+2SD as marginal and above bias+2SD as poor performance.

Poor performance indicates that the method is not acceptable for routine operation. A method with marginal performance provides the desired quality when everything is working correctly. However a total quality control strategy will be necessary to manage the method in routine operation. This total quality control strategy might include well-trained operators, expensive statistical quality control, aggressive preventive maintenance and continual efforts to improve method performance. A method with good performance can be well-managed in routine service with carefully planned statistical quality control. Finally, a method with excellent performance is easy to manage and can be controlled with minimal expense (15). Here is an example how to create a Method Decision Chart (figure 4):

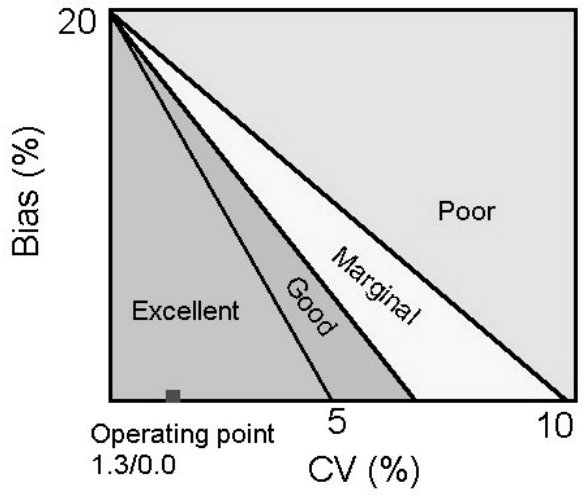


Figure 4: Example for a Method Decision Chart (WBC)
Total allowable error: 20%
Observed bias: 0.0%
Observed CV: 1.3%

The purpose of this study is the evaluation of the performance characteristics of seven different in-house and two laboratory hematology instruments, in respect to accuracy, precision, linearity, carry-over and ease of use. Comparing nine different instruments in one study is unique and allows objective comparison of the most important hematology instruments currently on the market. Furthermore, to my knowledge this is the first study in veterinary hematology utilizing the calculation of total errors for the assessment of instrument accuracy.

2. Material and Method

2.1. Sample specifications

From December 2004 to February 2005 2 ml EDTA-anticoagulant blood samples were collected for routine in-house blood testing from healthy (spay or castration presurgical examination (> 6 months old)) and diseased adult dogs and cats presented to 3 small animal clinics in the Fort Lauderdale area in Southern Florida. Two freshly prepared blood smears from each sample and left-over blood samples were stored at room temperature until transported by courier to the nearby laboratory[1] for comparative analyses.

All in-clinic hematology analyzers to be assessed in this study were set up with a software version for small animals in an air-conditioned laboratory at CDS. The Cell-Dyn 3500 (A) was available for immediate comparison while an aliquot blood sample was send to IDEXX laboratories overnight with next day morning delivery for analysis by the ADVIA™ 120 (B). Technical details and cell counting abilities of each hematology instrument are summarized in chapter 2.2 and table 2.

Calibrations as well as all maintenance procedures were performed according to the manufacturer's instructions. Control samples were run daily on those systems for which manufacturers provided control materials. In addition, the PCV was determined using the Microhaematocrit Centrifuge (StatSpin VT)[2].

All samples were analyzed within 3-6 hours after blood collection except the analysis on the ADVIA™ 120 was performed in a reference laboratory within 24 hours.

---

1 Clinical Diagnostic Solutions, Inc. 1800 NW 65th Ave, Plantation FL, 33313, USA
2 StatSpin VT Centrifuge, Statspin ®, Inc.

Prior to analysis tubes were placed on a roller plate (Coulter Mixer) for at least one minute to assure proper mixing.

The instrument sequence for each sample was as follows: LaserCyte® (C), Cell-Dyn 3500 (A), MS45 (E), Heska CBC-Diff (F), VetScan HMT (H), ForCyte (D), Scil Vet ABC (G), QBC® Vet Autoread™ (K). Every blood sample was then analyzed within minutes once on each system, except when there was insufficient sample volume or one of the instruments was not available. Sample volume permitting, repeat testing was performed when certain error messages were seen. The error messages were classified into two groups: all flags caused by a technical problem of the instruments or where the error message suggested to repeat the sample were classified as technical flags; all flags caused by abnormalities (e.g. morphological changes) of the blood samples were classified as morphological flags. When a technical flag occurred, the sample was analyzed again. If in the second measurement a technical flag appeared again, the results were rejected (see figure 5).

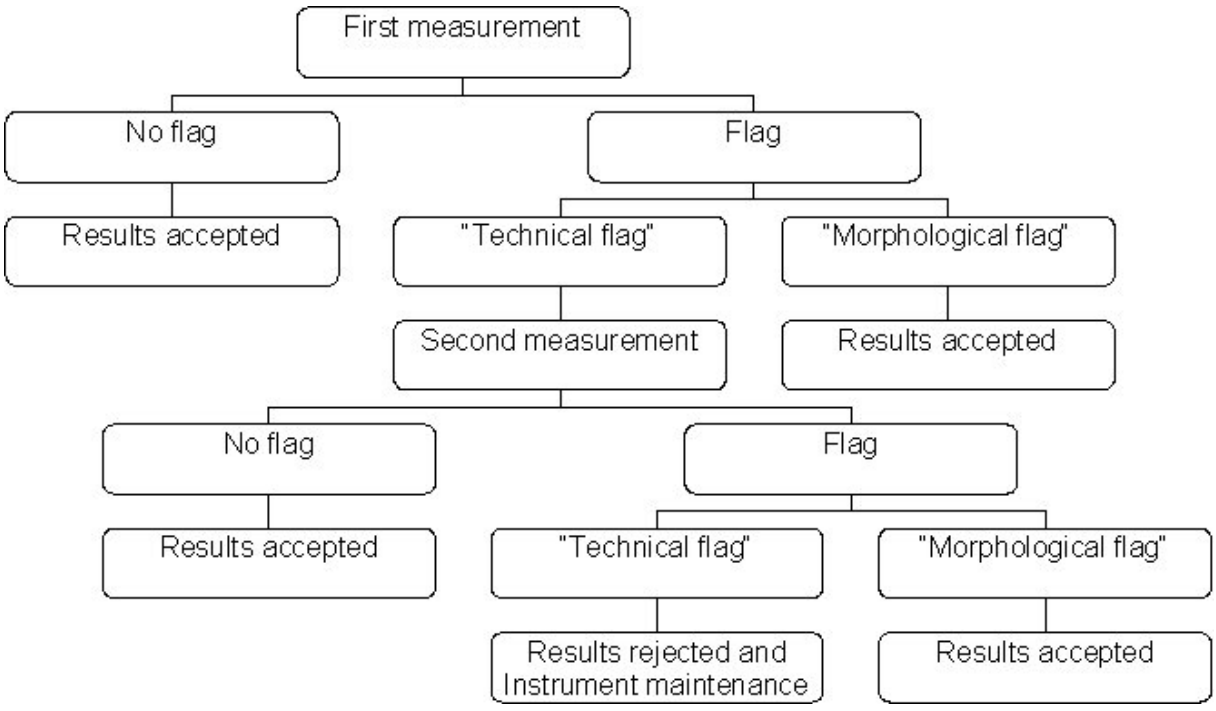


Figure 5: Algorithm for handling samples flagged by the hematology instruments

The blood smears were stained with a modified Wright Stain, according to manufacturer recommendations. In all thrombocytopenic cases an estimation of the platelet count on the blood smear was performed. If platelets were clumped,

»clumped platelets« were reported. If platelets were well distributed the number of platelets in ten 1000x fields were counted, the mean determined and multiplied with 20 (31). In case of 50,000/µl platelets or less the thrombocytopenia was classified as severe.

Two-hundred-cell-manual-differentials were performed and morphological changes reported. The absolute values for all cell types were calculated separately for each instrument by using the white blood cell counts of the accordant instrument. To determine values for granulocytes the absolute numbers of neutrophils, band neutrophils, eosinophils and basophils were added, for the manual differential, as well as for the instruments performing a 5-part-differential.
Neutropenia was classified as a neutrophil count of less than $2.5x10^9$/l for canine and feline samples, neutrophilia as a neutrophil count above $20.0x10^9$/l and an eosinophilia as an eosinophil count above $1.5x10^9$/l for canine respectively $1.7x10^9$/l for feline samples. Less than $1.0x10^9$/l lymphocytes in canine samples, and less than $1.5x10^9$/l lymphocytes in feline samples, was classified as lymphopenia (personal communication Urs Giger, Andreas Moritz, Dennis DeNicola, and Martina Becker).

For the precision, linearity and carry-over experiments 700ml of K-EDTA anticoagulant blood from two healthy dogs were collected and pooled. During all analyses the blood samples were stored at room temperature.
For the precision study a tube with 2ml of the pooled blood was prepared for each instrument. To eliminate the effects of cell-ageing the samples were run on all systems simultaneously.
The remaining blood volume was divided into 50ml tubes and centrifuged at room temperature at 500xG for 20 minutes. After centrifugation the plasma layers were removed and pooled into one tube, labeled as level 1. Afterwards the Buffy Coats were removed and pooled into another tube, labeled as level 5.
For the linearity study three additive levels were prepared by diluting level 5 with level 1 (level 2: 25% of level 5, level 3: 50% of level 5, level 4: 75% of level 5). Each level was analyzed two times followed by one analysis of level five. The order of the different levels was randomized.
For the carry-over experiment level 5 was analyzed twice on each instrument followed by three analyses of PBS (extra tube for each analysis). This sequence was repeated three times. As instrument H is not able to measure PBS the previously prepared level two was used instead of PBS.

Systems	Technology	Parameter			Sample volume	Time	Control material	Software Version
		CBC	Differential	Absolute Reticulocyte count				
In-house hematology instruments								
C[3]	Laser	Yes	5-part	Yes (Ca, Fe)	95µl	12min	Latex beats	1.37/1.1.87
D[4]	Laser and Impedance	Yes	5-part	Yes (Ca)	20µl	2min	-	-
E[5]	Impedance	Yes	5-part	No	16µl	~3min	-	5,04A
F[6]	Impedance	Yes	3-part	No	125µl	73sec	Vet-Con	V3.82s-r
G[7]	Impedance	Yes	3-part	No	12µl	~80sec	Minotrol-16™	-
H[8]	Impedance	Yes	3-part	No	10-20µl	~2min	Abaxis controls	3,51U
K[9]	Buffy Coat	Yes (no RBC)	2-part	No	111µl	~7min	Calibration Rod	4,3
Laboratory instruments								
A[10]	Laser and Impedance	Yes	5-part	No	~130µl	37sec	Para 12 Plus	Rev.H Ver 4.3
B[11]	Laser and Peroxidase	Yes	5-part	Yes (Ca, Fe)	157µl	50sec	Bayer Advia 120 controls	V2.2.2A

Table 2: Instrument specifications

3 LaserCyte®, IDEXX Laboratories, Inc., One Idexx Drive, Westbrook, Maine 04092, USA, www.idexx.com
4 ForCyte, Oxford Science Inc., One American Way, 178 Christian Street, Oxford, CT 06478, USA, www.oxfordscienceinc.com
5 MS45, Melet Schloesing Pharmaceuticals s.a., Rue du college 90, 2300 La Chaux de fonds, Switzerland, www.mslabos.com
6 Heska CBC, Boule Medicalab, Box 42065, 12613 Stockholm, Sweden, www.heska.com
7 Scil Vet ABC, Scil animal care company, Dina-Weissmann-Allee 6, 68519 Viernheim, Germany, www.scilvet.com
8 VetScan HMT, Abaxis, 1320 Chesapeake Terrace, Sunnyvale CA 94089, USA, www.abaxis.com
9 QBC® Vet Autoread TM, IDEXX Laboratories, Inc., One Idexx Drive, Westbrook, Maine 04092, USA, www.idexx.com
10 Cell-Dyn 3500, Abbott Laboratories, Abbott Park, Illinois, USA, www.abbottdiagnostics.com
11 ADVIA™ 120, Siemens Medical Solutions, Siemensst. 3, 35463 Fernwald , Germany, www.siememsmedical.de

2.2 Evaluated instruments

In the following chapter each evaluated instrument will be described shortly, including the displayed error messages.

2.2.1. Instrument A (Cell-Dyn 3500)

Figure 6: Cell-Dyn; Abbott Laboratories (32)

Instrument A (figure 6) is a well established instrument, used in large clinical pathology laboratories. It was originally developed for human medicine, but veterinary software with several different species makes the system applicable to veterinary medicine.
The system combines flow-cytometry and impedance technology. The red blood cell and platelet counts, as well as their parameters are determined by impedance technology. The white blood cells are analyzed with both methods. The »WIC-count« derives from impedance-technology; the »WOC-count« is determined by flow cytometry. The light scatter is measured at four different degrees: 0°, 10°, 90° and 90° depolarized light scatter. Multidimensional analysis of the light scatter allows the differentiation of the white blood cells as neutrophils, monocytes, lymphocytes, eosinophils and basophils. Comparison of both white blood cell counts is used for internal quality control. The »WOC-count« is routinely reported. The hemoglobin

content is measured spectophotometrically. The system generates different scatter- and histograms.

The following error messages are displayed by the instrument.

Error message	**Explanation**	**Kind of flag**	
		Technical	**Morphological**
WBC/WIC/WOC	Clinically significant difference between WIC and WOC count		✓
DFLT	Default criteria were used to determine the five-part-differential		✓
NWBC	Non-WBC population is present below the dynamic WBC threshold on the size/complexity scatter plot		✓
FWBC	Fragile WBC are suspected		✓
KWOC	Clinically significant difference between WIC and WOC count, and a kinetic decline in the WOC count rate		✓
NRBC	Nucleated red blood cells		✓
RRBC	Resistant red blood cells		✓
PLTR	Platelet recount		✓
Sampling Error	Sampling Error	✓	

Table 3: Error messages, displayed by instrument A

2.2.2. Instrument B (ADVIA™ 120)

Figure 7: ADVIA; Siemens Medical Solutions (photo: Bayer)

Instrument B (figure 7) is one of the most-advanced hematology system, which is available for veterinary medicine right now. Originally developed for human use, software for 21 different animal species was introduced several years ago. Because of the size, the quantity of samples run per hour and the price this system is usually limited to large clinical pathology laboratories.
Beside the routine parameters for cell count and white blood cell differential the system provides absolute reticulocyte counts, including a sub classification of reticulocytes as well as several platelet-parameters. The instrument is a flow-cytometry-based system, using light scatter, differential white blood cell lysis, cytochemical peroxidase staining and oxazine 750 staining to provide complete blood cell counts, white blood cell differential and reticulocyte counts of the peripheral blood (33). The hemoglobin value is determined photometrically by a modified Cyanmethemoglobin method. A number of scattergrams and histograms are provided by the system and can be used for further validation of the results (34). White blood cell counts are determined with the peroxidase method (WBCP) and the basophile method (WBCB), and comparison of both counts is used as internal quality control. Additionally, the Hb concentration is internally controlled by comparison of the mean corpuscular hemoglobin concentration MCHC and the corpuscular hemoglobin concentration mean (CHCM): CHCM is directly measured, based on cell-by-cell analysis, while MCHC is calculated, based on the Hb, MCV and RBC results.

The instrument generates a number of different error messages. The flag analysis is excluded, because the samples were measured on this system the next day. Therefore a higher number of error messages is expected (51).

2.2.3. Instrument C (LaserCyte®)

Figure 8: LaserCyte; IDEXX Laboratories (35)

Instrument C (figure 8) is the first in-house-hematology instrument using flow cytometry for cell differentiation. The instrument reports values for the complete blood cell count, reticulocytes in canine and feline species, as well as a 5-part-differential. Classical flow cytometry-based hematology instruments collect light that is scattered from cells at a variety of angles. These conventional systems utilize a collection lens which images the core stream, collects all the scattered light and passes it than through an aperture which determines the angle of collection. This filtered light is then focused down to a photo detector, which collects the entire cone of scattered light at the desired angles. The LaserCyte instead uses a lensless design, collecting signals at four different degrees: extinction (0°-0.5°), low angle forward scatter (1°-3°), high angle forward scatter (4°-9°) and right angle scatter (50°-130°). The extinction and the low angle forward scatter offer information about the cell size, the high angle forward scatter about internal granularity and the right angle scatter about the cell surface and the internal lobularity. Additionally the instrument determines the »time of flight«, which defines the time a particle needs to pass the

laser beam. Therefore the »time of flight« also gives some information about the cell size. Combining these five parameters allows the system to determine values for the blood cell counts, as well as for the differential. To count the reticulocytes the residual RNA is colored with New Methylene Blue. This stain is part of the CBC5R tube. This tube also contains Latex-particles, which are used at known concentration, and are to act as a cellular surrogate. This surrogate must have properties such that the instrument in at least one measurement technique can uniquely distinguish it from the cellular constituents of the sample. This cellular surrogate acts as an internal standard for quality control. This assures that the instrument's measurements are performing properly, and that all dilutions have been made properly. As another internal quality control the instrument measures the hemoglobin concentration photometrically in two different ways: in a conventional manner after lyses of the red blood cells and in an unconventional manner on unlysed red blood cells. The unconventional manner is not as accurate as the conventional one, because the red cell membranes scatter lights. Therefore the reported Hb-value is the one obtained after lyses of the red blood cells, but the comparison of both values allows the control of the dilution. The instrument generates a number of different scatter- and histograms, which can be printed on request.

Table 4 displays the different error messages, generated by the instrument.

Error message	Explanation	Kind of flag	
		Technical	Morphological
DB1/2, DB1/3, DB1-5	Differential algorithm issues, confirm with blood smear		✓
RB9	MCHC out of reportable range		✓
HB1	Hemoglobin out of reportable range		✓
PB2	MPV out of reportable range		✓
DB10	Possible rate analysis issue. Confirm differential with blood smear	✓	
HI	Possible dilution issue. Review blood smear and check WBC	✓	
RB2	Low RBC statistics. Distribution parameter not reported		✓
RB1	Too many RBC fragments. Confirm PLT value with blood smear		✓
DA1-5	Differential count too high. Confirm with blood smear		✓
RB3	Low PLT statistics. Distribution parameter not reported		✓
Hb air timing	Hemoglobin air timing		✓
HI1	Hemoglobin sheath timing variability		✓

Table 4: Error messages displayed by instrument C

2.2.4. Instrument D (ForCyte)

Figure 9: ForCyte, Oxford Science

Instrument D (figure 9) is an in-house hematology instrument which utilizes laser- and impedance-technology to perform a complete blood cell count, a five-part differential and a reticulocyte count in canine samples. The blood cell counts, as well as the red blood cell and platelet parameters are measured by impedance-technology. The laser-technology is used to perform the five-part differential.
The instrument displays only two different flags, so called »Rerun« and »Slide«. Neither the manual of the instrument nor the result printouts provide an explanation for these error messages.

2.2.5. Instrument E (MS45)

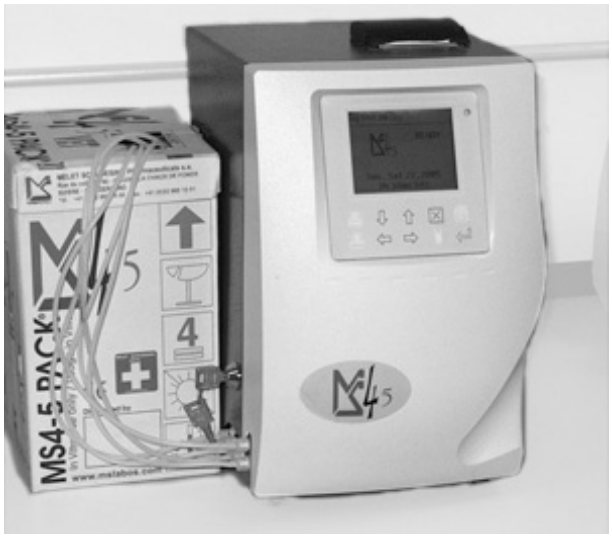


Figure 10: MS45, Melet Schloesing Pharmaceuticals (36)

Instrument E (figure 10) is the only evaluated impedance-based instrument which generates, besides a CBC a five-part differential. Histograms for white blood cells, red blood cells and platelets are provided. The following table lists the error messages, displayed by the system.

Error message	**Explanation**	**Kind of flag**	
		Technical	**Morphological**
# next to WBC, Eos, RBC or PLT	Clogging has been detected before analysis		✓
& next to WBC	Clogging has been detected during analysis	✓	
@ next to RBC or PLT	Homogeneity error has been detected during analysis	✓	
M next to differential	Monocyte value too high due to blasts or old blood		✓
R next to RBC, MCV, HCT and Hb	Result rejected due to error value of MCHC		✓
R next to WBC	Result rejected due to an analysis error		✓
R next to differential	Result rejected due to an analysis error		✓

Table 5: Error messages displayed by instrument E

2.2.6. Instrument F (Heska CBC)

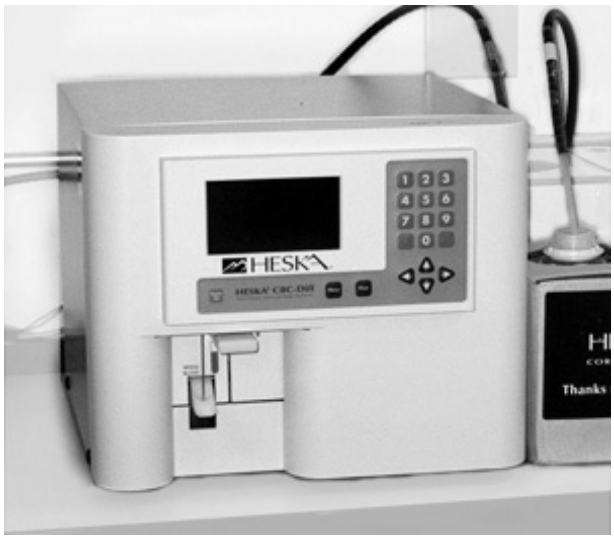


Figure 11: CBC, Boule Medicalab (37)

Instrument F (figure 11) employs impedance technology for cell counting and sizing, and a colorimetric method for measuring haemoglobin. Floating discriminators (thresholds) are used for discrimination of PLT and RBC, as well as for the generation of the white blood cell differential (three-part differential). In case no valley between the different cell populations can be found an error message (table 6) will be displayed and the floating threshold will be located at the point where the lowest number of cells is found.

Error message	**Explanation**	**Kind of flag**	
		Technical	**Morphological**
BD L/G/M	Lymphocyte and granulocyte populations are overlapped		✓
GM L/G/M	Large majority of the cells are seen as granulocytes		✓
LM L/G/M	Large majority of the cells are seen as lymphocytes		✓
DE PLT or RBC Rerun	Distribution Error (PLT and RBC cannot be differentiated)	✓	
FD RBC/PLT	Floating Discriminator		✓
SE WBC Rerun	Statistical Error	✓	

Table 6: Error messages displayed by instrument F

2.2.7. Instrument G (Scil Vet ABC)

Figure 12: ABC, Scil animal care company (38)

Instrument G (figure 12) is also an impedance-based in-house instrument, which provides a complete blood cell count and a three-part differential. In contrast to instrument F, the system uses fixed thresholds to discriminate the different cell populations.
Table 7 presents the error messages of instrument G.

Error message	**Explanation**	**Kind of flag**	
		Technical	**Morphological**
* next to WBC	Sample was analyzed three times and all three counts differed	✓	
$ next to WBC, PLT or RBC	Two of the three counts were within precision limits		✓
D next to PLT, RBC, Hb or HCT	Linearity range for the parameter has been exceeded		✓
MIC PLT	Minor crossover has occurred between PLT and RBC		✓

Table 7: Error messages displayed by instrument G

2.2.8. Instrument H (VetScan HMT)

Figure 13: HMT, Abaxis (39)

Instrument H (figure 13) is very similar to the instrument E. Except instrument H performs only a three-part differential. The error messages displayed are the same than generated by instrument E (Table 5).

2.2.9. Instrument K (QBC® Vet Autoread ™)

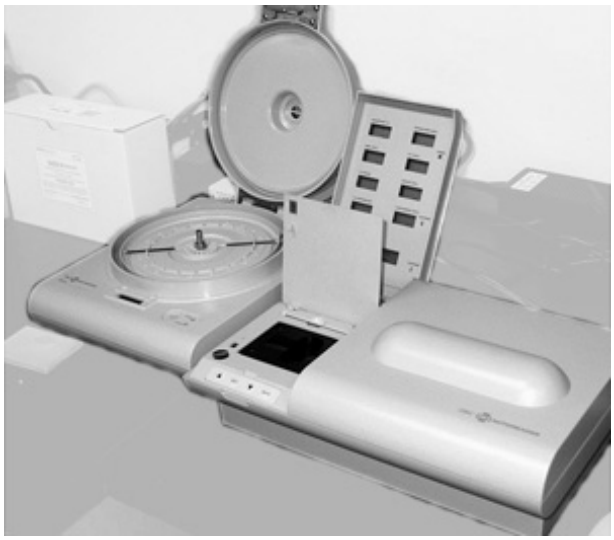


Figure 14: QBC, IDEXX Laboratories (40)

Instrument K (figure 14) is based on the principle that different blood cells have different densities and that they sort into individual layers when spun in a microhematocrit tube. A cylindrical float inserted into a capillary tube expands the buffy coat. The inner surface of the capillary tube is coated with acridine orange, a fluorescent dye that stains a variety of different cellular components (DNA, RNA, Lipoproteins, Glycosamines): normal erythrocytes are not stained by acridine orange, granulocytic cells fluoresce orange-yellow, lymphocytes and monocytes brilliant green and platelets pale yellow. The fluorescence emitted by the cells in the tube is examined and values for WBC, PLT, MCHC, Hb, HCT, Granulocytes and Lymphocytes/Monocytes are reported. The instrument is able to quantify reticulocytes, as a percentage of the hematocrit, within a specific range (0.2% to 4.0%). Furthermore a Buffy Coat profile for result verification is routinely displayed.

The following error messages are generated by the instrument:

Error message	Explanation	Kind of flag	
		Technical	Morphological
Granulocytes (1)	Red cells have not separated cleanly from granulocytes		✓
Buffy Coat (1)	Granulocytes have not separated cleanly from the lymph/mono layer and RBC layer		✓
Buffy Coat (3)	Lymph/mono layer has not separated distinctly from the other cell layers, possibly due to inadequate staining		✓
Buffy Coat (4)	Buffy Coat layer inconsistent due to clumped platelets, granulocytes, expired tube or stain on tube exterior	✓	
Buffy Coat (6)	Buffy Coat cell layers inconsistent, re-spin sample and retest	✓	
PLT (1)	Platelets found on top of float, re-spin sample and retest	✓	
Hb (1)	The presence of either immature red blood cells or of cells on top of the float may affect the hemoglobin measurement	✓	

Table 8: Error messages, displayed by instrument K

2.3 Statistics

Results were collected electronically and directly transferred to an excel spread sheet for analysis, for the following exceptions: PCV, the data generated by instruments D and K and the manual differential. This data had to be entered manually.

2.3.1 Cell count

The results were statistically analyzed by linear regression, Deming regression (normally distributed measurement errors), Passing-Bablok regression (not normally distributed measurement errors), Bland-Altman plots (41; 42) and columnar statistics using Microsoft Excel with Analyse-it® and GraphPadPrism version 4.0. The

correlation coefficients (r) and the standard deviations of the residuals ($S_{y|x}$) from the linear regression, the intercept and slope with 95% confidence intervals calculated by Deming regression or Passing-Bablok regression and the biases with 95% limits of agreement, calculated by Bland-Altman plots are reported for each instrument and each parameter. The mean was calculated for each instrument. Absolute values of total errors were calculated as follows: bias plus three times standard deviation of the replication experiment (15). The absolute errors were converted into percentage, by using the mean of the specific instrument. Total errors were compared to requirements for human medicine (24) and recommendations for veterinary medicine, published several years ago (20). The accuracy was determined by comparing results to the laboratory instrument B, which was identified as the comparative method in this study. Accuracy was assessed separately for dogs and cats for WBC, RBC, MCV, Hb, PLT and Reticulocytes for each point-of-care instrument and laboratory instrument A. The hematocrit was compared to the packed cell volume, performed with the Stat Spin VT Centrifuge.

2.3.2 Differential

Accuracy of instruments A, B, C, E, F, H, D and G was assessed in comparison to the manual differential. Canine and feline samples were analyzed separately. Absolute values were used for comparison. The absolute values were calculated separately for each instrument using the white blood cell counts of the accordant instrument. The results were statistically analyzed by linear regression, Passing-Bablok regression and bias plots (41; 42) using Microsoft Excel and Analyse-it® and GraphPad Prism version 4.0. The estimated intercept and slope, obtained by Passing-Bablok regression were only interpreted in cases with r-values above 0.975 (15; 29). The number of true and false neutropenias, neutrophilias, eosinophilias and lymphopenias was assessed for each instrument.

2.3.3 Replication experiment

For the replication experiment the standard deviations and coefficients of variation were calculated.

2.3.4 Linearity experiment

For the linearity study for each instrument the values for level two, three and four were calculated, based on the contributions of level one and level five in each of the other levels. Afterwards the mean of the measured values were compared to the calculated values using a regression analysis.

2.3.5 Carry-over experiment

The carry-over was calculated by subtracting the mean obtained from the third run of PBS from the mean of the first run of PBS and dividing by the mean of level five for WBC, RBC, Hb and PLT.

3. Results

Over a three month period blood samples from healthy and diseased dogs (n=260) and cats (n=110) were analyzed. Because of sample volumes and some instrument issues some blood samples were not analyzed on each instrument.

3.1 Instrument Specifications and IT-options

The in-house-hematology instruments vary in their specifications and offered IT-options (table 2).

Instrument C requires no maintenance and separate control material is not available. The instrument is connected to a provided PC, where an unlimited amount of data can be stored. The data can then be managed with the IDEXX VetLab®. The cycle time (12min) is relatively long in comparison to the other in-house-hematology instruments (73sec-7min).
Instrument D also requires no routine maintenance; control material is not available. The data is stored within the instrument and can be reaccessed.
Instruments E and H are self-cleaning; if the instruments is not used for more than one week minimal maintenance is required. For instrument H the manufacturer offers control material. Up to 100 analyses are stored in safeguard mode. Both instruments have a built-in printer; an external printer can optionally be connected.
At the beginning of each day a background count should be performed on instrument F. Additionally cleaning cycles are necessary monthly. Once a year, the hemoglobin photometer needs to be adjusted. The manufacturer also provides control material. Up to 250 analyses are stored within the instrument.
For instrument G a startup cycle at the beginning of each day and a standby cycle at the end of each day are necessary. A cleaning cycle needs to be performed once a

week. Control material is available. Up to 60 results can be stored with the Smart Card memory; optionally a PC can be connected for further data storage and a program for further data analysis.
Instrument K requires no maintenance and a calibration rod is available. The instrument itself cannot store data, but the system can be connected to the IDEXX VetLab®, where the data will be stored on the PC. The time to prepare the samples is much longer than for the other instruments.

3.2 Flags

3.2.1 Blood Cell count

The in-house hematology instruments generate a number of different flags associated to the cell count. The percentage of measurements flagged with a technical flag ranges from 1% to 36% (table 9). In-house instrument F flags a high percentage of feline samples with a technical flag (36%). This impedance-based instrument separates the different cell populations by floating thresholds, which are set at the point where the lowest number of cells is detected (valley) (figure 15). If the instrument is unable to find such a valley an error message is generated (figure 16). This error message clearly indicates to repeat the analysis. Most of these flags pass into a »floating discriminator flag (FD)« in the second measurement. This is also generated because no valley within the limits defined in the »Discriminator Setup« can be found. Therefore the floating threshold will be located at the point where the lowest number of cells is found and the »floating discriminator flag« will be displayed. The manufacturer considers this flag as a warning and not as an error flag. This flag is reflected in the high number of technical and morphological flags associated with the platelet count in feline samples (89%) (table 9).

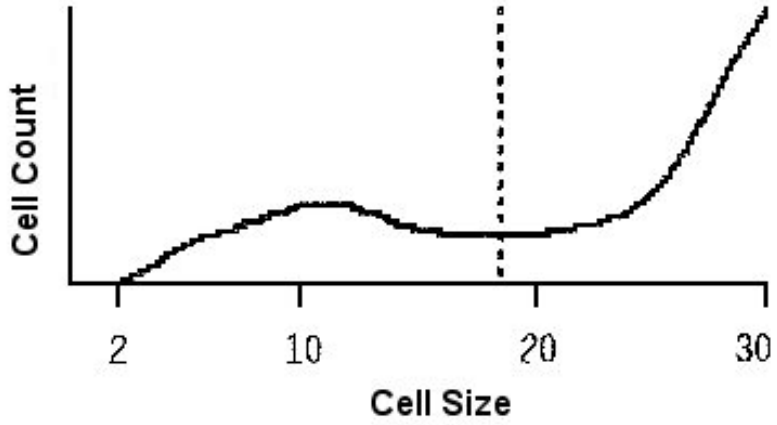


Figure 15: PLT-Histogram of instrument F for a feline sample without an error message. The instrument was able to define a valley between PLT and RBC

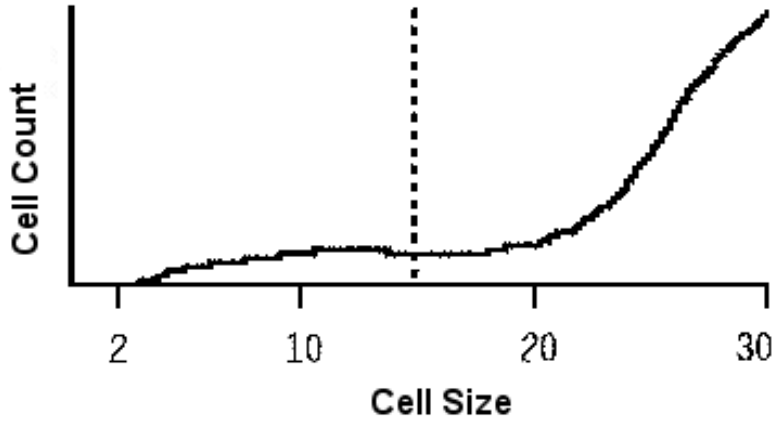


Figure 16: PLT-Histogram of instrument F for a feline sample with a »FD« error message. The instrument was unable to define a valley between PLT and RBC

Comparable to the FD flag of instrument F is the »MIC PLT flag« of instrument G which also leads to a high number of flags associated to the platelet count in feline samples (44%) (table 9). In this case a minor crossover has occurred between platelets and red blood cells. Unlike instrument F, instrument G uses a fixed threshold to separate platelets and red blood cells. Instrument K also displays a high number of flags associated to the platelet count as well as the white blood cell count in canine and feline samples (table 9). This instrument generates a number of sample alerts, mostly due to platelet aggregation or cell clumping. The manufacturer recommends

verifying the results by inspection of the buffy coat profile or examination of a blood smear. For some of these sample alerts the sample needs to be redrawn or analyzed a second time. These flags were classified as technical flags. The high number of flags associated to the Hb value in feline samples (40%) (table 9) is caused by cells on top of the float. Instrument K measures hemoglobin and the mean corpuscular hemoglobin concentration (MCHC) by a buoyancy[12] calculation from the amount that the float sinks into the red blood cells. If there is a greater concentration of lighter-density cells that have gathered at the top of the RBC layer (such as reticulocytes or nucleated red blood cells) or if clumped platelets lodge on top of the float the hemoglobin concentration and the MCHC maybe falsely decreased because the float sinks further into the red cells. Therefore when the system detects an MCHC outside what normally would be expected it will automatically flag the hemoglobin concentration (40). This technology explains the high number of flags associated to the Hb concentration in feline samples (table 9). The other in-house instruments display a relatively low number of flags associated with the cell count. Most of these flags are caused by technical failures and not by morphological changes of the samples. For comparison the flag analysis of the laboratory instrument A is also shown. It displays a high number of error messages associated with the white blood cell count (table 9). Instrument A measures the white blood cell count in the impedance (WIC) as well as in the optical (WOC) channel. If a significant difference exists, an algorithm decides if the WIC value or the WOC value is reported (32). Simultaneously an appropriate flag is displayed. Flag analysis for laboratory instrument B has not been done as the blood samples were analyzed on the next day and therefore significant higher numbers of flags are expected.

To determine if there is a difference in the accuracy between flagged and not-flagged samples, Bland-Altman analyses separating these two groups are shown for some instruments and some parameters (figure 17).

[12] buoyancy is the upward force on an object produced by the surrounding fluid, due to the pressure difference of the fluid between the top and the bottom of the object

Figure 17: Bland-Altman plots: separation between flagged (○) and not-flagged samples (●)

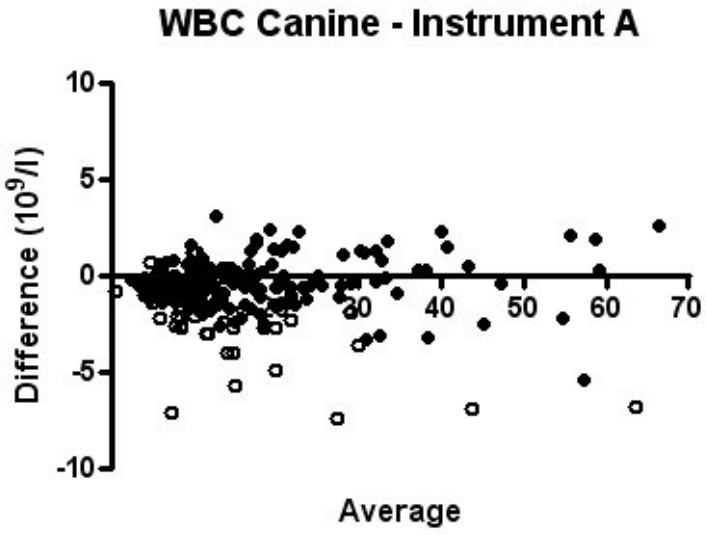


Figure 17.1: White blood cell count, Canine Instrument A versus Instrument B. The significant outliers are mostly correctly identified by flags.

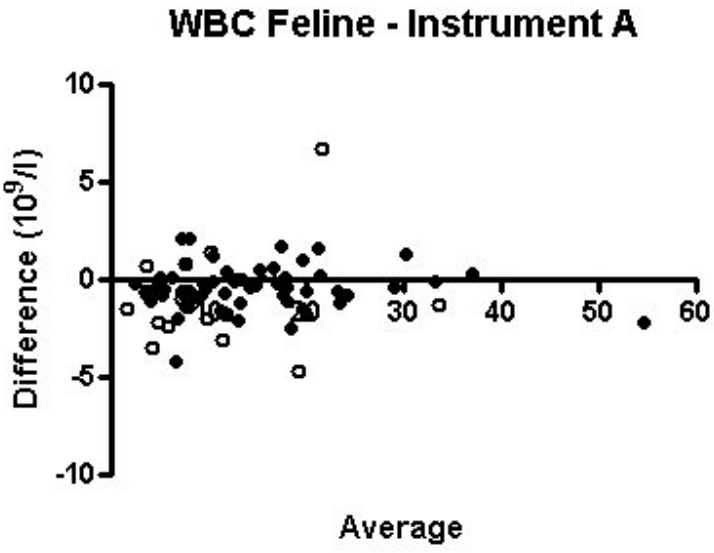


Figure 17.2: White blood cell count, Feline Instrument A versus Instrument B. Most of the significant outliers are flagged.

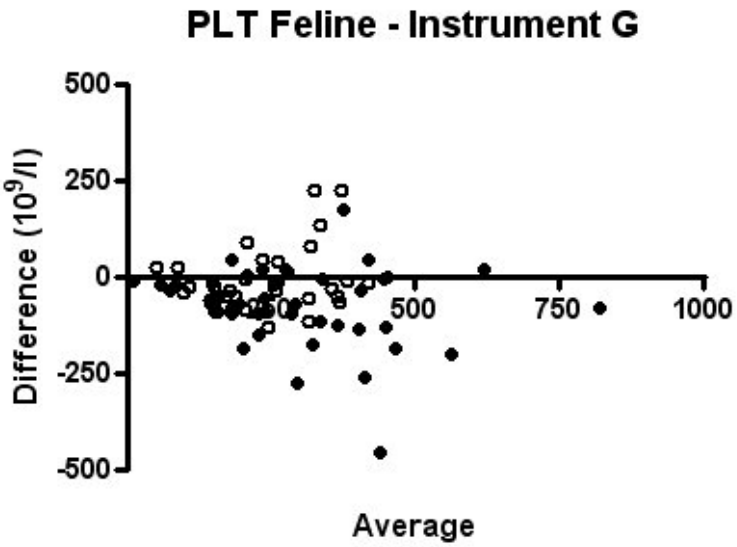


Figure 17.3: Platelet count, Feline Instrument G versus Instrument B. Some of the outliers are flagged, some are not.

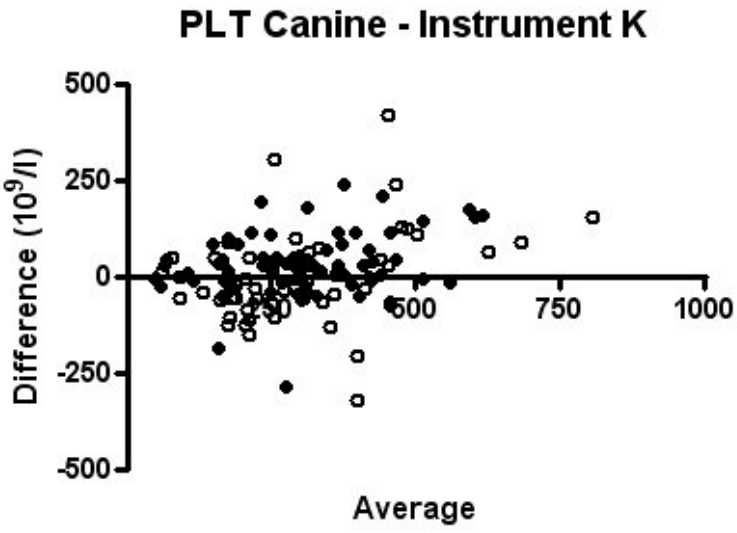


Figure 17.4: Platelet count, Canine Instrument K versus Instrument B. Some of the outliers are flagged, some are not.

Interestingly, there are a number of flagged values which appear to correlate well between in-clinic and laboratory method and are thus likely accurate, while values of some not-flagged samples differed between in-clinic and laboratory instruments (figure 17).

3.2.2 Differential blood cell count

The instruments display a number of different flags, associated with the differential (table 9). Instruments A, C, E and F flag a high percentage of canine and feline samples. To verify if the displayed error messages are appropriate, Bland-Altman analyses differentiating between flagged and not-flagged samples are performed. Comparable to the findings of the cell count, the Bland-Altman plots for in-house-instruments C, F and E show that a high number of samples with accurate results are flagged, but that also some samples with inaccurate results are present which are not flagged (figure 18). That means the instruments generate inaccurate results without displaying an error message. The number of these samples varies between the parameters, the species and the instruments. Instrument F more frequently identifies inaccurate results with flags than instruments C and E (figure 18). Laboratory instrument A also displays a high number of flags associated with some parameters and one or both species (figure 18.4). The instrument flags a high number of samples with accurate results and the number of inaccurate results without flags is lower than for instruments C, F and E. Figure 18 show examples of the Bland-Altman plots for each of the discussed instruments.

Figure 18: Bland-Altman plots: separation between flagged (○) and not-flagged samples (●)

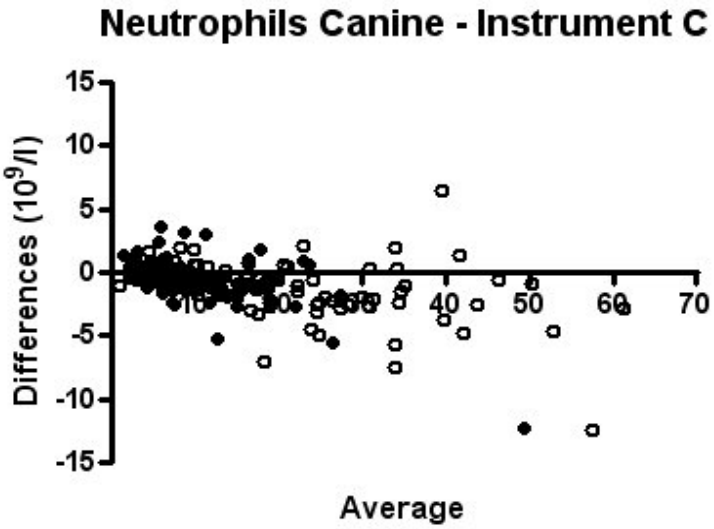


Figure 18.1: Neutrophil count, Canine
Instrument C versus manual differential.
Most of the significant outliers are flagged.

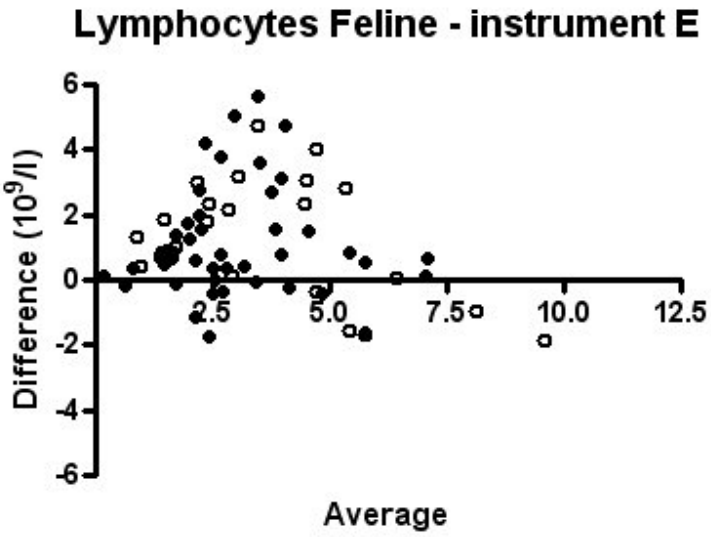


Figure 18.2: Lymphocyte count, Feline
Instrument E versus manual differential.
Some of the outliers are flagged, some are not.

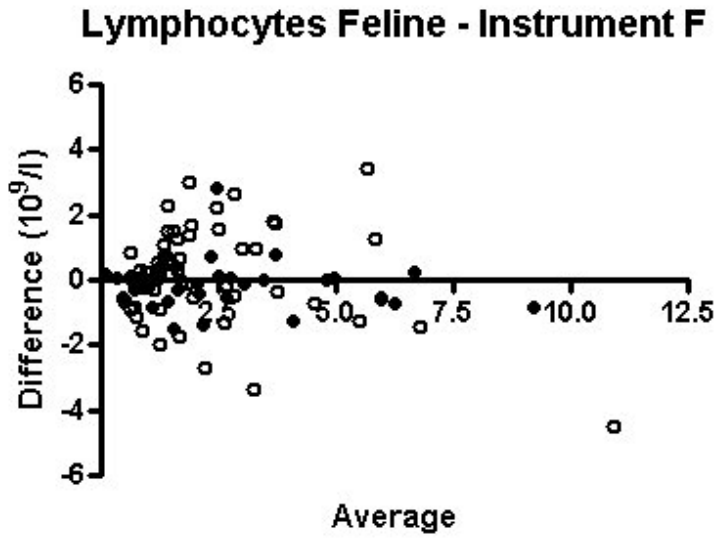


Figure 18.3: Lymphocyte count, Feline Instrument F versus manual differential. Most of the significant outliers are flagged.

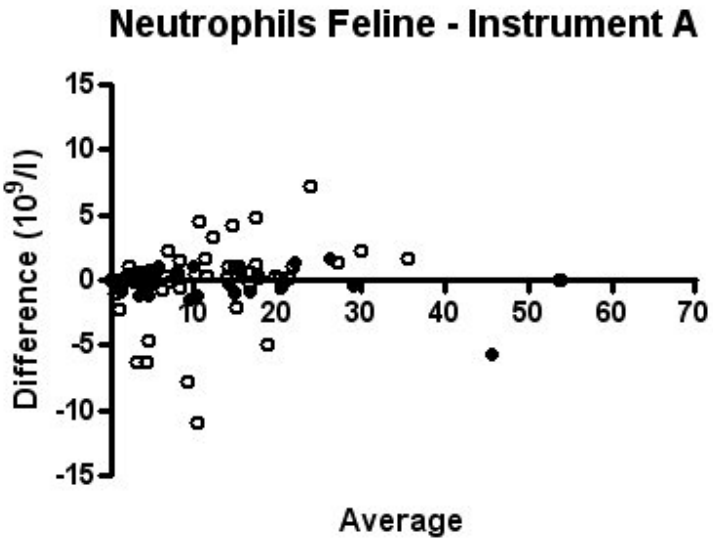


Figure 18.4: Neutrophil count, Feline Instrument A versus manual differential. Significant outliers are flagged.

The information offered by the manufacturers concerning the error messages varies: the error messages displayed by in-house-instrument C are caused by differential algorithm issues, and the manufacturer recommends confirming the results with a blood smear. Instrument E shows two different types of flags: one of these flags is indicative of blast cells or aged blood samples; the other one appears if the instrument recommends rejecting the results because of an analysis error (36). Three different types of error messages occurred for instrument F: »BD« means that the lymphocyte and granulocyte populations overlapped. Reasons for this overlap include extremely fragile granulocytes and aged blood. »LM« is displayed if a large majority of the cells are classified as lymphocytes. This flag may also indicate aged blood or white blood cells that have collapsed into the lymphocyte region of the histogram. »GM« is displayed if a large majority of the cells are identified as granulocytes or if no significant lymphocyte population is noted (37). Laboratory instrument A displays a »DFLT«-flag. This indicates that default criteria were used to determine the five-part-differential. This is caused by the presence of abnormal cell clusters that the instrument cannot reliably discriminate between, or by a low number of cells in a specific subpopulation. Descriptors in parentheses are added to the flag to indicate which subpopulations are suspect, based on the criteria used (32). For laboratory instrument B no flag analysis is performed because the samples were 24 hours old when measured. Therefore significant higher numbers of error messages are expected (51).

Table 9: Detailed flag analysis for all in-house hematology instruments and laboratory hematology instrument A

Hematology instruments	A		C		D		E		F		G		H		K	
Species	Ca	Fe	Ca	Fe	Ca	Fe	Ca	Fe	Ca	Fe	Ca	Fe	Ca	Fe	Ca	Fe
Sample number	259	110	260	109	90	44	178	72	242	95	249	107	236	97	173	50
% of flagged samples	65	89	61	64	8	20	7	39	61	98	10	58	6	19	61	80
% of measurements flagged with technical flags	2	5	3	6	2	2	3	6	1	36	4	2	6	6	14	8
% of measurements flagged with technical and morphological flags associated with each parameter																
WBC	44	72	1	3	8	20	3	14	1	-	6	22	2	5	21	28
RBC	2	4	1	4	8	20	3	3	-	2	1	3	2	3		
Hb	2	4	1	2	8	20	1	3	-	-	1	-	1	1	9	40
MCV	2	4	-	-	8	20	1	3	-	-	-	-	1	1		
HCT	2	4	-	-	8	20	1	3	-	-	1	-	1	1	-	-
PLT	2	4	1	3	8	20	3	1	3	89	4	44	3	2	57	68
Neutrophils	26	64	53	56	7	20	2	37								
Granulocytes									61	60	3	-	2	11		
Eosinophils	25	64	6	9	7	20	2	15								
Lymphocytes	29	66	12	11	7	20	2	37	61	60	3	-	2	11		
Monocytes	25	66	58	58	7	20	2	37	61	60	3	-	2	11		

3.3 Precision, Linearity and Carryover

In the linearity study all instruments achieve a high correlation with correlation coefficients of at least 0.993 for WBC (range 0-80x10^3/µl), Hb (range 0-22g/dl), RBC (0-8.5x10^6/µl), and PLT (0-1500x10^3/µl). Only instrument G obtains a correlation coefficient of 0.970 for the white blood cell count. The regression analysis indicates that instrument G has difficulties in detecting high white blood cell counts, because the data appears to plateau at the higher levels.
In the precision study the coefficients of variation for white blood cells range from 1.4% to 3.2%, for red blood cells from 0.5% to 4.1%, for hemoglobin concentration from 0.6% to 2.9%, for hematocrit from 0.7% to 3.9%, for MCV from 0.4% to 1.0% and for platelets from 2.2% to 9.2%, depending on the evaluated instrument (table 10).

Parameter	Instrument, Coefficients of variation (%)						
	A	C	D	E	F	G	H
WBC	1,4	3,1	2,3	3,2	2,6	2,2	1,4
RBC	0,5	4,1	2,8	1,2	2,0	2,7	1,4
Hb	0,6	2,9	1,3	1,1	2,3	1,8	0,7
HCT	0,7	3,9	2,8	1,6	2,5	2,7	2,1
MCV	0,6	0,4	1,0	0,9	0,7	0,6	1,0
PLT	2,2	4,7	4,4	6,7	6,4	7,2	9,2

Table 10: Precision for in-house and laboratory instrument A: Coefficients of variation

The required values for the carryover should not exceed 0.25% (34). This limit is achieved in all instruments for WBC, RBC, Hb and PLT, except for the following instruments and parameters: Instrument A achieves a value for the platelet count of 0.40% and instrument H clearly exceeds the limits for RBC (3.62%), hemoglobin (0.80%) and platelets (14.06%). But as instrument H is not able to measure PBS, level two is used instead of PBS. The data shows that the high values are mostly not caused by a carryover but by imprecision in the repeated measurements of the two levels. Removing one outlier of level two, results in improvement in the carryover for platelets (1.25%).

3.4 Accuracy

3.4.1 Cell count

The accuracy of the in-house-hematology instruments is evaluated in comparison to the results of laboratory instrument B, which is identified as the comparative method in this study. The comparison between both laboratory instruments A and B is also included. Canine and feline samples are analyzed separately.

In table 11 the results of linear regression, Deming or Passing-Bablok regression and Bland-Altman-analysis are summarized. Table 12 shows the means (including data range), absolute values and percentages of the total errors and recommendations for total allowable errors for human and veterinary medicine.

The bias represents the amount of systematic error (an error that is always in one direction). Systematic error can be constant or proportional. A constant systematic error is present if the error stays constant over a range of concentrations, whereas in a proportional systematic error the error changes as concentration changes. The type of systematic error can be identified, by assessing intercept and slope of regression analyses and/or by visual assessment of Bland-Altman plots. If the correlation coefficient is less than 0.975, the estimations of intercept and slope obtained by linear regression are not reliable (15; 29). It is recommended to improve the data or use alternate statistics, like Deming regression (15; 29). The comparison of intercept and slope, calculated by Deming or Passing-Bablok regression with the Bland-Altman plots for our data suggests that in the case of low r-values (under 0.975) the estimations of the Deming or Passing-Bablok regression are also not appropriate. Therefore the type of systematic error is assessed with the Bland-Altman plots. Figures 19 show examples for different types of systematic errors. In figure 19.1 the Bland-Altman plot of instrument H for MCV in canine samples is displayed: the bias is clearly positive with a value of 7.76fl. The Bland-Altman plot shows that the values are overestimated, independently of the concentration range. This is indicative of a constant systematic error. Because of low r-value (0.849), intercept and slope can not be interpreted. Figure 19.2 shows the Bland-Altman plot of instrument F for RBC in canine samples: the bias is positive ($0.44x10^{12}$/l), the r-value is above 0.975. The slope

is 1.07; the intercept is very close to zero ($0.06x10^{12}$/l). This combination is indicative of a proportional systematic error. As the slope is greater than one a positive proportional systematic error is present. This means that the overestimation is proportionately stronger at higher ranges. This assumption is confirmed by the Bland-Altman plot (figure 19.2). In figure 19.3 the Bland-Altman plot of instrument G for MCV in canine samples is displayed. It shows that the instrument underestimates the values in comparison to the reference method, resulting in a negative bias of -3.33fl. The underestimation is proportionately stronger at higher values, indicative of a negative proportional systematic error. The estimation of intercept and slope are not appropriate as the r-value is too low. In case of high r-values, a negative proportional systematic error would result in a slope smaller than one.

A random error is an error which can be either positive or negative and whose direction and magnitude cannot be predicted (figure 19.4).

Figure 19: Different types of errors

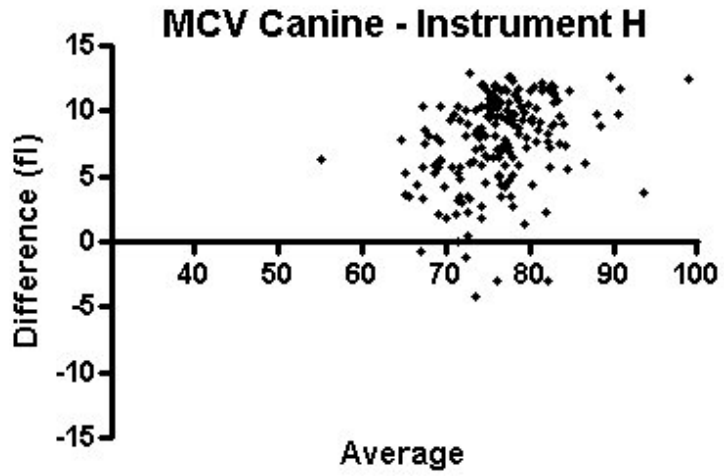


Figure 19.1: constant systematic error, indicated by clustering of data points above the zero bias, independent of concentration range

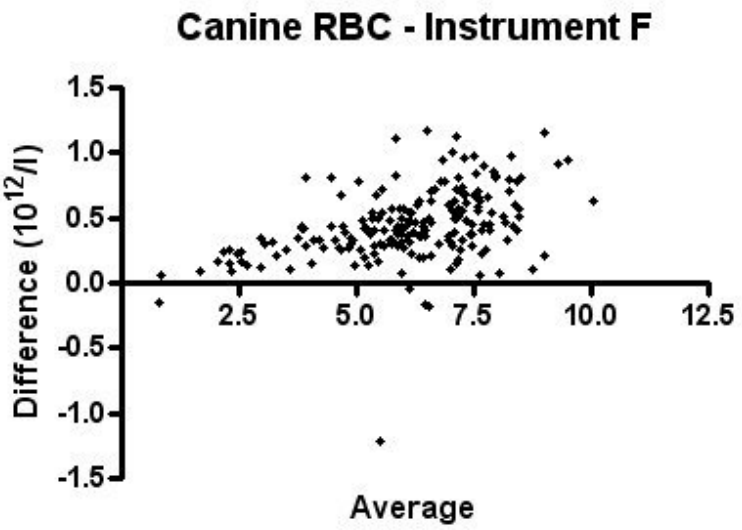


Figure 19.2: positive proportional systematic error, indicated by increasing positive differences with increasing concentration range

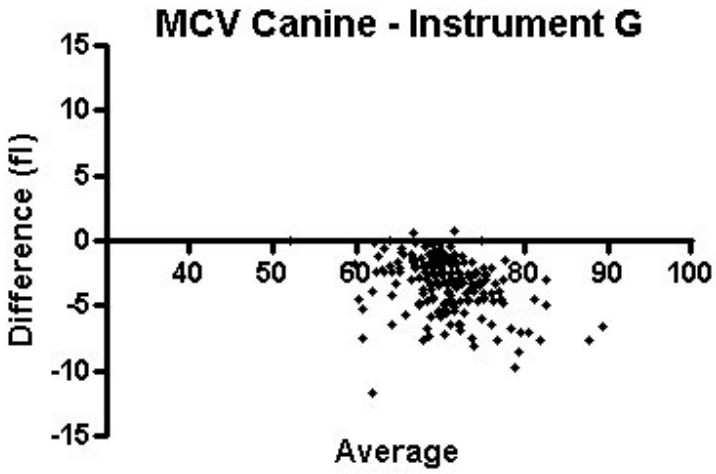


Figure 19.3: negative proportional systematic error, indicated by increasing negative differences with increasing concentration range

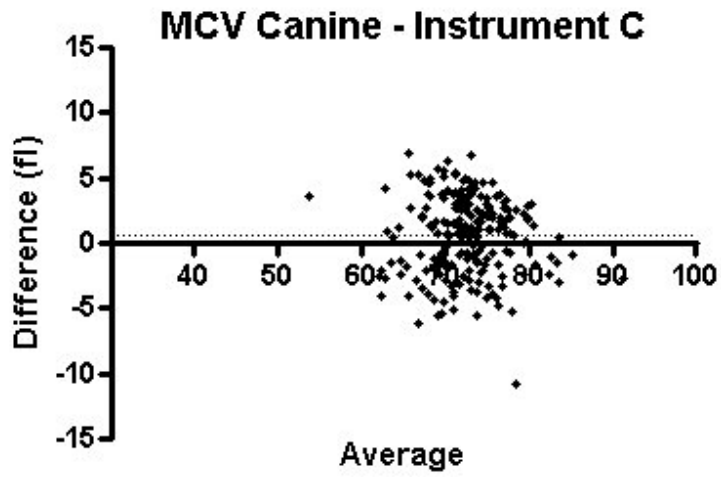


Figure 19.4: random error, indicated by clustering of data points around the zero line

In addition to the standard deviations, obtained in the replication experiment, the standard deviations of the residuals ($S_{y|x}$) and the 95% limits of agreement of the Bland-Altman analysis reflect the random error. The standard deviations of residuals ($S_{y|x}$) are influenced by two effects, namely by the total analytical imprecision (represented by the standard deviation calculated from the replication experiment [SD]) and by sample-related effects (29). When $S_{y|x}$>>SD sample-related effects are present. Furthermore, it needs to be considered that in the case of any relationship between the difference and the average of the two methods, the 95% limits of agreement might not be appropriate, as the observed range of differences depends on the average values (16). In our study $S_{y|x}$ exceeds SD in all parameters. That means sample-related effects are present.
The total error is the combined effect of random and systematic error, and the absolute value for total error can be estimated by the bias from method comparison plus three times the standard deviation from replication experiment (15). The absolute values are converted into percentage by using the mean of the measurements for the specific instrument: 100 divided by mean times absolute total error value. For example, instrument A achieved for white blood cell counts in canine samples a bias of -0.58x10^9/l. The standard deviation, according to the replication experiment is 0.12x10^9/l. Therefore the absolute value of the total error is 0.58x10^9/l + 3 x 0.12x10^9/l (bias plus three times SD) = 0.94x10^9/l. The mean of the measurements for which this total error is obtained, is 15.11x10^9/l. Hence the total error in percentage can be calculated as follows: 100 / 15.11x10^9/l x 0.94x10^9/l (100 divided by

mean times absolute value of total error) = 6%. A method is acceptable when the total error does not exceed the total allowable error (TE_a).

For white blood cell counts the total errors range from 5% to 12% (tables 12.2 and 12.2). Hence they all meet the CLIA requirements for human medicine (24), except the total errors of instrument K are higher (16% in canine and 21% in feline samples). In canine samples the slightly higher total error is caused by an elevated bias due to a constant systematic error, in feline samples a positive proportional error is responsible for the elevated bias and the higher total error. The total errors for red blood cell counts range from 2% to 18% (tables 12.3 and 12.4). In canine samples instruments C, F, D, and G exceed the CLIA requirements (24), as well as previous published recommendations for veterinary medicine (20). In feline samples all instruments except A and H show total errors above the recommendations published by Freeman et al. (20). In instrument C the higher total errors are caused by increased standard deviations, achieved in the replication experiment (table 10), indicating random error. Instrument F shows a positive proportional error in both species (tables 11.3 and 11.4), as does instrument G with feline samples (table 11.4). For instrument E in felines and for instrument D in both species a constant systematic error is present (tables 11.3 and 11.4). For hemoglobin, only instruments A, H and D meet the CLIA requirements, with values equal or below 7% (tables 12.3 and 12.4). For the other instruments the total errors range from 8% to 19%. Instruments G in both species, as well as instruments F and H with feline samples, show a positive proportional error (tables 11.5 and 11.6); instruments E and K in both species show a mixture of constant and proportional systematic error. For MCV no CLIA requirements are available. The total errors range from 2% to 9% (tables 12.7 and 12.8). Only instrument H shows higher total errors of 13% for canine and 19% for feline samples. In both species this is due to a strong constant systematic error. Instrument D in both species, instrument F in canines and instruments A and H in felines show also a constant systematic error. For instrument F in felines and for instrument G in canines and felines a mixture of constant and systematic error is obvious. Except for instruments B and K in both species and instrument A in canine samples, the total errors for hematocrit are very high with values from 11% to 26% (tables 12.9 and 12.10). Because the hematology instruments calculate the hematocrit by multiplying red blood cell count and MCV, the total errors for hematocrit reflect the errors in these two parameters. Only instrument K in canine samples meets the CLIA requirements (24); instrument B with canine and feline, instrument K with

feline and instrument A with canine samples fulfills the recommendations published by Freeman et al. (20). The total errors achieved for platelet counts vary between the instruments. The total error of instruments A, C, D and K are between 20% and 31% (tables 12.11 and 12.12); hence they are lower or only slightly above the CLIA requirements (25%). Instrument G shows a total error of 59% in both species, this is caused by high imprecision as well as by strong biases. Instrument H achieved unacceptable high errors of 68% and 120%, respectively.

The results are summarized in table 11 and table 12

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to laboratory instrument B

	A	C	D	E	F	G	H	K
N	204	209	70	141	203	204	189	136
r	0.994	0.990	0.992	0.989	0.990	0.988	0.992	0.942
$S_{y|x}$	1.27	1.55	1.74	1.78	1.10	1.11	1.53	3.62
Slope	1.01	1.04	1.03	1.03	1.03	1.13	1.03	0.97
95% CI	0.99 to 1.04	1.02 to 1.06	0.99 to 1.08	1.00 to 1.06	1.02 to 1.05	1.10 to 1.15	1.01 to 1.05	0.91 to 1.03
Intercept	-0.79	-0.26	0.29	0.65	-0.39	-0.21	0.72	2.06
95% CI	-1.00 to -0.45	-0.61 to 0.04	-0.45 to 0.81	0.16 to 1.15	-0.63 to -0.17	-0.46 to 0.07	0.35 to 1.09	0.92 to 3.20
Bias	-0.58	0.08	0.81	1.10	0.99	1.11	1.21	1.61
95% Limits of agreement	-3.09 to 1.94	-2.95 to 3.10	-2.61 to 4.22	-2.41 to 4.61	0.82 to 1.19	0.91 to 1.35	-1.82 to 4.23	-5.72 to 8.93

Table 11.1: WBC Canine (10^9/l)

	A	C	D	E	F	G	H	K
N	74	74	26	50	67	76	69	33
r	0.992	0.986	0.966	0.970	0.993	0.979	0.984	0.950
$S_{y|x}$	1.16	1.56	2.50	2.19	1.20	2.07	1.67	3.08
Slope	0.99	1.01	1.05	1.00	1.04	1.10	1.05	1.22
95% CI	0.96 to 1.02	0.96 to 1.05	0.93 to 1.17	0.93 to 1.08	1.01 to 1.07	1.05 to 1.16	1.00 to 1.10	1.07 to 1.36
Intercept	-0.37	-0.14	-0.34	0.45	-0.52	-1.20	0.11	-0.74
95% CI	-0.87 to 0.12	-0.66 to 0.31	-2.67 to 2.00	-0.66 to 1.57	-0.89 to -0.12	-2.08 to -0.33	-0.63 to 0.84	-3.19 to 1.71
Bias	-0.49	-0.15	0.54	0.51	-0.11	0.22	0.75	2.42
95% Limits of agreement	-2.77 to 1.80	-3.18 to 2.88	-4.27 to 5.36	-3.73 to 4.75	-2.57 to 2.35	-4.03 to 4.47	-2.53 to 4.04	-3.90 to 8.74

Table 11.2: WBC Feline (10^9/l)
N: Number of samples; r: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to laboratory instrument B

	A	C	D	E	F	G	H	K
N	208	213	74	142	201	202	186	
r	0.985	0.983	0.965	0.983	0.990	0.992	0.985	
$S_{y|x}$	0.27	0.31	0.48	0.29	0.26	0.22	0.25	
Slope	1.00	1.02	1.01	1.00	1.07	1.02	0.90	
95% CI	0.98 to 1.02	0.99 to 1.04	0.95 to 1.08	0.97 to 1.03	1.04 to 1.09	1.00 to 1.04	0.88 to 0.92	
Intercept	0.01	-0.05	0.34	0.17	0.06	0.08	0.55	
95% CI	-0.12 to 0.15	-0.21 to 0.11	-0.02 to 0.70	-0.02 to 0.37	-0.08 to 0.19	-0.01 to 0.17	0.41 to 0.69	
Bias	0.00	0.04	0.41	0.17	0.44	0.22	-0.04	
95% Limits of agreement	-0.52 to 0.53	-0.57 to 0.65	-0.51 to 1.34	-0.39 to 0.73	-0.10 to 0.98	-0.21 to 0.66	-0,64 to 0.56	

Table 11.3: RBC Canine (10^{12}/l)

	A	C	D	E	F	G	H	K
N	76	73	30	51	68	76	70	
r	0.979	0.981	0.899	0.971	0.985	0.983	0.975	
$S_{y|x}$	0.44	0.42	0.82	0.58	0.40	0.44	0.46	
Slope	1.06	1.00	1.00	1.00	1.07	1.09	0.90	
95% CI	1.00 to 1.11	0.95 to 1.05	0.81 to 1.19	0.93 to 1.07	1.02 to 1.12	1.04 to 1.13	0.85 to 0.95	
Intercept	0.13	0.03	0.74	0.56	0.41	0.27	0.78	
95% CI	-0.24 to 0.50	-0.31 to 0.36	-0.52 to 2.00	0.05 to 1.06	0.08 to 0.74	-0.07 to 0.60	0.43 to 1.14	
Bias	0.52	0.01	0.75	0.55	0.89	0.85	0.11	
95% Limits of agreement	-0.35 to 1.38	-0.80 to 0.82	-0.83 to 2.33	-0.59 to 1.68	0.08 to 1.70	-0.07 to 1.77	-0.91 to 1.12	

Table 11.4: RBC Feline (10^{12}/l)
N: Number of samples; r: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to laboratory instrument B

	A	C	D	E	F	G	H	K
N	206	208	71	141	201	202	184	144
R	0.995	0.988	0.987	0.978	0.990	0.994	0.993	0.983
$S_{y|x}$	0.38	0.65	0.67	0.78	0.60	0.47	0.46	0.74
Slope	0.97	1.05	0.95	0.96	1.03	1.05	0.97	0.95
95% CI	0.95 to 0.98	1.03 to 1.08	0.92 to 0.98	0.93 to 0.99	1.01 to 1.05	1.03 to 1.07	0.95 to 0.98	0.92 to 0.98
Intercept	0.17	-0.68	0.41	-0.87	-0.08	-0.27	0.70	-0.55
95% CI	-0.00 to 0.36	-1.06 to -0.29	-0.05 to 0.80	-1.40 to -0.33	-0.38 to 0.23	-0.50 to -0.08	0.56 to 1.03	-1.00 to -0.11
Bias	-0.30	0.13	-0.21	-1.49	0.40	0.47	0.33	-1.33
95% Limits of agreement	-1.09 to 0.49	-1.17 to 1.43	-1.58 to 1.61	-3.07 to 0.08	-0.80 to 1,61	-0.51 to 1.45	-0.60 to 1.27	-2.85 to 0.19

Table 11.5: Hb Canine (g/dl)

	A	C	D	E	F	G	H	K
N	73	75	27	50	65	74	67	34
r	0.987	0.983	0.920	0.973	0.989	0.987	0.989	0.983
$S_{y|x}$	0.50	0.63	0.73	0.83	0.52	0.58	0.53	0.57
Slope	0.99	0.99	1.03	1.00	1.06	1.05	1.05	0.93
95% CI	0.96 to 1.03	0.95 to 1.04	0.92 to 1.14	0.93 to 1.06	1.02 to 1.10	1.01 to 1.09	1.01 to 1.09	0.87 to 0.99
Intercept	0.16	0.16	-0.09	-1.14	0.10	-0.53	0.19	-0.50
95% CI	-0.29 to 0.61	-0.35 to 0.67	-1.31 to 1.12	-1.93 to -0.34	-0.39 to 0.58	-1.00 to -0.05	-0.26 to 0.64	-1.22 to 0.23
Bias	0.08	0.10	0.25	-1.19	0.74	0.07	0.73	-1.28
95% Limits of agreement	-0.89 to 1.06	-1.12 to 1.33	-1.14 to 1.65	-2.80 to 0.43	-0.32 to 1.80	-1.11 to 1.25	-0.32 to 1.79	-2.49 to -0.08

Table 11.6: Hb Feline (g/dl)
N: Number of samples; r: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to laboratory instrument B

	A	C	D	E	F	G	H	K
N	207	209	73	140	199	201	184	
r	0.938	0.813	0.879	0.820	0.893	0.920	0.849	
$S_{y|x}$	1.74	2.84	2.26	0.67	2.32	1.82	3.40	
Slope	1.03	1.15	0.97	0.88	1.02	0.89	1.22	
95% CI	0.98 to 1.08	1.04 to 1.26	0.85 to 1.10	0.78 to 0.99	0.95 to 1.09	0.84 to 0.95	1.11 to 1.33	
Intercept	-0.37	-10.15	-3.06	9.42	-4.77	4.27	-8.13	
95% CI	-4.15 to 3.40	-18.17 to -	-12.00 to 5.88	1.91 to 16.92	-9.95 to 0.41	0.44 to 8.11	-16.13 to -	
Bias	1.87	0.65	-4.29	1.01	-3.53	-3.33	7.76	
95% Limits of agreement	-1.60 to 5.34	-5.25 to 6.56	-9.56 to -0.27	-5.29 to 7.31	-8.25 to 1.20	-7.47 to 0.80	1.10 to 14.41	

Table 11.7: MCV Canine (fl)

	A	C	D	E	F	G	H	K
N	71	72	30	50	68	76	69	
r	0.953	0.759	0.901	0.947	0.937	0.946	0.948	
$S_{y|x}$	2.28	2.20	1.65	1.84	1.95	1.64	2.45	
Slope	1.25	0.93	0.88	0.82	0.94	0.88	1.26	
95% CI	1.15 to 1.34	0.74 to 1.12	0.71 to 1.04	0.74 to 0.90	0.86 to 1.03	0.81 to 0.95	1.15 to 1.36	
Intercept	-8.41	2.46	3.63	7.85	0.13	3.44	-4.31	
95% CI	-13.08 to -	-6.71 to 11.62	-4.42 to 11.68	3.83 to 11.86	-4.13 to 4.40	0.01 to 6.87	-9.41 to 0.78	
Bias	3.69	-1.15	-2.48	-1.22	-2.57	-2.31	8.29	
95% Limits of agreement	-1.12 to 8.49	-6.80 to 4.49	-6.14 to 1.17	-5.87 to 3.43	-6.66 to 1.51	-6.13 to 1.52	3.05 to 13.53	

Table 11.8: MCV Feline (fl)
N: Number of samples; r: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to the PCV

	A	B	C	D	E	F	G	H	K
N	245	202	249	80	177	233	239	234	162
r	0.971	0.982	0.957	0.948	0.959	0.979	0.978	0.975	0.990
$S_{y|x}$	2.70	2.19	3.56	4.20	3.35	2.54	2.48	2.79	1.71
Slope	1.03	1.04	1.05	1.03	1.05	1.05	1.03	1.09	1.02
95% CI	1.00 to 1.05	1.01 to 1.06	1.01 to 1.09	0.96 to 1.11	1.00 to 1.09	1.02 to 1.08	1.00 to 1.05	1.06 to 1.12	1.00 to 1.04
Intercept	0.80	-0.38	-0.06	-0.01	1.34	0.25	-0.54	2.11	-0.95
95% CI	-0.15 to 1.90	-1.58 to 0.82	-1.76 to 1.65	-3.06 to 3.03	-0.68 to 3.37	-0.96 to 1.45	-1.53 to 0.35	0.74 to 3.48	-1.83 to -0.22
Bias	2.29	1.09	1.93	1.25	3.27	2.23	0.83	5.73	0.03
95% Limits of agreement	-3.01 to 7.58	-3.21 to 5.40	-5.09 to 8.95	-6.29 to 9.42	-3.29 to 9.83	-2.48 to 7.30	-4.03 to 5.69	0.12 to 11.35	-3.32 to 3.38

Table 11.9: PCV Canine (%)

	A	B	C	D	E	F	G	H	K
N	99	75	100	41	70	88	101	96	50
r	0.950	0.981	0.943	0.887	0.957	0.975	0.969	0.927	0.991
$S_{y|x}$	3.20	1.90	3.15	3.46	3.14	2.18	2.49	4.12	1.32
Slope	1.14	1.06	0.98	0.89	1.06	1.05	1.05	1.13	1.05
95% CI	1.07 to 1.22	1.01 to 1.10	0.91 to 1.05	0.74 to 1.04	0.98 to 1.14	1.00 to 1.10	1.0 to 1.11	1.04 to 1.23	1.01 to 1.09
Intercept	1.98	-0.39	1.51	6.62	1.56	1.89	2.15	2.99	-0.61
95% CI	-0.56 to 4.53	-2.01 to 1.24	-0.79 to 3.80	1.79 to 11.46	-0.96 to 4.08	0.18 to 3.60	0.39 to 3.91	-0.13 to 6.11	-2.05 to 0.84
Bias	6.61	1.37	0.74	3.21	3.43	3.50	3.82	7.18	1.01
95% Limits of agreement	0.24 to 12.98	-2.38 to 5.12	-5.43 to 6.91	-3.88 to 10.32	-2.71 to 9.57	-0.83 to 7.82	-1.10 to 8.73	-0.89 to 15.25	-1.67 to 3.69

Table 11.10: PCV Feline (%)
N: Number of samples; r: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 11: Results of regression analysis and Bland-Altman analysis for each instrument in comparison to laboratory instrument B

	A	C	D	E	F	G	H	K
N	200	204	69	135	190	195	169	124
r	0.824	0.886	0.790	0.870	0.887	0.890	0.804	0.789
$S_{y|x}$	76.71	73.13	103.80	75.46	66.12	57.27	79.99	101.2
Slope	0.98	1.08	0.97	1.05	0.97	0.84	1.08	1.13
95% CI	0.88 to 1.07	1.01 to 1.16	0.79 to 1.15	0.95 to 1.15	0.90 to 1.05	0.78 to 0.90	0.96 to 1.22	0.97 to 1.29
Intercept	-24.26	-17.02	15.60	-20.00	-14.48	1.06	-19.74	-17.46
95% CI	-55.16 to 6.63	-42.76 to 8.73	-45.94 to	-54.33 to	-38.14 to 9.18	-18.88 to	-56.55 to 8.28	-67.43 to
Bias	-31.65	8.38	6.38	-3.64	-22.19	-47.93	-8.02	20.31
95% Limits	-188.60 to	-134.93 to	-198.08 to	-151.24 to	-152.03 to	-168.65 to	-171.95 to	-177.58 to
of agreement	125.31	151.69	210.83	143.96	107.65	72.79	155.92	218.19

Table 11.11: PLT Canine (10^9/l)

	A	C	D	E	F	G	H	K
N	75	76	30	50	67	74	69	34
r	0.766	0.883	0.837	0.779	0.716	0.767	0.685	0.847
$S_{y|x}$	105.3	100.4	147.5	103.7	117.1	91.23	96.65	83.58
Slope	1.05	1.18	1.30	0.97	0.80	0.72	0.59	1.05
95% CI	0.84 to 1.25	1.02 to 1.33	0.97 to 1.62	0.74 to 1.19	0.61 to 1.00	0.58 to 0.86	0.44 to 0.74	0.81 to 1.29
Intercept	-46.74	-27.51	-24.76	54.88	54.98	38.16	47.10	-1.21
95% CI	-113.90 to	-79.78 to	-153.20 to	-13.79 to	-7.68 to 117.6	-8.47 to 84.79	-2.85 to 97.05	-69.20 to
Bias	-32.67	22.98	77.20	45.44	-1.04	-44.71	-69.33	11.88
95% Limits	-244.75 to	-174.95 to	-214.19 to	-156.48 to	-240.39 to	-243.66 to	-294.05 to	-150.05 to
of agreement	179.41	220.91	368.59	247.36	238.30	154.23	155.39	173.82

Table 11.12: PLT Feline (10^9/l)
N: Number of samples; R: correlation coefficient (linear regression); $S_{y|x}$: standard deviation of residuals (linear regression); Slope and Intercept (Deming or Passing-Bablok regression)

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	C	D	E	F	G	H	K
Mean (Range)	15.11 (0.20-67.70)	15.55 (0.42-66.20)	19.53 (4.45-66.32)	16.53 (0.78-69.92)	16.23 (0.50-71.50)	18.28 (0.70-72.54)	16.97 (0.77-68.41)	17.57 (5.80-56.60)
TE (10^9/l)	0.94	0.83	1.35	1.97	1.62	1.59	1.57	2.81
TE (%)	6	5	7	12	10	9	9	16
TE_a (CLIA)	15	15	15	15	15	15	15	15
TE_a (F)	20	20	20	20	20	20	20	20

Table 12.1: WBC Canine

	A	C	D	E	F	G	H	K
Mean (Range)	13.52 (1.60-53.70)	13.44 (1.94-55.13)	18.23 (1.48-37.16)	12.99 (3.57-53.69)	13.89 (2.70-56.80)	14.05 (1.73-58.87)	13.97 (1.83-55.61)	17.13 (4.30-34.70)
TE (10^9/l)	0.85	0.90	1.35	1.38	0.74	0.70	1.11	3.62
TE (%)	6	7	7	11	5	5	8	21
TE_a (CLIA)	15	15	15	15	15	15	15	15
TE_a (F)	20	20	20	20	20	20	20	20

Table 12.2: WBC Feline

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	C	D	E	F	G	H	K
Mean (Range)	5.99 (1.94-9.48)	5.98 (0.84-9.80)	5.72 (1.50-9.81)	6.31 (0.85-9.73)	6.33 (0.74-10.40)	6.11 (0.94-10.07)	5.95 (0.81-9.12)	
TE (10^{12}/l)	0.12	0.82	0.98	0.41	0.86	0.79	0.31	
TE (%)	2	14	17	6	14	13	5	
TE_a (CLIA)	6	6	6	6	6	6	6	
TE_a (F)	10	10	10	10	10	10	10	

Table 12.3: RBC Canine

	A	C	D	E	F	G	H	K
Mean (Range)	7.54 (0.59-11.93)	6.91 (0.92-10.83)	7.21 (2.56-11.91)	7.28 (1.01-11.48)	7.85 (0.99-11.74)	7.68 (1.00-11.89)	6.98 (0.94-10.43)	
TE (10^{12}/l)	0.64	0.79	1.32	0.79	1.31	1.42	0.38	
TE (%)	8	11	18	11	17	18	5	
TE_a (CLIA)	6	6	6	6	6	6	6	
TE_a (F)	10	10	10	10	10	10	10	

Table 12.4: RBC Feline

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	C	D	E	F	G	H	K
Mean	14.70	14.88	12.94	13.90	15.12	15.26	15.37	13.33
(Range)	(5.00-24.70)	(2.82-24.01)	(3.40-19.80)	(1.60-24.20)	(1.80-26.20)	(2.49-26.41)	(2.90-24.60)	(2.80-21.30)
TE (g/dl)	0.60	1.27	0.78	2.00	1.48	1.34	0.69	1.90
TE (%)	4	9	6	14	10	9	4	14
TE_a (CLIA)	7	7	7	7	7	7	7	7
TE_a (F)	7	7	7	7	7	7	7	7

Table 12.5: HGB Canine

	A	C	D	E	F	G	H	K
Mean	11.55	11.35	11.13	9.89	12.22	11.50	11.96	9.84
(Range)	(2.00-17.60)	(1.72-16.79)	(5.00-16-70)	(1.10-16.20)	(2.20-18.30)	(2.09-17.38)	(1.70-17.40)	(1.80-15.20)
TE (g/dl)	0.38	1.24	0.82	1.70	1.82	0.94	1.09	1.85
TE (%)	3	11	7	17	15	8	9	19
TE_a (CLIA)	7	7	7	7	7	7	7	7
TE_a (F)	7	7	7	7	7	7	7	7

Table 12.6: HGB Feline

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	C	D	E	F	G	H	K
Mean	74.06	72.79	66.15	73.74	68.72	69.00	80.18	
(Range)	(55.60-95.20)	(55.58-89.98)	(49.90-74.00)	(57.50-88.70)	(49.60-88.70)	(52.00-86.00)	(58.30-105.1)	
TE (fl)	3.04	1.37	6.12	3.02	5.03	4.59	10.10	
TE (%)	4	2	9	4	7	7	13	
TE_a (CLIA)	na	na	na	na	na	na	na	
TE_a (F)	na	na	na	na	na	na	na	

Table 12.7: MCV Canine

	A	C	D	E	F	G	H	K
Mean	52.37	47.06	46.79	48.09	46.43	46.51	57.15	
(Range)	(39.40-94.70)	(38.39-53.38)	(39.40-59.50)	(38.20-76.80)	(36.50-80.30)	(38.00-78.00)	(43.20-102.7)	
TE (fl)	4.86	1.87	4.31	3.23	4.07	3.57	10.63	
TE (%)	9	4	9	7	9	8	19	
TE_a (CLIA)	na	na	na	na	na	na	na	
TE_a (F)	na	na	na	na	na	na	na	

Table 12.8: MCV Feline

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	B	C	D	E	F	G	H	K
Mean (Range)	43.92 (12.50-70.50)	43.08 (7.50-74.00)	43.50 (7.08-75.21)	38.16 (9.50-69.70)	45.89 (7.10-74.60)	43.39 (6.60-75.20)	41.90 (7.83-72.25)	47.29 (7.70-76.60)	41.13 (7.80-71.70)
TE (%)	3.22	3.09	6.73	4.79	5.70	5.77	4.67	8.85	2.25
TE (%)	7	7	15	13	12	13	11	19	5
TE_a (CLIA)	6	6	6	6	6	6	6	6	6
TE_a (F)	10	10	10	10	10	10	10	10	10

Table 12.9: HCT Canine

	A	B	C	D	E	F	G	H	K
Mean (Range)	38.72 (5.60-64.70)	33.20 (5.30-50.00)	32.63 (4.81-54.14)	34.44 (15.20-49.40)	34.20 (7.10-51.50)	35.61 (6.60-53.00)	35.27 (6.42-51.95)	38.91 (8.10-58.60)	35.19 (6.40-50.20)
TE (%)	7.54	2.27	5.54	6.75	5.86	7.07	7.66	10.30	3.23
TE (%)	19	7	17	20	17	20	22	26	9
TE_a (CLIA)	6	6	6	6	6	6	6	6	6
TE_a (F)	10	10	10	10	10	10	10	10	10

Table 12.10: HCT Feline

Table 12: Summary of calculated total errors (TE) and comparison to total allowable errors according to CLIA requirements (24) (TE_a (CLIA)) and recommendations published by Freeman et al. (20) (TE_a (F))

	A	C	D	E	F	G	H	K
Mean (Range)	269.9 (5.0-736.0)	309.3 (32.96-897.6)	307.1 (56.0-1093)	310.3 (17.0-825.0)	267.7 (5.0-771.0)	253.0 (19.0-658.0)	291.9 (53.0-770.0)	311.0 (45.0-882.0)
TE (10^9/l)	59.46	66.13	60.95	109.27	100.19	149.48	197.44	59.61
TE (%)	22	21	20	35	37	59	68	19
TE_a (CLIA)	25	25	25	25	25	25	25	25
TE_a (F)	na	na	na	na	na	na	na	na

Table 12.11: PLT Canine

	A	C	D	E	F	G	H	K
Mean (Range)	256.9 (34.0-896.0)	312.2 (19.21-1057)	422.2 (48.0-1084)	318.1 (5.0-730.0)	282.1 (3.0-950.0)	248.1 (7.0-780.0)	217.0 (73.0-1009)	262.9 (1.0-645.0)
TE (10^9/l)	60.48	80.73	131.77	151.07	79.04	146.26	258.75	51.18
TE (%)	24	26	31	47	28	59	120	19
TE_a (CLIA)	25	25	25	25	25	25	25	25
TE_a (F)	na	na	na	na	na	na	na	na

Table 12.12: PLT Feline

To verify the ability to correctly identify severe thrombocytopenic samples (platelet count equal or under 50x10^9/l), a comparison to platelet estimation on the blood smear is performed. The results are summarized in table 13. All systems missed some severe thrombocytopenias (table 13) but they all generated results under the lower limit of the reference range of 150x10^9/l. Except in one canine sample laboratory instrument B, as well as instruments C and K, displayed results within the reference range, and instrument E generated platelet counts within the reference range in two thrombocytopenic feline samples. It is important to consider that the number of canine and feline samples with severe thrombocytopenia is too low to draw any general conclusions.

Instruments	A	B	C	D	E	F	G	H	K
Dog	7 / 10	1 / 6	7 / 10	0 / 4	3 / 7	7 / 10	6 / 9	0 / 9	2 / 7
Cat	3 / 4	1 / 2	3 / 4	1 / 2	1 / 4	1 / 3	1 / 3	0 / 4	1 / 2

Table 13: Number of recognized severe thrombocytopenias / total number of severe thrombocytopenias; (severe thrombocytopenias count ≤ 50x10^9 platelets per µl)

Reticulocytes

Beside the two laboratory analyzers only two point-of-care instruments can determine absolute reticulocyte counts in canine samples (C and D). Reticulocyte counts in the range of 0-647x10^9/l were identified with 126 samples ≥ 60x10^9/l. The correlation coefficients were fair for instruments C (r=0.789) and D (r=0.662). The $S_{y|x}$-values were between 25x10^9/l and 30x10^9/l for both instruments, indicating a moderate random error. Based on Bland-Altman analysis instrument C and D had a negative bias of -36.20x10^9/l and -62.10x10^9/l, respectively, mostly caused by a proportional systemic error and preventing a 95% limit of agreement determination.

In addition, feline reticulocytes counts could only be determined by laboratory instrument B, but not A, and exclusively point-of-care instrument C. A good correlation of absolute reticulocyte counts between laboratory and point-of-care instrument was found (r=0.862). The $S_{y|x}$-values was 20x10^9/L, the 95% limits of agreement was between -49x10^9/l and 47.50x10^9/l with a neglible bias (-0.75x10^9/l).

3.4.2 Differential blood cell count

Except for neutrophils/granulocytes the correlation between the instruments and the manual differential are weak, therefore the estimates of intercept and slope, obtained by Passing-Bablok regression are not accurate. Hence the data interpretation is based on Bland-Altman analyses and the ability to correctly identify clinically relevant stages (for example neutropenia). Because of the high imprecision of the reference method (see chapter 5) and the expected high imprecision with low values, the calculation of total errors is not advisable.

Neutrophils/Granulocytes (table 14.1, 14.2, 14.3 and 14.4)

The correlation coefficients for canine and feline samples equal or exceed 0.975 in all instruments except for the following: Instrument D in canine and feline samples, instrument G in canine and instruments A and F in feline samples. The biases range from -3.31x10^9/l to 2.14x10^9/l. Instrument E shows negative biases in both species (-1.67x10^9/l respectively -1.69x10^9/l). The Bland-Altman plots as well as the estimation of intercept and slope reveal that this is due to a proportional systematic error, with larger underestimation of higher values. The same type of error can be observed for instrument F in feline samples. Instrument G shows positive biases in both species (2.14x10^9/l [dog] and 1.10x10^9/l [cat]) caused by constant systematic errors. Both constant and proportional systematic errors cause the negative biases obtained by instrument D in canine samples. The same instrument shows also a clearly negative bias for feline samples. This is caused by few outliers (see below). For the other instruments no major biases are present.

For all instruments generating a 5-part-differential the ability to detect clinically relevant neutropenias and neutrophilias is verified and summarized in tables 15.1 and 15.2.

Instruments	A	B	C	D	E
Dog	5 / 6	4 / 5	7 / 8	0 / 2	3 / 4
Cat	14 / 18	8 / 9	16 / 18	3 / 6	6 / 8

Table 15.1: Number of recognized neutropenias / total number of samples with neutropenia, according to the manual differential (≤ 2.5x10^9/l)

Between 0 and 89% of the samples with neutropenia, according to the manual differential are correctly classified by the instruments (table 15.1). All instruments report values above the threshold of $2.5x10^9$/l in few samples with neutropenia in the manual differential, but the differences are minor and therefore usually acceptable. A notable exception is instrument C. In two canine and one feline samples the overestimation by the instrument is major with values between $3.83x10^9$/l and $4.13x10^9$/l. Clear overestimation of neutrophils is identified in one canine sample for instrument E ($4.19x10^9$/l instead of $2.23x10^9$/l obtained by manual differential). All instruments also generate a few false neutropenias, but in most samples the discrepancies with the manual differential are minor. However, in three feline samples instrument A displays values between $0.10x10^9$/l and $2.20x10^9$/l, when the manual differential obtained values between $6.35x10^9$/l and $7.48x10^9$/l. Instrument C clearly underestimates the value in one feline sample ($2.11x10^9$/l instead of $7.17x10^9$/l). And instrument D shows discrepant results in two canine and two feline samples, with values between $0.74x10^9$/l and $1.60x10^9$/l, when the manual differential displays values between $7.3x10^9$/l and $16.67x10^9$/l. Instruments A and D display error messages, related to the differential in all the samples described above. Instrument C displayed an error message associated to the differential in one feline samples. The remaining samples of instrument C and E are not flagged.

Between 44 and 97% of the samples with neutrophilia, according to the manual differential, are correctly classified by the instruments (table 15.2). In a few samples results under $20.0x10^9$/l are generated by the instruments, but in all instruments except one the differences are only minor. In four feline samples instrument D generates values clearly dissimilar to the manual differential: the instrument reports values between $3.72x10^9$/l and $11.37x10^9$/l whereas the manual differential reveals values between $20.52x10^9$/l and $29.17x10^9$/l. Two of the four samples are flagged.

Instruments	A	B	C	D	E
Dog	36 / 37	29 / 30	43 / 48	17 / 20	23 / 28
Cat	14 / 15	6 / 9	9 / 10	4 / 9	4 / 7

Table 15.2: Number of recognized neutrophilias / total number of samples with neutrophilia, according to the manual differential (> $20.0x10^9$/l)

Eosinophils (table 14.5 and 14.6)

In both species and all instruments except instrument B the biases are very close to zero ranging from -0.22x10⁹/l to 0.29x10⁹/l. The 95% limits of agreement are relatively narrow. Laboratory instrument B shows a bias of 1.03x10⁹/l in feline samples, associated with very wide 95% limits of agreement (-2.04x10⁹/l to 4.1x10⁹/l). This positive bias is caused by 22 feline samples, for which the instrument generates values between 1.85x10⁹/l and 8.25x10⁹/l, whereas the manual differential reports values below 1.7x10⁹/l. That means instrument B is producing a high number of false eosinophilias in feline samples.
Concerning the ability to detect eosinophilias no conclusive statement can be made with our data, as only a few eosinophilic samples are available. But it is remarkable that in all 9 samples with values above 1.5x10⁹/l (ranging from 1.70x10⁹/l to 4.55x10⁹/l) according to the manual differential, instrument C produces values below 0.93x10⁹/l. Taking the 95% confidence limits for 200-cell-manual-differentials (43) into account, 4 of the 9 samples would continuously be classified as eosinophilic. In 2 of these 4 samples the scattergrams show cell populations located in the region of eosinophils. But the instrument classified these populations as neutrophils (figure 20).

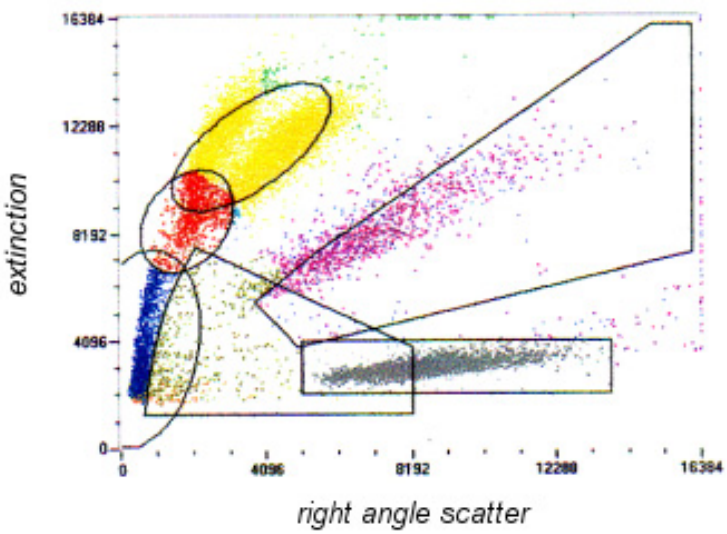


Figure 20: Scattergrams of instrument C for the white blood cells: There are two yellow populations present; the populations on the upper left corner are the eosinophils, but misclassified as neutrophils.

In contrast, instrument E reports eosinophil counts above 1.5x10^9/l in 16 canine samples, when the manual differential generates values below 1.46x10^9/l. Considering the 95% CI for 200-cell-manual-differentials (43), in 8 of the 16 samples, the instrument still clearly overestimates the values, leading to false eosinophilias. In 2 of the 8 samples the instrument reports values of 6.49x10^9/l and 7.20x10^9/l.

Lymphocytes (table 14.7 and 14.8)

In both species the biases range from -1.12x10^9/l to 2.74x10^9/l. The 95% limits are relatively wide und vary between the instruments. The differences between the upper and lower limit are between 3.45x10^9/l and 21.54x10^9/l. In-house-instrument E shows positive biases in both species (1.77x10^9/l in canine and 1.19x10^9/l in feline samples), associated with relatively wide 95% limits of agreement. The Bland-Altman plots indicate that low lymphocyte counts are overestimated by instrument E. Thirty-one canine samples with lymphopenia (below 1.0x10^9/l) according to the manual differential are analyzed using instrument E. In 29 of these samples instrument E generates values above 1.0x10^9/l; in 23 of these samples, the values are higher than 2.0x10^9/l (maximum value of 12.05x10^9/l). A similar phenomenon is obvious in feline samples for instrument E, as well as in canine and feline samples analyzed by instrument D. The range of values obtained by instrument D includes values above 15x10^9/l, when the manual differential reveals values below 1.0x10^9/l (dog) and 1.5x10^9/l (cat). As a result of this misinterpretation instrument D also shows clearly positive biases in both species. In contrast, instrument G obtains negative biases in both species (-1.04x10^9/l in canine and -1.12x10^9/l in feline samples). In the Bland-Altman plots it is obvious that the negative biases are caused by proportional systematic errors, with larger underestimation of higher values. Hence in 76 canine and 27 feline samples the instrument generates values below the threshold for lymphopenia, when lymphopenia is not obtained by the manual differential. In 21 of the 76 canine samples, the manual differential lymphocyte counts are above 2.0x10^9/l. The other instruments (A, B, C, F, and H) also misclassified some lymphopenic samples, according to the manual differential. In a few samples the differences between instruments and manual differential are major (maximum values of 15.04x10^9/l compared to 0.78x10^9/l).

Monocytes (table 14.9 and 14.10)

The biases are between -0.79x10^9/l and 0.69x10^9/l, except for instrument C in both species and instrument F in felines. Their biases are around 1.20x10^9/l, due to constant and proportional systematic errors. The overestimation is larger at higher values. In canine samples, instruments A, E, G and H show a clear relationship between differences and averages in Bland-Altman plots: the underestimations increase as the averages increase.

Table 14: Results of the statistical analysis for each instrument in comparison to the manual differential

	A	B	C	D	E	F	G	H
N	243	213	257	88	172			
r	0.987	0.992	0.989	0.958	0.978			
Slope	1.03	0.95	0.92	0.88	0.83			
95% CI	1.01 to 1.06	0.93 to 0.97	0.90 to 0.93	0.82 to 0.94	0.82 to 0.86			
Intercept	0.04	0.16	0.36	0.96	0.53			
95% CI	-0.10 to 0.20	0.00 to 0.31	0.16 to 0.49	0.42 to 1.36	0.33 to 0.72			
Bias	0.63	-0.48	-0.70	-1.50	-1.67			
95% Limits of agreement	-3.49 to 4.75	-3.08 to 2.12	-4.31 to 2.91	-8.93 to 5.92	-7.27 to 3.93			

Table 14.1: Neutrophils Canine

	A	B	C	D	E	F	G	H
N	243	213	257	88	172	242	243	235
r	0.989	0.992	0.990	0.958	0.979	0.975	0.974	0.987
Slope	1.03	0.97	0.93	0.87	0.87	0.91	1.04	0.99
95% CI	1.01 to 1.04	0.95 to 0.98	0.92 to 0.95	0.82 to 0.92	0.85 to 0.90	0.88 to 0.94	1.02 to 1.07	0.97 to 1.01
Intercept	-0.10	0.12	0.10	0.89	0.35	0.26	1.06	0.27
95% CI	-0.23 to 0.08	-0.30 to 0.30	-0.02 to 0.24	0.39 to 1.39	0.15 to 0.58	0.00 to 0.53	0.81 to 1.36	0.06 to 0.53
Bias	0.52	-0.35	-0.69	-1.52	-1.29	-0.97	2.14	0.22
95% Limits of agreement	-3.43 to 4.48	-2.97 to 2.26	-0.89 to -0.48	-2.31 to -0.74	-6.03 to 3.44	-1.30 to -0.64	1.81 to 2.47	-0.01 to 0.44

Table 14.2: Granulocytes Canine
N: Number of samples; r: correlation coefficient (linear regression); Slope and Intercept (Passing-Bablok regression)

Table 14: Results of the statistical analysis for each instrument in comparison to the manual differential

	A	B	C	D	E	F	G	H
N	103	76	105	39	68			
r	0.973	0.979	0.985	0.764	0.986			
Slope	1.04	0.92	0.91	0.83	0.88			
95% CI	1.01 to 1.07	0.88 to 0.96	0.89 to 0.94	0.78 to 0.88	0.71 to 1.00			
Intercept	-0.09	0.20	-0.07	-0.16	0.55			
95% CI	-0.30 to 0.08	-0.03 to 0.48	-0.35 to 0.03	-0.45 to 0.27	-0.73 to 1.40			
Bias	-0.07	-0.80	-0.97	-3.17	-1.69			
95% Limits of agreement	-4.51 to 4.38	-4.47 to 2.87	-1.30 to -0.64	-5.15 to -1.20	-2.16 to -1.21			

Table 14.3: Neutrophils Feline

	A	B	C	D	E	F	G	H
N	103	76	105	39	68	90	105	94
r	0.978	0.989	0.984	0.762	0.986	0.967	0.990	0.988
Slope	1.04	1.01	0.94	0.86	0.85	0.86	1.01	1.01
95% CI	1.01 to 1.06	0.99 to 1.04	0.92 to 0.98	0.71 to 0.98	0.81 to 0.89	0.81 to 0.91	0.99 to 1.04	0.98 to 1.05
Intercept	-0.03	0.09	-0.11	0.62	0.15	0.16	0.67	0.34
95% CI	-0.32 to 0.15	-0.17 to 0.25	-0.36 to 0.08	-0.65 to 1.15	-0.23 to 0.55	-0.15 to 0.50	0.39 to 0.90	-0.10 to 0.72
Bias	0.07	0.27	-0.61	-3.31	-1.45	-1.47	1.10	0.83
95% Limits of agreement	-0.33 to 0.46	-2.37 to 2.90	-4.01 to 2.79	-15.40 to 8.78	-5.57 to 2.68	-2.04 to -0.90	0.81 to 1.39	0.51 to 1.14

Table 14.4: Granulocytes Feline
N: Number of samples; r: correlation coefficient (linear regression); Slope and Intercept (Passing-Bablok regression)

Table 14: Results of the statistical analysis for each instrument in comparison to the manual differential

	A	B	C	D	E	F	G	H
N	243	213	257	88	172			
r	0.792	0.816	0.10	0.577	0.266			
Bias	-0.22	0.10	-0.06	-0.06	0.24			
95% Limits of agreement	-0.93 to 0.49	-0.52 to 0.72	-1.28 to 1.17	-0.97 to 0.86	-1.66 to 2.14			

Table 14.5: Eosinophils Canine

	A	B	C	D	E	F	G	H
N	103	76	105	39	68			
r	0.883	0.111	0.646	0.142	0.712			
Bias	0.03	1.03	0.29	-0.15	0.13			
95% Limits of agreement	-0.45 to 0.50	-2.04 to 4.10	-0.60 to 1.18	-1.03 to 0.72	-0.73 to 0.99			

Table 14.6: Eosinophils Feline
N: Number of samples; r: correlation coefficient (linear regression)

Table 14: Results of the statistical analysis for each instrument in comparison to the manual differential

	A	B	C	D	E	F	G	H
N	243	213	257	88	172	242	243	235
r	0.787	0.667	0.792	0.080	0.293	0.427	0.676	0.451
Bias	-0.33	0.28	-0.27	1.79	1.77	0.51	-1.04	0.87
95% Limits of agreement	-2.05 to 1.40	-2.43 to 2.99	-2.09 to 1.54	-5.52 to 9.09	-3.64 to 7.18	-3.35 to 4.36	-3.59 to 1.51	-3.24 to 4.99

Table 14.7: Lymphocytes Canine

	A	B	C	D	E	F	G	H
N	103	76	105	39	68	90	105	94
r	0.796	0.871	0.769	0.190	0.656	0.811	0.876	0.779
Bias	-0.58	-0.31	-0.55	2.74	1.19	0.03	-1.12	-0.74
95% Limits of agreement	-3.26 to 2.11	-2.58 to 1.97	-3.36 to 2.27	-8.03 to 13.51	-2.19 to 4.56	-2.50 to 2.55	-4.05 to 1.80	-3.98 to 2.50

Table 14.8: Lymphocytes Feline
N: Number of samples; r: correlation coefficient (linear regression)

Table 14: Results of the statistical analysis for each instrument in comparison to the manual differential

	A	B	C	D	E	F	G	H
N	243	213	257	88	172	242	243	235
r	0.731	0.786	0.903	0.770	0.640	0.686	0.183	0.790
Bias	-0.03	-0.03	1.20	0.04	-0.29	0.69	-0.92	-0.79
95% Limits of agreement	-2.77 to 2.71	-2.38 to 2.31	-1.53 to 3.93	-3.45 to 3.54	-2.34 to 1.76	-2.35 to 3.74	-4.81 to 2.97	-3.80 to 2.22

Table 14.9: Monocytes Canine

	A	B	C	D	E	F	G	H
N	103	76	105	39	68	90	105	94
r	0.349	0.686	0.687	0.420	0.417	0.427	0.434	0.202
Bias	0.52	-0.08	1.19	0.65	0.33	1.22	0.07	-0.04
95% Limits of agreement	-2.38 to 3.41	-0.92 to 0.77	-1.08 to 3.46	-1.17 to 2.46	-0.86 to 1.52	-1.34 to 3.77	-1.03 to 1.16	-1.23 to 1.16

Table 14.10: Monocytes Feline
N: Number of samples; r: correlation coefficient (linear regression)

4. Discussion

4.1 Accuracy

Purchasing an automated blood counting system is a large capital expense that carries with it maintenance, supply and personnel costs (44). Therefore the decision about the type of system to be used depends on the accuracy and various factors, including the number of samples to be analyzed, the desired parameters, costs and the ease of use.
The accuracy of a method should ideally be assessed in comparison to a reference method (15). The International Council for Standardization in Hematology (ICSH) defines accuracy as 'a measure of agreement between the estimate of value and the true value' (28). The true value must be obtained by a reference method and the only hematology parameters that can be estimated accurately are hemoglobin concentration, PCV, red and white cell counts and the differential leukocyte count (28). Alternatively the comparability can be assessed. This is defined as the ability of the instrument to produce results which agree satisfactorily with those obtained by procedures in routine use (28). The comparability is an accepted method to evaluate hematology instruments (28). But in the case of differences between the methods, it is necessary to identify which method is inaccurate (15). We have chosen one laboratory instrument as the comparative instrument for the CBC which is widely used and accepted in veterinary medicine.
In the past most method comparison studies in veterinary medicine focused on result interpretation using regression analysis, based on an objective classification scheme introduced more than 10 years ago (45). But different authors, as well as journal editorial boards recommend that the correlation coefficient cannot be used to judge the agreement between methods (15-17). Instead the r-value can only be used to judge the acceptability of the concentration range of the data being used to calculate the regression statistics (15). The decision on the acceptability of a method should be made, based on the observed total error. If the observed total error is smaller than the

medically allowable error, the method performance is acceptable (15). For human medicine several requirements and recommendations for allowable errors are available (21-27). These regulatory requirements are lacking in veterinary medicine. Therefore the assessment should be based on clinical quality requirements determined by clinical interpretation of the data (20). For the differential cell count, calculation of total errors seems inappropriate, as imprecision of the reference is extremely high. Furthermore, due to the nature of low values for the differential blood cell count, the standard deviations, obtained in the replication experiment are high as well, leading to high total errors.

To my knowledge this is the first study, applying total errors in veterinary hematology. For the cell counts we calculated the total errors as recommended by Westgard (15), using the standard deviation observed in the replication study as an estimation of the random error. In general the $S_{y|x}$-values and 95% limits of agreement clearly exceed the standard deviations. The estimate of random error (standard deviation) obtained by the replication study likely represents the lowest random error, because it used a single sample from a healthy animal with parameters that are within reference intervals. The $S_{y|x}$ and 95% limits of agreement which also provide an estimation of random error are based on a range of samples from healthy and diseased animals and likely more accurately reflect the random error achieved in clinical practice populations. Furthermore, our comparison method also includes some amount of error. Calculating a single value for total error with data covering the whole range is not ideal. The goal should be to calculate total errors for each medical decision level and perform separate replication experiments for every decision level. However the CLIA requirements, as well as the recommendations for veterinary medicine, published by Freeman et al. do not distinguish between different medical decision levels.

Based on the recommendations for veterinary medicine (20), the accuracy of the evaluated instruments is acceptable for the white blood cell counts within reference range, as the slightly higher total error of instrument K with feline samples is unlikely to be significant. Regarding the red blood cell parameters, some instruments exceed the recommended total allowable errors for some or all red blood cell parameters. Therefore, using the recommended total allowable errors as the basis, the accuracy of the instruments is not acceptable for red cell parameters. However, in-house-instruments should not be expected to have the same performance capability as larger instruments or reference laboratory instruments. Furthermore, the given recommendations are based on quality control material, and

morphologically alterations like in blood samples from diseased animals are not expected. Therefore alterations in total allowable error requirements are appropriate and users need to be aware of those parameters that are not measured as accurately by in-house-instruments. If we imagine, that all evaluated instruments meet the total allowable errors for red cell parameters (except MCV for instrument H), the following total allowable errors would be needed: RBC 20%, Hb 20%, MCV 10% and HCT 25%. With the following red blood cell parameters, RBC $6.5x10^{12}$/l, Hb 13g/dl, MCV 65fl and HCT 42%, the maximal calculated deviations (calculated with the defined maximum total allowable errors) would be: RBC $6.5x10^{12}$/l ± 1.3, Hb 13g/dl ± 2.6, MCV 65fl ± 6.5 and HCT 42% ± 10.5. In this example the deviation for hematocrit is major and of clinical relevance. As the hematocrit is calculated by the RBC and the MCV, the high total errors of hematocrit are the results of the errors in these two parameters. Instrument A achieved a relatively low total error for hematocrit in canine samples (7%), whereas the total error in feline samples is major and not acceptable (19%). This is caused by elevated total errors in MCV, as well as in RBC. The summations of the total errors for MCV and RBC are also the cause for the unacceptable high total errors for hematocrit achieved by instruments E, F, D and G. Instrument C instead shows very little total errors for MCV, whereas the total errors for RBC are unacceptable high (due to random error) resulting in elevated total errors for hematocrit. In contrast instrument H shows very little total errors for RBC, but the total errors for MCV are high, due to strong constant systematic error. As most veterinary clinicians use hematocrit for the assessment of the red blood cell parameters, and the total errors for hematocrit are mostly unacceptable high we recommend further verification, by performing spun PCVs or by verifying the accuracy of the hematocrit by comparison with hemoglobin values (hematocrit in % should be about three times the hemoglobin concentration in g/dl). For example, in canine samples analyzed in our study the regression and Bland-Altman analysis between Hb times three and HCT reveals correlation coefficients ≥ 0.89 and biases between -4.74 and 3.33. Only for instruments B and K the total errors for hematocrit are acceptable in both species.

For instruments A, C, D and K the total errors for platelet counts are acceptable compared to laboratory instrument B. For the other instruments (E, F, G and H) the total errors clearly exceed the CLIA recommendations (24) and the values, proposed by Freeman et al. (20). But it should be considered, that this may also be due to imprecision and inaccuracy of the comparative method, especially as the samples were 24 hours old when measured on instrument B (34).

For all instruments and all cell count parameters additional studies are needed to characterize calculated total errors at medical decision limits outside of reference interval.

The result analysis and the discussion above are based on the interpretation of total allowable errors. As previously indicated, a classification scheme for correlation coefficients is used traditionally to assess accuracy of laboratory instruments (45). Applying this classification scheme to the results obtained for hematocrit, remarkable differences concerning the interpretation are obvious. Except for instrument H and D in feline samples, all correlation coefficients are above 0.93 and can be classified as excellent. In contrast applying total allowable errors, instruments B and K are the only systems with acceptable results in both species. Calculating correlation coefficients in method comparison is appropriate, but the way of result interpretation needs to change. A correlation coefficient can only be used to assess the acceptability of the data range. It can not be used, to judge if an instrument is performing well. Hence, if total errors are not calculated, data interpretation has to focus on intercept and slope (in case of high r-values) and Bland-Altman analysis.

Regarding the differential, r-values above 0.975 are only obtained for neutrophils/ granulocytes. Therefore for the parameters of the white blood cell differential regression analysis are only appropriate for the neutrophils/granulocytes. For all other parameters the r-values are low. This does not indicate that the instruments are not performing well, but it indicates that the concentration range of our data is not sufficiently large to perform linear regression. Hence, the judgment of acceptability is based on bias, 95% limits of agreement and the ability to recognize pathological stages.
A 200-cell-manual-differential was used as the reference method. This is a somewhat inaccurate and irreproducible method that is subject to errors that cannot be totally eliminated (46). These errors include human inconsistency in cell interpretation, as well as inconsistent cell distribution on the blood film and high imprecision of the manual differential (43). Nevertheless, the National Committee for Clinical Laboratory Standards (NCCLS) recommends the manual differential as the reference method for the evaluation of automated differential cell counts (47). But inaccuracy and imprecision of the reference method need to be considered when interpreting the results obtained by the instruments. Slight deviations above or below the thresholds of pathological stages are not necessarily indicative of poor performance.

In general, the biases obtained for neutrophils or granulocytes are acceptable. In the result section, there are few samples of different instruments with discrepant results in comparison to the manual differential described. Although the differences in these samples are clinical relevant and sole interpretation of the numerical reports would cause clinical misinterpretation, the current expectations in hematology analyzers are not to completely substitute the manual evaluation of blood smears. And in face of different species and morphological changes of cells, discrepant results are somehow expected and natural. Worldwide, the majority of clinical pathologists agree that even with large hematology analyzers the differential blood cell count cannot be used for clinical interpretation without further verification (ASVCP (American Society of Veterinary Clinical Pathology) list serve, 2006). Further verification can be realized in different ways: interpretation of error messages, evaluation of scattergrams and histograms and evaluation of the blood smear. The discussion on the ASVCP list serve in 2006 clearly revealed, that in many large clinical pathology laboratories using advanced hematology instruments blood smears from all samples are reviewed for possible discrepancies. Hence, it would be exaggerated to expect perfect accuracy from small in-house-hematology instruments. Instead the possibilities to verify the numerical results should be improved, for example by offering the scattergrams of the laser-based instruments to the customers and most important by improving flag algorithms (see later). An example, emphasizing the importance of blood smear evaluation is displayed in table 16. In the manual differential numerous medium-sized lymphoblasts are present.

Parameter (10^9/l)	A *	B *	C *	E	F *	G	H
WBC	23.80	31.22	32.72	27.49	29.5	30.87	30.85
Neutrophils/Granulocytes	22.50	15.69	14.88	24.38	25.5	30.20	28.41
Eosinophils	0.60	0.32	0.40	0.60			
Lymphocytes	0.70	8.56	13.78	1.79	1.4	0.60	1.73
Monocytes	0.0	1.08	3.52	0.71	2.60	0.1	0.71
Basophils	0.0	2.81	0.13	0.0			

*Table 16: Reported results of a canine sample with numerous lymphoblasts in the manual differential (Instruments marked with * displayed an error message associated to the differential)*

The only instruments reporting a lymphocytosis are instruments B and C. An error message is displayed by instruments A, B, C and F. Relying on the results reported by instruments E, G and H without further verification would cause a tremendous

clinical misinterpretation. All histograms of the impedance-based instruments show that they were not able to differentiate the white blood cells (figure 21).

Figure 21: Histograms of the impedance-based instruments for a canine sample with numerous lymphoblasts in the manual differential. In all histograms of the white blood cells it is obvious that no valley is present, and therefore the instruments are not able to distinguish between the different cell populations.

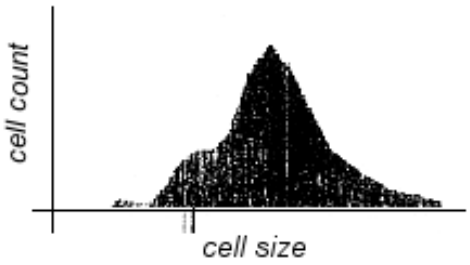


Figure 21.1: Histogram of instrument E

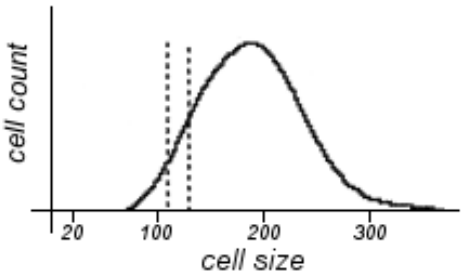


Figure 21.2: Histogram of instrument F

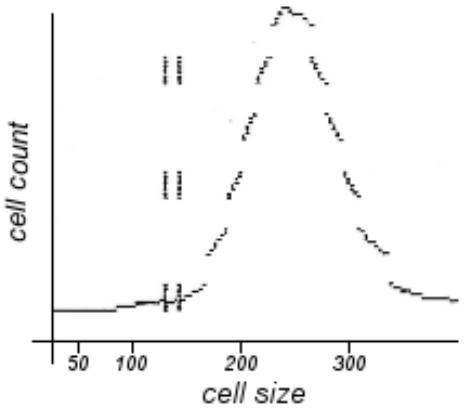


Figure 21.3: Histogram of instrument G

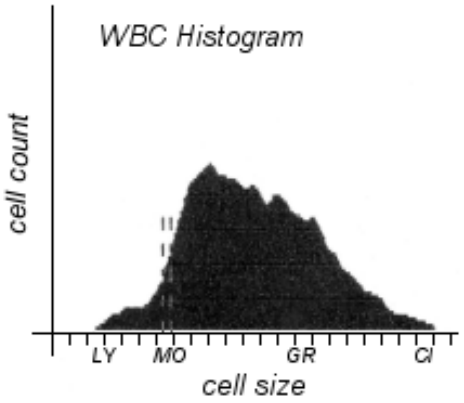


Figure 21.4: Histogram of instrument H

The scattergram of instrument C reveals changes which might be indicative of large abnormal cells, and can be detected by an experienced and trained user (figure 22): the blue population (lymphocytes) on the left side is enlarged, with numerous events scattering upwards. This is comparable to the scattergram of instrument B (figure 23) with events spreading into the large unstained cells (LUC) gate (34).

Figure 22: Scattergrams (time of flight versus right angle scatter) of instrument C (blue=lymphocytes; yellow=neutrophils; red=monocytes; grey=latex particles; pink=interferences)

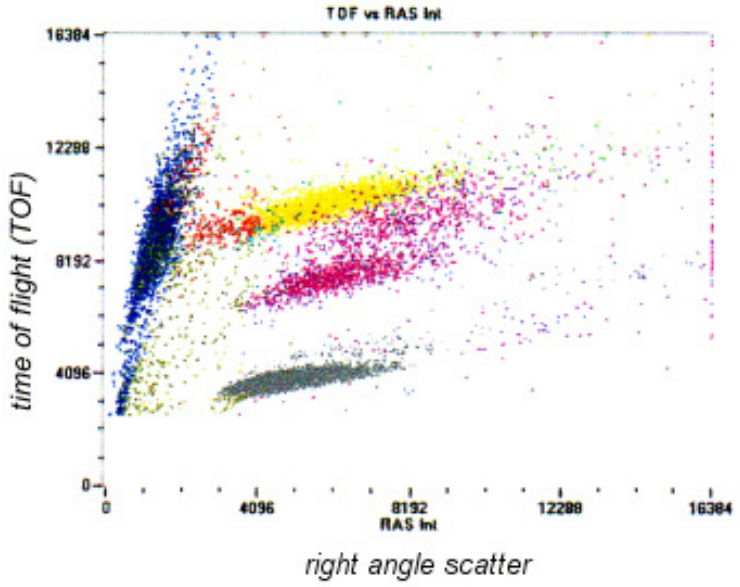


Figure 22.1: Scattergram of instrument C for a canine sample with numerous lymphoblasts in the manual differential

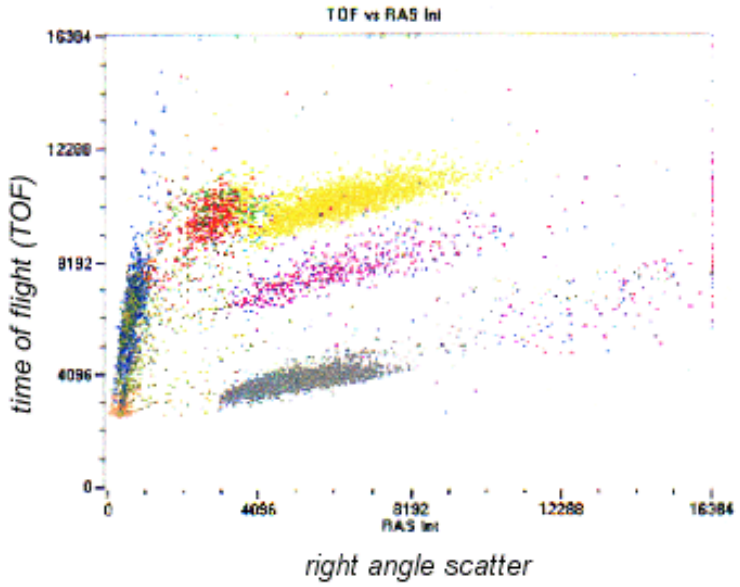


Figure 22.2: Scattergram of instrument C for a sample from a healthy dog

Figure 23: Scattergrams (Peroxidase channel) of instrument B (1=neutrophils gate; 2=monocytes gate; 3=lymphocyte gate; 4=eosinophil gate; 5=LUC gate)

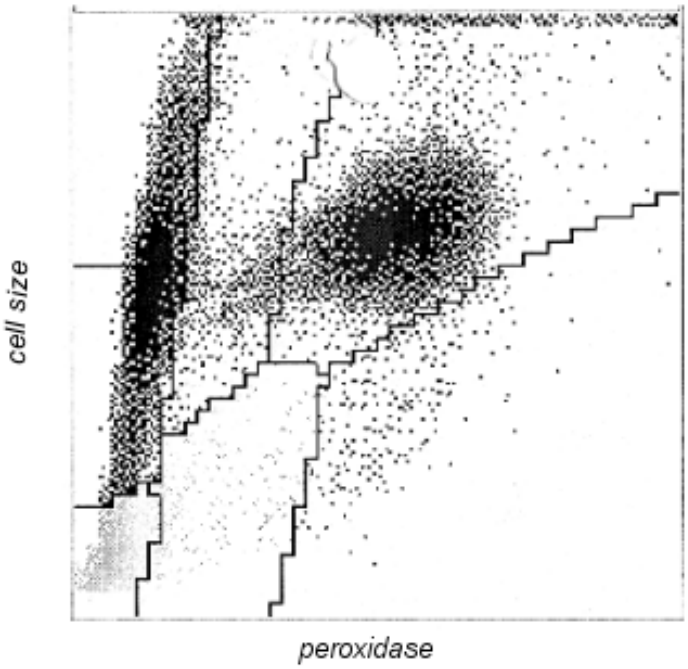


Figure 23.1: Scattergram of instrument B for a canine sample with numerous lymphoblasts in the manual differential

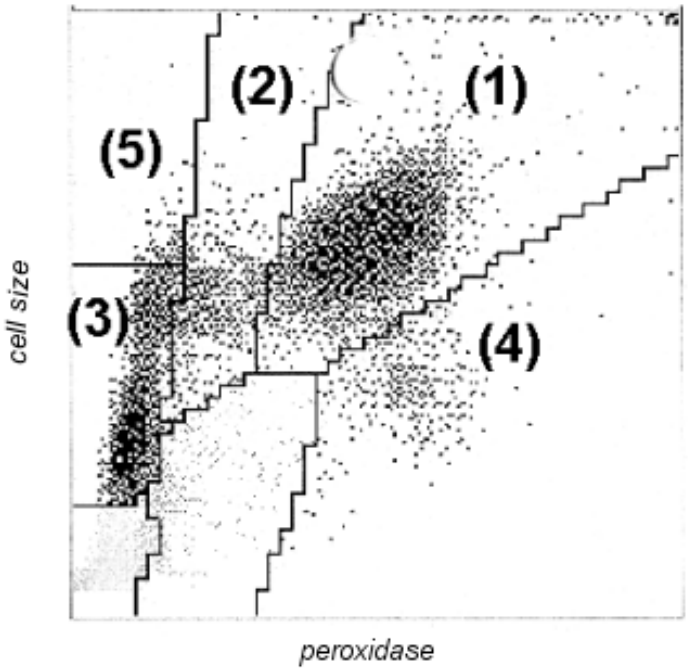


Figure 23.2: Scattergram of instrument B for a sample from a healthy dog

As indicated above, this sample shows that relying on the reported results without further verification can cause clinically important misinterpretations. The graphical reports offer the possibility to confirm the accuracy of the reported results, but cannot replace the review of a blood smear.

Although the number of samples with eosinophilia is relatively low in our study, we could show that the in-house instruments offering a 5-part differential have difficulties to accurately measure eosinophils. Instrument C is not able to determine high eosinophil counts in canine samples, and instrument E produces a high number of false eosinophilias in canine samples. Thus, with the software algorithms in use at the time of data collection, the advantage of 5-part differentials is not truly available to the users. And software modifications are likely needed to improve the eosinophil counts in these instruments. A high number of false eosinophilias is also obvious for laboratory instrument B in feline samples. But a previous study evaluating the accuracy of this instrument could show, that the eosinophil counts in feline samples increased up to 250% during blood storage at 4 °C (48).

4.2 Instrument flags

Flags are important information to alert users of hematology instruments that a system is malfunctioning or that morphological abnormalities of blood cells or software related algorithm issues are present. As indicated above, the high number of flagged samples for some parameters and some instruments and absence of flags in some cases with problematic results suggest that the error messages are of limited use. The NCCSL Document H20-A points out that flagging clinically abnormal sample for visual review is an integral part of performance evaluation for instrumental methods (47). Thereby different levels have been developed: Instruments which tabulate the usually circulating cells and flag for review any abnormal leucocytes or other variations from normal (cell count ↑,↓); Instruments which classify normal and abnormal leucocytes and are suitable for screening purposes; Instruments which classify normal and abnormal leucocytes and are suitable for diagnostic purposes (47). For veterinary medicine the only level which is probably achievable is the described level one: tabulation of the usually circulating

cells and flag for review any abnormal leukocytes or other variations from normal. As pointed out above, verification of the results generated by hematology analyzers is primarily done with evaluation of graphical reports and blood smears. Error messages as verification tools are not widely used and the most likely reasons are the inappropriate flagging algorithms.
Concerning the usefulness of the displayed flags, there is no difference between in-house and laboratory hematology instruments and the described level with tabulation of the usually circulating cells and flag for review any abnormal leukocytes or other variations from normal, might not be achievable in veterinary medicine. Although clear statements from the manufacturer how to interpret their error messages are needed. With this knowledge the users of the different instruments would be able to define their clinical approach how to interpret and verify the results of their hematology instruments. Based on the results of this study, I would recommend the following. Numerical results of every evaluated hematology instrument need further verification. For more advanced hematology instruments, like laboratory instruments A and B, the scattergrams should be used to distinguish if blood smear evaluation is needed or if the results can be released without further blood smear evaluation. Additional studies are needed to verify if the scattergrams of instrument C offer the same diagnostic value as the scattergrams of laboratory instrument B. Right now, the scattergrams are not routinely reported and available for the user. Comparison of the two different white blood cell counts (WOC versus WIC and WBCB versus WBCP) can also be used as verification tools for WBC counts of the laboratory instruments A and B. For impedance-based instruments which distinguish cells solely based on cell size, the displayed histograms offer not the same diagnostic value as the scattergrams. Although they can be used to verify if the instruments were able to distinguish the different cell population, blood smear evaluation is recommended for every sample. This does not mean that a manual differential has to be performed for every sample; it just means that a short evaluation, concerning the accuracy of the results, left shift and morphological changes is indicated. With more accurate error messages, they could also be used as verification tool, but with the results of this study this cannot be recommended, neither for in-house nor for laboratory hematology instruments.

4.3 Precision, Linearity and Carryover

The precision reflects the ability of an assay to get the same result if a sample is analyzed several times (52). The need for analytical precision is dependant on the degree of variation that can be accepted as a random variation (52). Allowable error recommendations for hematology parameters are available for human medicine (49). The results in our study mostly exceed these limits slightly, but are still comparable to those obtained in previous studies in veterinary medicine (34). The manufacturers of instrument A, D, F and G provide precision data in their manuals. For instrument F and G slightly higher values were obtained in our study than reported by the manufacturers.
The slightly higher coefficient of variation for the red blood cell count of instrument C might be explainable by the different technologies. Instrument C is the only in-house systems which analyzes the red blood cells with laser technology. Mild fluctuations regarding the noise level of the laser are expected. This may explain the slightly higher imprecision in comparison to the other systems where the red blood cells are measured with impedance technology.
In the tested ranges (see chapter 4.3) all hematology instruments achieved excellent linearity. Only instrument G was not able to measure high white blood cell counts precisely. When high white blood cell counts are obtained, the instrument displays a flag, indicating that the sample was analyzed three times but all three counts differed and were outside the system's precision limit. The manufacturer of instrument G gives a linearity range from 5 to 80x10^3/µl for the white blood cell count with a limit of +/-0.2 respectively +/-3%. Our study cannot confirm this linearity range; our data shows a linearity range from 0 to 45x10^3/µl.
Except for instrument H no hematology instrument shows a marked carryover of WBC, RBC, Hb and PLT. This is an important requirement for hematology systems to make sure that no cells are carried in the next patient analysis, leading to falsely elevated counts. Instrument H is not able to detect PBS, therefore level two was used instead of PBS. As indicated above, the high carryover was mostly caused by an imprecision in the repeated measurements of the two levels. Hence the achieved values for instrument H might not be comparable to those of the other instruments.

4.4 Ease of use

Regarding the ease of use, Instrument C has a significant longer cycle time than the other systems (table 2). But as the user can walk away while the instrument is analyzing the sample, the longer cycle time doesn't seem to be an important disadvantage, especially if the sample load does not exceed four samples per hour. The PC with the IDEXX VetLab® routinely connected to instrument C offers a lot more possibilities to evaluate and store the data. If an optional PC is not purchased the absence of the ability to store data is a clear disadvantage of instrument K. Except for instrument K all evaluated in-house-hematology systems require minimal maintenance and minimal hands-on time to analyze samples. Training should be offered by the manufacturers for all instruments, not only how to run samples but also how to maintain the instruments and how to handle flagged samples.
The ability to measure absolute reticulocyte counts on instruments C and D reduces the workload in clinical practices as manual reticulocyte counts are time consuming. It also provides the ability to more accurately classify, treat and monitor anemic patients. The MCV cannot replace the absolute reticulocyte counts to classify anemic patients as it is an average of the volume of all erythrocytes and therefore changes far later than reticulocytes occur. Also the total errors for MCV achieved by some of the impedance-based instruments are relatively high, and can cause clinical misinterpretation in anemic patients.

5. Summary

Objective

Compare the total blood cell counts and leukocyte differentials of seven in-house-hematology instruments and two laboratory systems.

Material and Methods

Over a three month period fresh K-EDTA anticoagulant blood samples from healthy and diseased dogs (n=260) and cats (n=110) were analyzed. Beside precision, linearity and carry-over, the accuracy was evaluated for each instrument. For the WBC, RBC and PLT concentrations, as well as for the MCV and Hb concentration laboratory instrument B (ADVIA ™120) was used as comparative method; the accuracy of the HCT was assessed in comparison to spun-PCV. A 200-cell manual differential was used as reference method for the leukocyte differential.

Statistics

For all parameters linear regression, Deming regression (in case of normally distributed measurement errors) or Passing-Bablok regression (in case of not normally distributed measurement errors) and Bland-Altman analysis was performed. The results of Deming regression or Passing-Bablok regression were used

to assess type of occurring errors in case of r-values ≥ 0.975. In case of r-values < 0.975, the Bland-Altman analysis was solely used for categorization of errors. For the cell counts, total errors were additionally calculated and compared to requirements in human medicine and recommendations in veterinary medicine. For the leukocyte differential, the ability to correctly identify pathological stages was also assessed.

Results

Canine and feline samples were analyzed separately. In comparison to laboratory instrument B, the biases for WBC counts ranged from -0.6x10^9/l to 2.4x10^9/l, for RBC counts from 0 to 0.9x10^{12}/l, for Hb from -1.5g/dl to +0.7g/dl, for MCV from -4.3fl to +8.3fl, for HCT versus PCV values from 0.1% to 7.2% and for PLT counts from -69.3x10^9/l to +77.2x10^9/l. Calculation of the total errors revealed, that for the white blood cell count all instruments met the CLIA requirements (Clinical Laboratory Improvement Amendments) for human medicine. Only instrument K achieved slightly higher total errors. For the red blood cell parameters, the results vary between the instruments and the parameters. Especially for HCT the total errors exceeded CLIA requirements and recommendations for veterinary medicine. Thus, verification of the results is recommended, and can be achieved by the comparison of HCT and Hb concentration or a spun-PCV. Concerning platelet counts, only instruments A, C, D and K met or slightly exceeded the CLIA requirements. Only two of the in-house-instruments are able to perform reticulocyte counts. The results were overall acceptable, although both instruments showed a negative bias in comparison to laboratory instrument B. Regarding the leukocyte differential, the biases for neutrophils/granulocytes ranged from -3.31x10^9/l to 2.14x10^9/l, for eosinophils from -0.22x10^9/l to 1.03x10^9/l, for lymphocytes from -1.12x10^9/l to 2.74x10^9/l and for monocytes from -0.92x10^9/l to 1.22x10^9/l. For in-house and laboratory instruments some samples are present, where the results differed significantly from the manual differential. The in-house instruments, providing a 5-part differential have difficulties in detecting eosinophils in canine samples.
Some of the instruments flagged a high percentage of samples with error messages,

including samples with accurate results in comparison to the different comparative methods. Additionally in most of the instruments samples with inaccurate results, which are not marked with an error message are present.

Conclusion

Overall the data shows, that in general a good agreement was achieved between the hematology instruments and the different comparative methods. For most total blood cell counts and RBC parameters point-of-care analyzers performed similarly well as their large laboratory counterparts. The correlation for total WBC counts was better than for RBC counts and related parameters. The determination of the PLT counts is marginal for both in-clinic and laboratory methods. The ability to measure canine and feline reticulocyte is a clear advantage of the laboratory and the two laser-based point-of-care instruments to better assess erythroid regeneration. Although the reference method for the leukocyte differential has some clear limitations, the data shows that in general good agreement was achieved by the in-house and laboratory instruments. But accepting the results of the differential without further verification is not reasonable and can cause clinically important misinterpretation. Comparable to the total blood cell counts, there is no difference obvious between laboratory and in-house-instruments and between laser- and impedance-based systems with regard to accuracy of the numerical results of the automated differential cell counts. The flags generated by the various instruments appear only marginally helpful since there were many false flags and unflagged samples with inaccurate results. While in-clinic and laboratory hematology analyzers provide many useful hematological parameters and seem to perform similarly well, there remains a need for result verification. Beside brief blood smear evaluation, the scattergrams generated by laboratory instruments A and B offer important information and further studies are necessary to investigate if the scattergrams reported by instrument C offer the same amount of information. Accurate display of flags to identify those samples requiring review of the peripheral blood smear would be an important improvement. Based on the results of our study we recommend reviewing a blood smear for every sample, regardless of whether or not a flag is displayed.

6. Zusammenfassung

Ziel der Studie

Ziel der Studie ist der Vergleich von sieben In-Haus-Hämatologiegeräten und zwei Laborgeräten hinsichtlich Blutzellzählung und Differentialblutbild.

Material und Methoden

Über einen Zeitraum von drei Monaten wurden frische K-EDTA antikoagulierte Blutproben von gesunden und kranken Hunden (n=260) und Katzen (n=110) untersucht. Neben Präzision, Linearität und Verschleppung wurde für jedes Instrument die Übereinstimmung mit einer Vergleichsmethode verifiziert. Als Vergleichsmethoden kamen Laborgerät B (ADVIA™ 120) für Leukozyten-, Erythrozyten und Thrombozytenzahl sowie für MCV und Hämoglobin-Konzentration zum Einsatz. Ein Mikrohämatokrit (PCV) stellte die Referenzmethode für den HCT dar. Für die Leukozytendifferenzierung wurde ein manuelles 200-Zell Differentialblutbild verwendet.

Statistik

Für alle Parameter wurden lineare Regressionen, Deming Regressionen (bei normalverteilten Daten) oder Pasing-Bablok Regressionen (bei nicht normalverteilten

Daten) berechnet sowie eine Bland-Altman-Analyse durchgeführt. Bei r-Werten ≥ 0.975 wurden die Ergebnisse der Deming oder Passing-Bablok Regression zur Klassifikation vorliegender Fehler, interpretiert. Bei r-Werten < 0.975 erfolgte die Beurteilung ausschließlich anhand der Bland-Altman Analyse. Für die Blutzellzählung wurden zusätzlich »total errors« berechnet und die Ergebnisse mit Anforderungen aus der Humanmedizin und Empfehlungen aus der Veterinärmedizin verglichen. Das Differentialblutbild wurde ferner anhand der Fähigkeit pathologische Veränderungen zu erkennen, beurteilt.

Ergebnisse

Die Analyse von Hunde- und Katzenproben erfolgte separat. Im Vergleich zu Instrument B erzielten die Geräte folgende Mittelwertsabweichungen (Bias): für WBC zwischen -0.6x10^9/l und 2.4x10^9/l, für RBC zwischen 0 und 0.9x10^{12}/l, für Hb zwischen -1.5g/dl und +0.7g/dl, für MCV zwischen -4.3fl und +8.3fl, für HCT im Vergleich zu PCV Werten zwischen 0.1% und 7.2% und für PLT zwischen -69.3x10^9/l und +77.2x10^9/l. Die Berechnung der »total errors« zeigte, dass für die Leukozytenzahlen, die humanmedizinischen Anforderungen der CLIA (Clinical Laboratory Improvement Amendments) von allen Geräten erfüllt wurden. Nur Instrument K erreichte etwas höhere Werte in beiden Spezies. Für die Parameter des roten Blutbildes variierten die Ergebnisse je nach Gerät und Parameter. Vor allem die »total errors« für HCT überschritten die CLIA Anforderungen und eine Verifikation der Ergebnisse ist angeraten. Dies kann durch einen Vergleich von Hb und HCT Werten oder durch einen Mikrohämatokrit erfolgen. Hinsichtlich der Thrombozytenzahl erreichten nur Instrumente A, C, D, und K Werte unter oder nur geringgradig über den CLIA-Anforderungen. Nur zwei der In-Haus-Hämatologiegeräte sind in der Lage absolute Retikulozytenzahlen zu bestimmen. Obgleich beide Geräte einen deutlich negativen Bias im Vergleich zu Instrument B aufweisen, sind die Ergebnisse akzeptabel. Für das Differentialblutbild liegen die Mittelwertsabweichungen zwischen -3.31x10^9/l und 2.14x10^9/l für neutrophile Granulozyten/Granulozyten, zwischen -0.22x10^9/l und 1.03x10^9/l für eosinophile Granulozyten, zwischen -1.12x10^9/l und 2.74x10^9/l für Lymphozyten und zwischen

-0.92x10^9/l to 1.22x10^9/l für Monozyten. Sowohl die In-Haus Hämatologiegeräte, als auch die Laborgeräte gaben für einige Proben Ergebnisse aus, die sich signifikant vom manuellen Differentialblutbild unterschieden. Die In-Haus Geräte, die ein 5-Zell-Differentialblutbild bestimmen, haben Schwierigkeiten eosinophile Granulozyten beim Hund korrekt zu identifizieren.
Einige der Instrumente markieren einen hohen Prozentsatz der Messungen mit Fehlermeldungen. Dabei sind Proben mir akkuraten Ergebnissen im Vergleich zu den Vergleichsmethoden enthalten. Zusätzlich gilt für die meisten Geräte, dass einige Proben mit nicht akkuraten Ergebnissen vorhanden sind, die nicht mit einer Fehlermeldung markiert sind.

Schlussfolgerung

Insgesamt zeigen die vorliegenden Daten, dass eine gute Übereinstimmung zwischen den Hämatologiegeräten und den verschiedenen Vergleichsmethoden erzielt wurde. Für die meisten Parameter der Blutzellzählung und der Erythrozytenparameter sind die Ergebnisse der In-Haus Geräte vergleichbar zu denen der Laborgeräte. Die Korrelationen für die Leukozytenzahl waren besser, als die für die Erythrozytenzahl und Erythrozytenparameter. Die Bestimmung der Thrombozytenzahlen ist grenzwertig für In-Haus und Laborgeräte. Die Möglichkeit absolute Retikulozytenzahlen zu bestimmen, ist ein deutlicher Vorteile der Laser-basierten Instrumente, da es eine objektivere Beurteilung der erythroiden Regeneration erlaubt. Obgleich die Referenzmethode für das Differentialblutbild einige Limitationen enthält, zeigen die vorliegenden Daten, dass die In-Haus und Laborgeräte insgesamt eine gute Übereinstimmung erzielten. Allerdings ist die Akzeptanz der numerischen Daten ohne weitere Verifizierung nicht angemessen und kann zu klinisch wichtigen Fehlinterpretationen führen. Ähnlich wie für die Blutzellzählung, sind für die numerischen Ergebnisse des Differentialblutbildes keine Unterschiede zwischen In-Haus und Laborgeräten sowie zwischen Impedanz- und Laser-basierten Instrumenten offensichtlich. Die Fehlermeldungen, die von den verschiedenen Instrumenten generiert werden, sind wenig hilfreich, da zahlreiche Proben mit akkuraten Ergebnissen und Fehlermeldungen, sowie einige Proben ohne

Fehlermeldungen, jedoch mit inakkuraten Ergebnissen vorhanden sind. Obgleich In-Haus und Laborgeräte zahlreiche nützliche hämatologische Parameter bestimmen, ist eine Ergebnisvalidierung notwendig. Neben einer kurzen Evaluation des Blutausstriches, bieten die Scattergramme von Instrument A und B wertvolle Informationen und weitere Studien sind notwendig um festzustellen, ob die Scattergramme von Instrument C in gleichem Maße hilfreich zur Ergebnisvalidierung sind. Eine wichtige Verbesserung wäre akkurates Kennzeichnen von Proben, die eine Blutausstrichevaluation benötigen. Basierend auf den Ergebnissen dieser Studie empfehlen wird die Evaluation eines Blutausstriches für jede Probe, unabhängig davon ob eine Fehlermeldung vorhanden ist oder nicht.

7. Reference List

(1) Bienzle D, Stanton JB, Embry JM, Bush SE, Mahaffey EA
Evaluation of an in-house centrifugal hematology analyzer for use in veterinary practice
Journal of the American Veterinary Medical Association 2000; 217(8):1195-1200

(2) Guelfi JF, Trumel C
Comparative study of the hemogram: QBC VetAutoread versus manual methods
Revue de Medecine Veterinaire 1995; 146(11):765-770

(3) Hofmann-Lehmann R, Wegmann D, Winkler GC, Lutz H
Evaluation of the QBC-Vet Autoread hematology system for domestic and pet animal species
Comparative Hematology International 1998; 8(2):108-116

(4) Papasouliotis K, Cue S, Graham M, Sparkes AH, Gruffydd-Jones T
Analysis of feline, canine and equine hemograms using the QBC VetAutoread
Vet Clin Pathol 1999; 28(3):109-115

(5) Tasker S, Cripps PJ, Mackin AJ
Evaluation of methods of platelet counting in the cat
Journal of Small Animal Practice 2001; 42(7):326-332

(6) Wegmann D, HofmannLehmann R, Lutz H
Brief evaluation of the QBC-Vet Autoread-System
Tierarztliche Praxis 1997; 25(2):185-191

(7) Trumel C, Geffre A, Vergez F, Diquelou A, Bourges-Aballa N
Correlation study of Hemavet analyser (Drew) against Vet ABC (Scil) in dogs, cats, horses, cows and sheep
Proceedings 7th ESVCP meeting, Utrecht 2005, p 134

(8) Vergez F, Bourges-Aballa N, Diquelou A, Geffre A, Trumel C
Comparison of the results obtained with the Abacus Junior Vet and Vet ABC hematology analysers in dogs, cats, horses, sheep and pigs
Proceedings 7th ESVCP meeting, Utrecht 2005, p 131

(9) Schwendenwein I, Jolly M
Automated differentials by an impedance on site hematology system
Proceedings 3rd ESVCP meeting, Edinburgh 2001, p 21

(10) Pinches MDG, Burley K, Cue SM, Crawford EM, Papasouliotis K
Comparison of white blood cell differential counts determined by the in-house LaserCyte Hematology Analyser and manual method
Proceedings 7th ESVCP meeting, Utrecht 2005, p 107

(11) Gröndahl G, Bottema B, Lillihöök I
Comparison of standard hematologic parameters determined on an impedance-based point-of-care hematology analyser using floating discriminators and an automated hematology analyser commonly used in veterinary reference laboratories
Proceedings BSAVA congress, Birmingham 2004,

(12) Dewhurst EC, Crawford E, Cue S, Dodkin S, German AJ, Papasouliotis K
Analysis of canine and feline hemograms using the VetScan HMT analyser
J Small Anim Pract 2003; 44(10):443-448

(13) Papasouliotis K, Cue S, Crawford E, Pinches M, Dumont M, Burley K.
Comparison of white blood cell differential percentages determined by the in-house LaserCyte hematology analyzer and a manual method
Vet Clin Path 2006; 35(3):295-302

(14) Guija de Arespacochaga A, Schwendenwein I
Automatisierte Differenzialblutbilder mit einem Impedanz-Hämatologiesystem (ABACUS Junior VetTM, Diatron)
Tierarztliche Praxis 2005; 33:446-457

(15) Westgard JO
Basic Method Validation 2nd edition
Westgard QC 2003

(16) Hollis S
Analysis of method comparison studies
Ann Clin Biochem 1996; 33:1-4

(17) Jensen AL, Kjelgaard-Hansen M
Method comparison in the clinical laboratory
Vet Clin Path 2006; 35(3):276-286

(18) ISO/FDIS 15189
Quality management in the medical laboratory (2000), p 20ff

(19) Statland BE
Clinical Decision Levels for Laboratory Tests 2nd edition
Medical Economics Books 1987

(20) Freeman KP, Gruenwaldt J
Quality Control Validation in Veterinary Laboratories
Vet Clin Path 1999; 28 (4): 150-155

(21) Commission of Office Laboratory Accreditation (COLA)
Accreditation Manual (Columbia 1998)

(22) College of American Pathologists (CAP)
Laboratory Accreditation Program (Northfield 1998)

(23) Joint Commission on Accreditation of Healthcare Organizations (JCAHO)
Accreditation Manual for Pathology and Clinical Laboratory Services (Oakbrook Terrace 1998)

(24) U.S. Department of Health and Human Services
Regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA)
Fed Regist 1992; 57: 7002-7186

(25) Fraser CG, Petersen PH, Ricos C, Haeckel R
Proposed quality specifications for the imprecision and inaccuracy of analytical systems for clinical chemistry
Eur J Clin Chem Clin Biochem 1992; 30: 311-317

(26) Petersen PH, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, Thienpont L
Proposed guidelines for the internal quality control of analytical results in the medical laboratory
Eur J Clin Chem Clin Biochem 1996; 34: 983-999

(27) Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J, Perich C, Simon M
Current databases on biological variation: pros, cons and progress
Scand J Clin Lab Invest 1999; 59: 491-500

(28) International Council for Standardization in Hematology (ICSH)
Guidelines for the evaluation of blood cell analysers including those used for differential leucocyte and reticulocyte counting and cell marker applications
Clin Lab Haematol 1994; 16: 157-174

(29) Stöckl D, Dewitte K, Thienpont LM
Validity of linear regression in method comparison studies: is it limited by the statistical model or the quality of the analytical input data?
Clin Chem 1998; 44 (11): 2340-2346

(30) Westgard JO
A method evaluation decision chart (MEDx Chart) for judging method performance
Clin Lab Sci 1995; 8: 277-283

(31) Stockham S, Scott M
Fundamentals of Veterinary Clinical Pathology 1st edition, p 41
Blackwell Publishing 2002

(32) Abbott
Cell-Dyn 3500 System Operator's Manual, Chapter 3

(33) Kunicka J
Routine Use of Advanced Cell Analysis in the Hematology Laboratory
Lab Hematol 2005; 11(46)

(34) Moritz A
Der Einsatz lasergestützter Multiparameter-Hämatologiesysteme in der Veterinärmedizin
Habilitationsschrift Giessen 2000

(35) Idexx Laboratories
LaserCyte - Background / System Overview (2004)

(36) Melet Schloesing Pharmaceuticals
MS4-5 Operator's Manual

(37) Boule Medicalab
Heska CBC-Diff - User Manual, p 33ff

(38) Scil Animal Care Company
Vet ABC-Diff Hematology Analyzer - Operator's Manual, p 2.1ff

(39) Abaxis
VetScan H Operator's Manual, Chapter 7

(40) Idexx Laboratories
QBC Vet Autoread - Operator's Manual, p 0-1ff

(41) Jensen AL, Bantz M
Comparing laboratory tests using the difference plot method
Veterinary Clinical Pathology 1993; (46-48)

(42) Bland JM, Altmann DG
Statistical methods for assessing agreement between two methods of clinical measurement
Lancet 1986; 307-310

(43) Rumke CL
Variability of results in differential counts on blood smears
Triangle, Sandoz Journal of Medical Science 1960

(44) Feldmann BF, Zinkl JG, Jain NC
Schalm's Veterinary Hematology 5th edition, Chapter 1
Lippincott, Williams & Wilkins 2000

(45) Tvedten H.
The complete blood cell count and bone marrow examination. General comments and selected techniques
In Willard, Tvetden, Turnwald: Small Animal Clinical Diagnosis by Laboratory Methods 2nd edition
W B Saunders 1994

(46) Stiene-Martin EA, Lotspeich-Steininger CA, Koepke JA
Clinical Hematology: Priniciples, Procedures, Correlation 2nd edition
Lippincott, Williams & Wilkins 1998

(47) Koepke JA, Bentley SA, Pierre R.V, Richardson-Jones A, Ross DW, Saunders AM, Simson E, van Assendelft OW, Wilding P
Reference Leukocyte Differential Count (Proportional) and Evaluation of Instrumental Methods, NCCLS Document H20-A (1992)

(48) Meyer K
Softwareadaptation und Evaluation des ADVIA 120 bei der Tierart Katze
Dissertation Giessen 2005

(49) Burtis C, Ashwood ER
Tietz Fundamentals of Clinical Chemistry 5th edition, Chapter 13
W B Saunders 2001

(50) Roleff S, Arndt G, Bottema B, Junker L, Grabner A, Kohn B
Clinical evaluation of the CA530-VET hematology analyzer for use in veterinary practice
Veterinary Clinical Pathology 2007; 36 (2): 155-166

(51) Clark P, Mogg TD, Tvedten HW, Korcal D
Artifactual changes in equine blood following storage, detected using the ADVIA 120 hematology analyzer
Veterinary Clinical Pathology 2002; 31 (2): 90-94

(52) Stockham S, Scott M
Fundamentals of Veterinary Clinical Pathology 1st edition, p 14
Blackwell Publishing 2002

8. Abbreviations

CHCM	corpuscular haemoglobin concentration mean
CLIA	Clinical Laboratory Improvement Amendments
CV	coefficient of variation
Hb	haemoglobin
HCT	hematocrit
ICSH	International Council for Standardization in Hematology
MCHC	mean corpuscular haemoglobin concentration
MCV	mean corpuscular volume
NCCLS	National Committee for Clinical Laboratory Standards
PBS	phosphate buffered saline
PCV	packed cell volume
PLT	platelet
r	correlation coefficient
RBC	red blood cell
$S_{a,tot}$	total analytical imprecision
$S_{y\|x}$	standard deviation of residuals

SD	standard deviation
S_d	standard deviation of the differences
TE_a	total allowable error
WBC	white blood cell
WIC	WBC count of instrument A, derived from impedance-technology
WOC	WBC count of instrument A, determined by flow cytometry

9. Figures

The attached CD contains all graphs of the Regression and Bland-Altman analyses.

10. Tables

11. Acknowledgments

This study was financially supported by IDEXX Laboratories, Inc., Westbrook, Maine.

There are several people who I would like to thank for various supports:

- Prof. Dr. Andreas Moritz

- Prof. Dr. Urs Giger

- Dr. Kathleen Freeman

- John Roche

- Dr. Dennis DeNicola

- Pamela Vachon

- Michelle Haynes

- James Russell

- All my colleagues at the Small Animal Clinic, Internal Medicine in Giessen

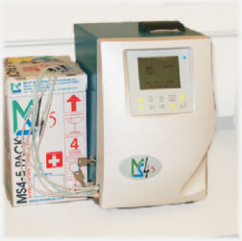
- All employees at Clinical Diagnostic Solution in Florida

- Alexander Stirn and my parents

Erklärung

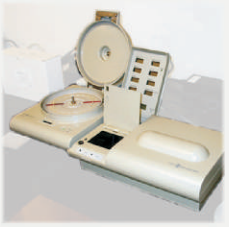
Ich erkläre: »Ich habe die vorliegende Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.«

Martina Becker, 2007









édition scientifique
VVB LAUFERSWEILER VERLAG

VVB LAUFERSWEILER VERLAG
STAUFENBERGRING 15
D-35396 GIESSEN

Tel: 0641-5599888 Fax: -5599890
redaktion@doktorverlag.de
www.doktorverlag.de

ISBN 3-8359-5243-9