

University of Nottingham
UK | CHINA | MALAYSIA

NOVEL APPROACHES TO QUALITY CONTROL AND HARMONISATION OF ANALYSERS WITHIN A VETERINARY LABORATORY SYSTEM

Susan Daly

BSc(Hons) MSc Biomedical Science

ID 20219842

Faculty of Veterinary Medicine and Science

This thesis is submitted in full satisfaction of the requirements for the degree of Doctor of Philosophy at the University of Nottingham.

March 2024

ACKNOWLEDGEMENTS

This research has been a journey but one I have not walked alone thanks to the support and guidance of many people who I want to thank.

Firstly, this research would not have been possible without my inspirational supervisors, Dr Kathy Freeman and Dr Peter Graham who championed this research and offered their time, expert advice, knowledge, encouragement, guidance and support. I feel I most definitely struck gold with you both, and I am truly grateful. With that in mind, thank you to Dr Nuthar Jassam for the opportunity to discuss laboratory harmonisation and accommodating a visit to her laboratory.

A great thanks to my fellow colleagues at VPG and especially those on my team who joined the research at various stages of the journey. A special thanks in particular to Amy Browne, who implemented this research with ease but who also supported my PhD journey, it would have been a lonely road without you.

Finally to my family, this was a sacrifice for us all and you came on the journey and never complained. To my parents for their encouragement and support, you ensured I was never afraid to ask questions and to always reach for the stars. To my (Fairy) Godmother who has been my cheerleader throughout and who has cheered all the way through my academic career, especially now. To my Mother and Mother-in-law for stepping in for my children when I needed them. To my Husband for your devoted support and encouragement for my quest for knowledge that unbalanced our family life

but you kept everything going. To my wonderful daughters, you have grown up alongside this PhD and I am truly blessed to have you by my side, you've kept my motivation up and my spirits high.

It has been a long and at times challenging journey but it has been more enjoyable due to the support from you all in different ways. Thank you.

This thesis is dedicated to my daughters, Ashleigh and Robyn.

We are stronger together.

ABSTRACT

The purpose of the research project was to develop a novel laboratory quality system to harmonise the five separate laboratories operating a commercial diagnostic service. Methods for initial and ongoing maintenance of harmonisation of haematology analysers within the laboratories were developed and implemented.

This project involves novel studies, as there are no relevant publications in the veterinary literature on these topics. The literature review presented in chapter 1 provides an overview of harmonisation followed by sections that cover the methods, tools and approaches used to assess the performance of the analytical network and to implement innovative quality control (QC) methods for veterinary haematology analysers.

Chapter 2 provides the first report of a comparison of performance evaluations for a network of six Sysmex analysers. A three-step process was used to evaluate and optimise performance to achieve harmonisation using quality parameters such as biologic variation for bias and CV, sigma metrics and total error as recommended by American Society of Veterinary Clinical Pathology (ASVCP).

Chapter 3 describes a quality control validation approach to maintain the harmonisation status and support the transition to harmonisation to ensure suitable clinical interpretation across the laboratory network. The results

demonstrated that using a QC-rule based on the performance of individual analysers rather than manufacturers' acceptable limits improved the probability of error detection (P_{ed}) and reduced the probability of false rejection (P_{fr}) for veterinary haematology analysers. In conjunction with a high sigma strategy, achievement of biologic variation-based goals and TE_a as recommended by ASVCP or expert opinion were used to establish acceptable performance for harmonisation across the analytical network.

Chapter 4 presents a first report of validating repeat patient testing-quality control (RPT-QC) using retained patient specimens for a network of haematology analysers as an alternative to commercial quality control material. Veterinary specific quality control material is not readily available which presents a difficult challenge for veterinary laboratories as a commutable matrix would be from animal origin. The repeat patient testing quality control (RPT-QC) could detect unstable analytical performance and RPT-QC limits were successfully generated using 40 data points rather than using 20 data points that was previously recommended. RPT-QC limits were required for each individual analyser and sigma metrics provided a useful performance tool to indicate suboptimal performance.

Chapter 5 outlines the process of implementation of RPT-QC, including an evaluation in parallel with the commercial quality control material to give confidence that RPT-QC can be used as practical alternative to commercial quality control material and that each individual analyser requires validation.

Chapter 6 outlines the design, evaluation and implementation of an internal laboratory comparative (ILC) testing program to maintain harmonisation by comparing the network analysers to the designated reference analyser. The ILC program used surplus canine samples with results across a range of clinical importance. Evaluation of the ILC results indicated that the ILC testing program could be used as an alternative to multiple subscriptions to an external quality assurance (EQA) program. This successful implementation offered a species-specific advantage and could determine whether performance issues were attributed to bias and/or imprecision and whether the interpretation of results was interchangeable between the network analyser and the reference analyser.

Chapter 7 outlines a pilot program for comparative testing offered to veterinary practices with in-clinic point of care testing (POCT) haematology analysers in comparison to the network reference haematology analyser. This study aimed to determine if the results of the POCT analysers could be harmonised with the reference analyser to ensure that results were interchangeable and clinically useful for diagnosis and monitoring, whether the results were generated by the POCT analyser or the reference analyser.

Chapter 8 is a discussion of the innovative approaches used in this study and reflections that may be useful for other veterinary laboratories desiring to achieve harmonisation for a network of analysers, wish to design their own comparative testing programs and/or provide comparative testing for clients with POCT analysers.

This chapter also describes the limitations and challenges associated with this research and possible future directions.

ABBREVIATIONS

AACC – American Association for Clinical Chemistry

ASVCP – American Society of Veterinary Clinical Pathology

BV – biologic variation, random, physiological fluctuations for an analyte around a homeostatic set point

CLSI – Clinical and Laboratory Standards Institute

CV – coefficient of variation, the standard deviation for a number of results divided by their mean, usually expressed as a percentage

CV_A – analytical coefficient of variation, variation attributed to the analyser

CV_G – group coefficient of variation, variation in homeostatic setting points between individual

CV_I – Individual Coefficient of Variation, random, physiological fluctuations of an analyte within an individual around a homeostatic setting point

DACC – Division of Animal Clinical Chemistry

DPM – defects per million

EDTA –ethylenediaminetetraacetic acid

EFLA – European Federation of Clinical Chemistry and Laboratory Medicine

EQA – external quality assurance

EU – European Union

ICHCLR – International Consortium for Harmonisation of Clinical Laboratory Results

IFCC – International Federation of Clinical Chemistry and Laboratory Medicine

ILC – interlaboratory comparison

INR – International Normalised Ratio

IQC – internal quality control

ISO – International Organisation of Standardisation

OOC – out of control

P_{ed} – sensitivity of a control rule in error detection or probability of detecting abnormal results

P_{fr} – specificity of a control rule in detecting error

PT – proficiency test

POC – point of care

POCT – point of care testing

QALS – Quality Assurance Laboratory Standards

QC – quality control

QCM – quality control material

QGI – quality goal index

QMS – quality manual system

RI – reference interval

RMP – reference measurement procedure

RPT – repeat patient testing

RPT-QC – repeat patient testing-quality control

SD – standard deviation, statistical value for the dispersion or spread of a set of measurements from the mean value (average)

SI – International System of Units

SQC – statistical quality control

TE – total Error, the sum or combined effect of random and systematic errors

TE_a – total allowable error, an quality assessment of imprecision and bias of one method compared to an established method

TE_{obs} – total observed error, Sum of measured random error (bias & imprecision) calculated from analytical performance data

GLOSSARY OF TERMS

Accuracy: The closeness of agreement between results of a measurement and the true concentration of the measurand.

Analytical quality requirements: Refers to a guidelines for the amount of variation acceptable in results for a given test which must not compromise interpretation.

Analytical error/variation: Variation in measurements/results due to analyser variation, comprises bias (accuracy) and imprecision (repeatability).

Bias: or 'systematic error' or 'inaccuracy' is the difference between the true value and the estimated value, this may be 'constant bias' (across all concentrations) or 'proportional bias' (error changes as when concentration changes).

Biologic variation: Expected, inherent, physiological fluctuations in measurands, can be 'between-subject', 'inter-individual' or 'group' biologic variation (CV_G) and 'within-subject' or 'intra-individual' biologic variation (CV_I).

Calibration traceability: ability to trace the readings of a measuring instrument to a known reference standard.

Coefficient of variation: the relative standard deviation expressed as a percentage, a statistical measure of imprecision including biologic variation.

Comparison testing: Where results from 2 or more analysers, instruments, methods are compared to each other to assess any performance differences. This can be in the form of EQA/PT programs.

Commutable/commutability: to yield equivalent results mathematically between different measurement procedures for a material such as reference, quality material and representative samples from individuals.

Inaccuracy: Refers to the 'trueness' of a result or the closeness the result is to the 'true' results (a reference method) for the measurand.

Levey-Jennings chart: control chart commonly utilised in healthcare laboratories to plot the result observed for a stable control material known as 'true' value versus time.

Measurand: quantity, object, or property intended to be measured.

Precision: Opposite of imprecision, Closeness of agreement between repeated results using the same specimen and conditions and expressed as CV of the test. May be repeated the same day (repeatability study) or over a number of days (within laboratory precision).

Quality control material: a variation of test material used to monitor the entire quality control process.

Reference intervals: a statistical interval that includes an upper and lower limit which is derived from clinically healthy reference population.

Repeat patient testing: Repeated testing for an individual patient under specified conditions that are used for statistical QC using the principle the samples deteriorate that under defined conditions (such as storage, fixed time point) that cause an expected degree of variation in the results whereby results exceeding this expected variation indicates unstable performance for the testing system.

Six sigma: A goal for performance that requires 6 SDs of variation to fit within tolerance limits and used as a quality marker for world class performance.

Standard deviation: a statistic describing the spread of a measurement set from the mean value of normal or Gaussian distribution.

Thromboplastin: enzyme released from damaged cells, especially platelets, which converts prothrombin to thrombin during the early stages of blood coagulation.

Transference: term used to describe a method for establishing whether reference intervals are appropriate by comparing results from 20 healthy patients from the population. The reference intervals are valid if no more than 2 results are outside the interval, and the transference is rejection is more than 2 results fall outside the interval.

Westgard rules: ‘ A control procedure that uses a series of control rules to test the control measurements. A 12s rule is used as a warning, followed by use of 13s, 22s, R4s, 41s, and 10x as rejection rules’

A note on style and terminology

There may be some repetition across the chapters as each chapter is a standalone representation of work that was published from this research. Throughout this thesis the term ‘in-clinic’ is used to refer to veterinary practices or hospitals that have point of care haematology analysers.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **II**

ABSTRACT ..... **V**

ABBREVIATIONS ..... **IX**

GLOSSARY OF TERMS ..... **XI**

A NOTE ON STYLE AND TERMINOLOGY XIII

LIST OF FIGURES ..... **XX**

LIST OF TABLES ..... **XXII**

CHAPTER 1: INTRODUCTION **27**

1.1 BACKGROUND AND LABORATORY QUALITY METHODOLOGY 27

1.1.1 Analytical performance evaluation *32*

1.1.2 Quality goals for veterinary haematology testing *38*

1.1.3 Quality control validation *42*

1.1.4 Sigma metrics *45*

1.1.5 Repeat patient testing vs commercial QCM *49*

1.1.6 External quality assurance *52*

1.1.7 Reference intervals *54*

1.1.8 Aims of this research project *57*

CHAPTER 2: EVALUATION OF SYSMEX XT-2000*i*V ANALYSER PERFORMANCE ACROSS A NETWORK OF FIVE VETERINARY LABORATORIES USING A COMMERCIALLY AVAILABLE QUALITY CONTROL MATERIAL **60**

2.1 INTRODUCTION 62

2.2 MATERIALS AND METHODS 64

2.2.1 Analysers *64*

2.2.2 Quality control material *65*

2.2.3 Data collection *65*

2.2.4 Quality evaluations *66*

2.3 RESULTS 69

2.3.1 Sigma metrics and QGI *69*

2.3.2 Biologic variation goals *77*

2.4 DISCUSSION 81

2.5 CONCLUSION 88

CHAPTER 3: QUALITY CONTROL VALIDATION FOR A VETERINARY LABORATORY NETWORK OF SIX SYSMEX XT-2000/V HAEMATOLOGY ANALYSERS 90

3.1 INTRODUCTION 92

3.2 MATERIAL AND METHODS 96

3.2.1 Haematology analysers *96*

3.2.2 Quality control material *97*

3.2.3 Evaluation of manufacturer's acceptable limits as control rules using EZRules3© *97*

3.2.4 QC validation – evaluation of control rules based on observed individual analyser performance *98*

3.2.5 Review of criteria for measurands with poor performance *99*

3.3 RESULTS 99

3.3.1 QC validation for manufacturer's acceptable limits *99*

3.3.2 QC validation based on observed analyser performance for 6 Sysmex analysers *103*

3.3.3 Evaluation of QC rules *103*

3.3.4 Sigma metric monitoring *104*

3.3.5 Review and optimisation of criteria for measurands with poor performance *104*

3.4 DISCUSSION 105

3.5 CONCLUSION 111

CHAPTER 4: REPEAT PATIENT TESTING-QUALITY CONTROL SHOWS PROMISE AS AN ALTERNATIVE TO COMMERCIAL QUALITY CONTROL MATERIAL FOR A NETWORK OF 4 SYSMEX XT2000/V HAEMATOLOGY ANALYSERS 112

4.1 INTRODUCTION 114

4.2 MATERIALS AND METHODS 116

4.2.1 Analysers and measurands 116

4.2.2 RPT-QC limits generation 117

4.2.3 QC validation for calculated RPT-QC control limit interval 121

4.2.4 RPT-QC and sigma metrics 122

4.2.5 RPT challenges 123

4.3 RESULTS 123

4.3.1 RPT-QC limit generation 123

4.3.2 QC Validation for RPT-QC intervals 124

4.3.3 RPT-QC and sigma metrics 125

4.3.4 RPT-QC challenges 130

4.4 DISCUSSION 130

4.4.1 Implications for our network of analysers using RPT-QC 134

4.5 CONCLUSION 136

CHAPTER 5: REPEAT PATIENT TESTING-QUALITY CONTROL COMPARED TO COMMERCIAL QUALITY CONTROL MATERIAL FOR THE SYSMEX XT-2000/V HAEMATOLOGY ANALYSER IN A MULTI-SITE VETERINARY LABORATORY 138

5.1 INTRODUCTION 140

5.2 MATERIALS AND METHODS 142

5.2.1 Determine if differences between RPT-QC repeat analyses for each of the time intervals differed across the network. 143

5.2.2 Determine if unified RPT-QC limits can be applied across the network of analysers. *144*

5.2.3 Comparison of RPT-QC with a commercial quality control material *144*

5.2.4 Automated spreadsheet development *146*

5.2.5 Technician experience and feedback *146*

5.2.6 Statistical analysis *147*

5.3 RESULTS 147

5.3.1 Evaluation to determine if unified RPT-QC limits could be used across the network of analysers. *150*

5.3.2 Automated spreadsheet development. *155*

5.3.3 Technician feedback. *157*

5.4 DISCUSSION 158

5.5 CONCLUSIONS 163

CHAPTER 6: MAINTAINING HARMONISATION OF SYSMEX XT-2000/V ANALYSERS IN A NETWORK OF VETERINARY LABORATORIES USING AN INTERLABORATORY COMPARATIVE TESTING PROGRAM 165

6.1 INTRODUCTION 166

6.2 MATERIALS AND METHODS 168

6.2.1 Multi-level performance evaluation *170*

6.2.2 Single analysis events *172*

6.2.3 Performance goals and ranking *173*

6.2.4 Reference analyser EQA performance *178*

6.3 RESULTS 178

6.3.1 Multi-level performance evaluation *178*

6.3.2 Single level evaluations *179*

6.3.3 Reference analyser EQA performance *186*

6.3.4 Performance goals and sigma metrics *187*

6.4 DISCUSSION 190

6.5 CONCLUSION 195

CHAPTER 7: COMPARATIVE TESTING FOR IN-CLINIC POINT OF CARE HAEMATOLOGY ANALYSERS WITH A COMMERCIAL LABORATORY REFERENCE ANALYSER 197

7.1 INTRODUCTION 198

7.2 MATERIALS AND METHODS 201

7.2.1 Study design for in-clinic and reference analyser comparative testing 201

7.2.2 Performance evaluation 203

7.2.3 Performance grading key 205

7.2.4 In-clinic POC analyser result harmonisation 206

7.2.5 Technician/Clinician feedback 207

7.3 RESULTS 208

7.3.1 Study design 208

7.3.2 Performance evaluation 208

7.3.3 Quality goal performance 212

7.3.4 Harmonisation 217

7.3.5 Participant feedback 220

7.4 DISCUSSION 221

7.5 CONCLUSION 227

CHAPTER 8: DISCUSSION AND LIMITATIONS 229

8.1 REVIEWING THE SCOPE OF THE THESIS 229

8.2 REVISITING RESEARCH FINDINGS 231

8.3 IMPORTANCE AND RELEVANCE TO THE WIDER VETERINARY LABORATORY SECTOR 239

8.4 CHALLENGES AND LIMITATIONS 244

8.5 OVERALL LEARNING OPPORTUNITIES 246

8.6 CONCLUSION 249

8.7 FUTURE DIRECTIONS 251

8.7.1 Recommendations for future research 251

CHAPTER 9: REFERENCES 253

APPENDIX A: PUBLICATION OF CHAPTER 2.....267

APPENDIX B: PUBLICATION OF CHAPTER 3.....279

APPENDIX C: PUBLICATION OF CHAPTER 4.....292

APPENDIX D: PUBLICATION OF CHAPTER 5.....302

APPENDIX E – SUPPLEMENTARY TABLES 315

APPENDIX F – SUPPLEMENTARY FIGURES.....338

LIST OF FIGURES

Figure 4.1: Initial scatterplot for RBC using 19 data points (one outlier removed) from the difference of Day1-Day2 samples used to calculate control limits for 1-3s rule.....120
Figure 4.2: Subsequent scatterplot for RBC using 39 data points (one outlier removed) from the difference of day1-day2 samples used to calculate control limits for 1-3s rule.....121
Figure 5.1: Levey Jenning (L-J) chart for commercial quality control material for haematocrit using a 1-2.5s QC rule. 153
Figure: 5.2 L-J chart for RPT-QC Day1-Day2 for HCT using a 1-2.5s QC rule. 154
Figure 5.3 (A-D): Images of the automated data management spreadsheet. 155

Figure 6. 1: Schematic for a harmonised network of analysers showing the network analysers linked to the reference analyser which remains subscribed to a EQA program. .170
Figure 6. 2: Automated spreadsheet showing result entry, quality goals and performance grading criteria.....176
Figure 6. 3: Automated spreadsheet showing area on spreadsheet for technical comments. 177
Figure 6. 4: Automated spreadsheet showing area for Pathologist comment. 177

Figure 7. 1: The performance grading key and corresponding criteria to achieve performance grades; “excellent”, “good”, “fair”, “poor”. 205

Figure F4. 1: An example of data calculations using 20 data points for HGB Day1-Day2 showing D1 (day1) measurements, D2 (day2) measurements, Diff (difference from Day1 minus Day2), Aver D1,D2 (average taken of Day1 and Day2 measurement) and Diff sq (Difference squared). 339

Figure F6. 1 (A- E): Bar charts of single ILC results for the erythroid measurands for the analyser network for each ‘bin’.....339
Figure F6. 2 (A-E): Bar charts of single ILC results for the leukocyte measurands for the analyser network.....345

Figure F6. 3 (A-E): Bar charts of single ILC results for the platelet measurands for the analyser network.....350
Figure F6. 4: An image of the reference analyser EQA spreadsheet results prior to routine service. 355
Figure F6. 5: An image of the reference analyser EQA (RIQAS) spreadsheet results following the service.....356

LIST OF TABLES

Table 2.1: Initial results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 1 (November) evaluation. 70

Table 2.2: Follow up results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 2 (December) evaluation. 72

Table 2.3: Follow up results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 5 (March) evaluation. 73

Table 2.4: Summary of investigation of measurands with sigma metrics < 6 for 6 Sysmex XT-2000iV Hematology Analysers in a system of veterinary laboratories using a commercially available quality control material. 74

Table 2.5: Summary of CV biologic variation goals that reached optimally or desirably acceptable levels for 8 measurands where biologic variation data is available, assessed during three evaluations. 77

Table 2.6: Summary of bias biologic variation goals that reached optimally or desirably acceptable for 8 measurands where biologic variation data is available assessed during three evaluations. 78

Table 2.7: Summary of 0.33 x CV_I biologic variation goal for multiple instruments within a laboratory system that reached optimal or desirable acceptable goals for 8 measurands where biologic variation data is available, assessed during three evaluations. 79

Table 3.1: Summary of QC validation for Sysmex quality control Level 2, Manufacturer's acceptable quality control material (QCM) ranges using one month of quality control data from the system reference laboratory Sysmex XT 2000iV haematology analyser (analyser 1). 101

Table 4.1: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for White Blood Cell Count (WBC) 126

Table 4.2: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Red blood cell count (RBC). 127
Table 4.3: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Haemoglobin concentration (HGB). 128
Table 4.4: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Haematocrit (HCT). 129
Table 5.1: Kruskal-Wallis Test comparing median differences across the analyser network, significance set to $P < 0.05$. 149
Table 5.2: Difference of means test showing standard error for the network analysers (3, 4a, 4b) when compared to the reference analyser (1) for each time interval and measurand (RBC, HGB, HCT, WBC) 151

Table 6. 1: Five levels of haematology results of clinical interest generated based on haematocrit (HCT)(%) created by addition or removal of canine plasma. 172
Table 6. 2: Performance criteria key for Internal Laboratory Comparative Testing assessment and RIQAS EQA assessment for the reference analyser. 175
Table 6. 3: Measurands with poor performance for the erythroid measurands for the network analysers using quality goals of TE_a (ASVCP/expert opinion), sigma metric, Desirable CV (biologic variation), Desirable Bias (biologic variation), clinical interpretation and quality goal index (QGI). 181
Table 6. 4: Fair or poor performance for the leukocyte measurands for the network analysers, using quality goals of TE_a (ASVCP/expert opinion), sigma metric, Desirable CV (biologic variation), Desirable Bias (biologic variation), clinical interpretation and quality goal index (QGI). 183

Table 6. 5: Fair and poor performance summary for platelets for the network analysers, using quality goals of TE_a (ASVCP/expert opinion), sigma metric, clinical interpretation and quality goal index (QGI) 185
Table 6. 6: TE quality goals for internal laboratory comparison testing for very low, low, normal, high and very high haematocrit bins (%). 188

Table 7. 1: Presents the observed mean, observed bias, and observed CV (%) generated by the individual in-clinic POC haematology analysers that participated in the comparative testing program using different canine samples that were chosen within the manufacturer's provided POC reference interval before sending to the reference laboratory for comparison to a Sysmex XT-2000iV 209
Table 7. 2: The results for 10 haematology measurands that could achieve/not achieve total allowable error as recommended by ASVCP or by expert opinion analysed on 7 in-clinic point of care haematology analysers using individual normal canine samples 213
Table 7. 3: Performance rank for each POCT haematology analyser using the performance grading system compared to the scoring points system to determine best performance for 7 in-clinic POC haematology analysers when compared to a reference haematology analyser. 216
Table 7. 4: Pre-and Post-harmonisation results evaluated with canine samples using point of care testing haematology analysers showing the measurands that were not within combined inherent imprecision when compared against a reference analyser and required harmonisation. 217

Table E3. 1: Summary of the haematological erythroid measurand QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All instruments used the same quality control material lot number. The number of quality control materials is one (n=1). The number of quality control materials is one (n=1). Ped < 0.85 are highlighted in bold type. 317
Table E3. 2: Summary of Platelet (PLT) and Plateletcrit (PCT) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All

instruments used the same quality control material lot number. The number of quality control materials is one (n=1). Ped < 0.85 is highlighted by bold type. 323
Table E3. 3: Summary of White Blood Cell Count (WBC) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All instruments used the same quality control material lot number. The number of quality control materials is one (n=1).....325

Table E5. 1: SD, CV, mean difference and RPT-QC limits for the reference (1) and network analysers (3, 4a, 4b) for each time interval for RBC, HGB, HCT, WBC.....326
Table E5. 2: Example calculations for RBC Day1-Day4, n=20 with outlier (Y) in row 14 in bold type. 328
Table E5. 3: RBC mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule 329
Table E5. 4: HGB mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule 330
Table E5. 5: HCT mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule. 331
Table E5. 6: WBC mean difference, upper and lower RPT-QC limits and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule 332
Table E5. 7: Comparison of RPT-QC and commercial QCM over a 4-month period for the reference analyzer (analyzer 1).....333

Table E6. 1: Mean and CV(%) range for the erythroid measurands generated from replicate analyses of 5 different ‘bins’ of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.....334

Table E6. 2: Mean and CV(%) range for the leukocyte measurands generated from replicate analyses of 5 different ‘bins’ of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers. 335
Table E6. 3: Mean and CV(%) range for the platelet measurands generated from replicate analyses of 5 different ‘bins’ of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers. 336

Table E7. 1: The number (n) of each measurand evaluated on 7 point of care haematology analysers using different canine EDTA samples that could achieve > 4 sigma metric and using the QGI whether those that didn't achieve > 4 sigma was due to bias, imprecision or both bias and imprecision. 337

CHAPTER 1: INTRODUCTION

1.1 Background and Laboratory Quality Methodology

Laboratory harmonisation in veterinary medicine have not been previously documented but the concepts of harmonisation and standardisation from human medicine are applicable (Miller *et al.*, 2011a; Berg and Lane, 2011; McLawhon, 2011; Plebani, 2013; Greenberg, 2014; Thienpont *et al.*, 2017) and should be utilised within veterinary laboratories and veterinary laboratory networks. The association for diagnostic and laboratory medicine (ADLM, formerly AACC) includes the division of animal clinical chemistry (DACC) and has given full support for harmonisation for all areas of laboratory testing based on a global goal to improve overall patient care (ADLM). This is not an easy task, but it is a worthwhile one since ensuring that a network of veterinary laboratories have taken the necessary steps to harmonise their measurement procedures gives a degree of confidence that a patient result is equivalent within that network. A lack of harmonisation means that different results from patient samples may be obtained from different laboratories (Miller *et al.*, 2011a), potentially increasing the risk of clinical misclassification of the patient, flawed accumulation of patient data used for patient monitoring as well as financial burdens and suffering associated with resolving discrepancies, possible delays in treatment or inappropriate or unnecessary treatment based on inaccurate or misinterpreted results.

Harmonisation and standardisation are terms that are often used interchangeably as both ultimately use measures to ensure that results from different laboratories are equivalent. However, they differ in that standardisation refers to measurands where reference material or reference measurement procedures (RMPs) exist and are traceable to the International System of units (SI), while harmonisation refers to those measurands that do not have RMPs but are traceable to reference material or by consensus to an all-methods mean or a designated standard (Miller *et al.*, 2011a). Standardisation has gone through numerous stages commencing in the 1950s when it was recognised from external quality assurance (EQA) that results were not equivalent between different laboratories. Efforts to address and challenge this were seen from 1960s (Radin, 1967), but were hampered until the 2000s when commutable human matrices for EQA and for reference materials was established (Miller *et al.*, 2006; Miller *et al.*, 2013). The 1998 European Union (EU) Directive 98/79/EC created a requirement for international standards of metrological traceability that led the International Organisation for Standardisation (ISO) Technical Committee to publish the Clinical Laboratory testing and in vitro diagnostic test system (ISO, 2003; ISO, 2009). Metrological traceability as defined by ISO is "*the property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties*" (ISO, 2003). The term *harmonisation* didn't appear in the laboratory literature until the 2010 AACC (Now ADLM) International congress whose theme was to

improve clinical laboratory testing with harmonisation. Following this conference, the recommendations were published in 2012 as a roadmap to harmonisation for clinic laboratories (Miller *et al.*, 2011a). These required an organisation to be established to determine the measurands requiring harmonisation, how this would be achieved at a global level and for procedural development for measurands that did not have reference material or reference measurement procedures (Miller *et al.*, 2011b). This led to the establishment of The International Consortium for Harmonisation of Clinical Laboratory Results (ICHCLR) to implement the recommendations from the AACC (AACC). It was recognised that an ISO standard would be required for a harmonisation protocol to achieve metrological traceability and therefore ISO/CD 21151 was established (ISO, 2020).

Today, harmonisation (and standardisation) in human medicine are ultimately based on ISO 17511 traceability principles (ISO, 2003) and are divided into five categories: the first three categories are for calibration traceability for which reference materials are well developed (Armbruster and Miller, 2007; Panteghini, 2009), category four refers to measurands for which limited reference material is available for calibration and category five for measurands which do not have calibration reference materials and are, therefore, more difficult to harmonise (Ekins, 1991).

A limitation to harmonisation throughout the decades has been material commutability and this is a term that is defined by the International Vocabulary of Metrology as *" a property of a reference material, demonstrated by the closeness of agreement between the relation among the measurement results for a stated quantity in this material, obtained*

according to 2 given measurement procedures, and the relation obtained among the measurement results for other specified materials" (ISO, 2007).

This was not recognised when calibration traceability was first presented by Radin in 1967 but has since been a major area of research and the impact of a commutable matrix was not well understood until the 2000s (Miller *et al.*, 2006; Miller *et al.*, 2013). Commutability of reference materials, whether used to validate accuracy or used to calibrate measurement, are essential for metrological traceability. It became evident when patient samples showed different results with different measurement procedures calibrated to the same reference material (Miller *et al.*, 2009; Rose, 1998; Stephan *et al.*, 2006; Sturgeon *et al.*, 2009), demonstrating that the reference material was, in fact, non-commutable. There is, to the author's knowledge, only a single paper in the veterinary literature has investigated the commutability of commercially available biochemistry quality control materials with feline plasma (Baral *et al.*, 2013). However, it has been recognized that different results for bias and imprecision are obtained with quality control material (QCM) and pooled patient samples for some measurands (Flatland *et al.*, 2014b; Tecles *et al.*, 2020).

Harmonisation is divided into three phases: pre-analytical (Lippi *et al.*, 2013; Lippi *et al.*, 2015), analytical (Miller *et al.*, 2014a) and post-analytical (Jones and Legg, 2019). The pre-analytical phase has not been as well established as the analytical phase and that may be due to the degree of control required of factors that are external to the clinical laboratory (Plebani *et al.*, 2014). This phase includes patient preparation, sample collection and sample transportation. Some efforts in human medicine have been made to

standardise sample collection procedures (Simundic *et al.*, 2014) and to optimise sample transport protocols (Zaninotto *et al.*, 2012a; Zaninotto *et al.*, 2012b). In veterinary medicine this is usually addressed in laboratory user manuals or information sheets provided by individual commercial laboratories to their veterinary clients.

The analytical phase is dedicated to the harmonisation of the laboratory results and much research has taken place in human laboratory medicine (Braga and Panteghini, 2013; Weykamp *et al.*, 2014). A successful and impressive harmonisation example has been the International Normalised Ratio (INR) which is a mathematical expression of prothrombin time into a ratio which allows comparison for all thromboplastin from patients receiving vitamin K antagonists (Tripodi *et al.*, 2016; Meijer *et al.*, 2018). Successful harmonisation for HbA_{1c} was also achieved and an international network was established for global HbA_{1c} standardisation (Weykamp, 2013). Two laboratories in the Netherlands are the recognised and certified reference methods for HbA_{1c} and the improvements included standardising the measurement units. This successful harmonisation reports impressive improvements to interlaboratory CVs from 30% to 3% (Jansen *et al.*, 2018). Currently, assays optimised, standardised, or harmonised for human medicine are often adopted without further verification by veterinary testing laboratories.

The post-analytical phase includes methods for interpreting the test results which include reporting units, reference intervals (RIs) or clinical decision limits. Test results are generally transferred electronically and a need for harmonisation of laboratory reporting has been recognised (Miller *et al.*,

2014b). In veterinary medicine, country or regional practices for reporting units and reference intervals may vary and there has not been success in adoption of SI units globally for veterinary testing, despite some suggestions to do so. Clinical decision limits or cut-off values in veterinary medical laboratory testing are often based on expert consensus documents, many of which do not include input from laboratorians or include very limited information about important laboratory and analytical performance considerations (Flatland *et al.*, 2020; Sparkes *et al.*, 2015; Behrend *et al.*, 2018).

The analytical phase and post analytical phases were the focal points for the research in this thesis as they include a method harmonisation process and harmonisation of the reference intervals (RIs) to allow common RIs to be used on an ongoing basis.

Therefore, this review is divided into sections relating to efforts to harmonise a network of veterinary haematology analysers:

1.1.1 Analytical performance evaluation

To achieve optimal patient care and comparability of results for laboratories for the analytical phase of harmonisation, its necessary to determine the total analytical error or total error (TE) and its components: bias and imprecision to characterise and assesses the instruments' capability to produce accurate results. First introduced in 1974 (Westgard *et al.*, 1974), total analytical error is a quantitative measurement and is comprised of (a) random error

(precision or imprecision) which determines the repeatability or closeness or agreement between independent results and (b) systematic error (trueness, bias, or inaccuracy) which determines how much variation is observed compared to a 'true value'. Therefore, TE is the combination of random and systematic error. Its associated quality requirement is called total allowable error and TE can be expressed as a percentage or in measurand units. It's quite useful when expressed as a percentage as it can be compared with other quality goals which are also often expressed as a percentage.

TE(%) is calculated as follows.

$$TE(\%) = \%Bias + 2CV\% \text{ (Flatland } et\ al.\text{, 2018a)}$$

Note: Multipliers other than 2 can be used.

Today, clinical laboratory equipment manufacturers will make general claims for bias and imprecision for the model of analyser but rarely for TE; clinical laboratories should verify these claims on their own specific analysers as TE varies amongst different analysers of the same make and model when assessed using the same specimen (Jensen and Kjelgaard-Hansen, 2006; Farr and Freeman, 2008).

a. Imprecision

Precision is a qualitative assessment that relies on the amount of random error rather than the trueness of the results. So, for quantitative assessment, the counterpart known as imprecision is of great interest. Imprecision is determined from repeatability of replicates for a measurement either during validation, external quality assurance (EQA), internal testing, or from internal

quality control (IQC) during routine operation. Although precision studies are carried out by the manufacturers, ongoing imprecision should be observed routinely to monitor the stability of the analytical system. Imprecision can be monitored from the statistical QC by calculating the standard deviation (SD). SD divided by the mean of results is used to determine the coefficient of variation (CV) (Jensen and Kjelgaard-Hansen, 2006).

$$CV = SD/mean \times 100$$

Imprecision goals must be defined according to specified conditions or a CV limit for the measurand.

This was initially proposed by Tonks in 1963 whereby limits for maximum allowable random error were set based on reference interval limits; this however was flawed because imprecision in the RIs was not accounted for (Tonks, 1963). The first medically significant CVs were based on consensus recommendations for 16 biochemistry analytes evaluated by 100 general practitioners and 125 internists (Barnett, 1968). These early attempts at establishing CVs from medically significant changes in test results and their conclusions should be viewed cautiously as they are opinions and based on results of unknown imprecision. Acceptable CVs needed to be developed for each analyte based on medical significance, and proposed in 1970 that biologic variation needed to be included in the assessment of imprecision so that the analytical CV (CV_A) is not greater than half the previously determined intra-individual variation (CV_I) for each analyte (Cotlove *et al.*, 1970). This was further developed by Fraser (1992) into optimal, desirable, and minimal imprecision performance goals. This system ensures that analytes/measurands with can be required to reach a defined quality

standard regardless of whether they have large or small BV (Fraser *et al.*, 1997). In the veterinary literature, most precision studies (Rishniw *et al.*, 2012; Rishniw and Pion, 2016) for haematology have not sufficiently taken BV into account.

b. Bias

Bias is also referred to as “inaccuracy” and expresses systematic error. It may be due to ‘constant’ bias where the error is consistent for all concentrations or ‘proportional’ bias when the error varies depending on the concentration (Jensen and Kjelgaard-Hansen, 2006; Flatland *et al.*, 2014b). Overall, bias refers to the difference in values measured by the test or instruments compared to the ‘true’ value (Stockham and Scott, 2008). The true value, in theory, should be from a definitive method which has been well researched, well standardised, and stable (Jensen and Kjelgaard-Hansen, 2006) and values from the definitive method can then be assigned to reference materials for use when comparing results from a new method. The ‘true’ value is a qualitative concept, and its quantitative counterpart is the lack of trueness which is a measurement of systematic error. Bias is one of the biggest challenges in the clinical laboratory and its elimination or minimisation is essential for the laboratory service and patient care (Kallner *et al.*, 2005; CLSI, 2013). Manufacturers perform studies to provide claims for their instrument’s performance, but this should be validated or verified by the laboratory and routinely monitored thereafter by one or more of the following methods; a definitive or reference method, a certified reference material, assayed quality control material(s) and/or external quality assessment

materials or comparison with a well characterised trusted field method (Flatland *et al.*, 2014b).

Ideally a minimum of 40 patient samples should be tested when introducing a new method or comparing a recognised one (Wayne, 2002) and the average concentration from the patient samples from the method to be tested should be compared to the average concentrations from the reference method using the same patient samples (Theodorsson and Magnusson, 2017). The concentrations from the patient samples should represent the clinically relevant range of results. Bias may also be estimated from certified reference materials (if available). The statistical analysis of bias is not predetermined; where a new method is being adopted, regression analysis (Stöckl *et al.*, 1998) and simply plotting data graphically in comparison (Bland and Altman, 1986) are recommended practical initial steps (Jensen and Kjelgaard-Hansen, 2006).

Bias in the units of the assay can be calculated by.

$$\text{Bias} = \text{Value}_{\text{measured}} - \text{Value}_{\text{target}}$$

or % Bias.

$$\%\ \text{Bias} = [(\text{Value}_{\text{measured}} - \text{Value}_{\text{target}})/\text{Value}_{\text{target}}] \times 100.$$

The target value may be the manufacturers target mean for assayed QCM, mean of a peer group from EQA data or the average of paired measurements carried out in a comparison study (Jensen and Kjelgaard-Hansen, 2006; Flatland *et al.*, 2018a). If multiple methods for assessment of bias provide estimates that are similar or complementary, this provides confidence that the estimate of bias is accurate. If various methods for assessment of bias give

discordant results, then further investigation to determine a reason for the discrepancies and the appropriate figures to use for various applications is needed. For QC validation purposes, estimates of imprecision and bias determined from the quality control material should be used (Tecles *et al.*, 2020). For interpretation of patient results, patient-derived information is needed (Tecles *et al.*, 2020).

In clinical chemistry, bias in a conglomerate of laboratories has been recognised as the most important bias to resolve (Theodorsson and Magnusson, 2017) and causes are applicable also to haematology. Important considerations include pre-analytical factors, sample matrix, and commutability. Pre-analytical factors include sample collection, sample transportation and sample interferences. Sample matrix is another important consideration when assessing bias (i.e., the components present in the sample other than the measurand have been known to contribute to bias (Sobas *et al.*, 2010; Unsal *et al.*, 2010; Miller *et al.*, 2006). As noted above, commutability is a qualitative concept that describes the extent to which control material, calibrators, or reference materials have matrix properties like patient samples. Patient samples are the most commutable material for comparability of methods (Miller and Myers, 2013).

Bias may be short term (within day) or long term (within months or across years) and it has been proposed that the short-term bias eventually reflects random effects in the long term (Theodorsson and Magnusson, 2017). Long term observations of bias may vary and contribute to part of the random error of measurement uncertainty. These bias components in well-controlled

systems are usually small and may be due to reagent lots changes, recalibrations and/or operator differences.

1.1.2 Quality goals for veterinary haematology testing

Analytical quality requirements are benchmarks that analytical performance should achieve and are also known as quality specifications or quality goals. Following a consensus in 1999, the 'Milan hierarchy' of 2015 describes three levels of quality approaches where level one is the most stringent and level three is the least (Sandberg *et al.*, 2015). The highest level of the hierarchy is not easily obtained since it is based on 'evaluation of the effect of clinical performance on clinical outcomes in specific clinical situations'. Such studies are rare because of the complexity of this model including the need for volumes of data relating to the effect of analytical performance on clinical outcomes (Petersen and Fraser, 2010). Additionally, this data is rarely available for veterinary medicine. The next level in the hierarchy is based on evaluation of analytical performance and on clinical decisions based on biologic variation (BV) data. The third level is for state of the art models that are employed by PT/EQA programs and use TE_a but the first two levels are prioritised over this (Sandberg *et al.*, 2015). Determining quality specifications based on biologic variation is becoming more widely used in recent years, including by the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and Clinical and Laboratory Standards Institute (CLSI) (Kjelgaard-Hansen and Jensen, 2010) and is receiving more

attention in veterinary medicine following the establishment of a veterinary biologic variation website and database (VetBiologicalVariation.org).

Biologic variation-based imprecision performance goals are a way of ensuring that the result generated is a reflection of within/intra-individual physiological fluctuations that occurs around a homeostatic/physiologic set point for measurands rather than analytical noise generated by analytical variation (Fraser *et al.*, 1992). Assessing BV requires the determination of individual variation CV_I (%), group variation CV_G (%) and analytical variation CV_A (%). Initially proposed by Cotlove and Harris (1970) and applied to human medicine, it was further refined by Fraser to clarify that healthy and sick patient groups can be distinguished (Fraser *et al.*, 1992; Fraser, 1992b). Applications of biologic variation-based performance goals in veterinary medicine are relatively recent but increasing in number (based on publications available) (Blois *et al.*, 2015; Wright *et al.*, 2019; Baral *et al.*, 2013; Rishniw and Pion, 2016). Much of this may be attributed to recommendations on how to perform BV studies becoming easier to access (Freeman *et al.*, 2017).

The use of biologic variation (BV) in quality goal setting provides acceptance limits determined from the physiological variation of the measurands (Baral *et al.*, 2013). BV quality goals are applied to maintain laboratory analytical performance and set limits to ensure that analytical error has minimal impact on clinical decisions (Petersen and Fraser, 2010). The BV goals quality standards are set to optimal, desirable, and minimally acceptable quality goals for imprecision (CV), Bias and total allowable error (TE_a).

As described above, total allowable error is a quality goal that combines both imprecision and bias to determine if analytical quality is acceptable, that it is fit for purpose and unlikely to cause erroneous clinical decisions. In veterinary publications, TE_a recommendations are usually consensus-based quality goals using expert opinion (what clinicians "think" the goals should be) and state of the art performance (what laboratories seem to be capable of), which differ from human medicine where goals may include those based in total or in part on analytical error derived from BV data (Ricós *et al.*, 2015; Westgard

QC, 2015; Aarsand AK).

The acceptable limits for imprecision and bias are calculated from within and between individual BV values which, depending on the quality standard level, have different multipliers summarised as follows:

$$\text{Allowable Imprecision} = X * CV_I$$

Where CV_I is the intra-individual BV and X ranges from 0.25 to 0.5 to 0.75 for optimal, desirable, and minimal quality standards, respectively.

$$\text{Allowable bias} = Y * \sqrt{CV_I^2 + CV_G^2}$$

Where CV_G is between-individual biologic variation and Y varies from 0.125 to 0.25 to 0.375 for optimal, desirable, and minimal quality standards, respectively (Fraser, 2001).

The "optimal" biologic variation quality goals are more stringent and therefore the "desirable" goals are usually considered to be sufficient, depending on how stringent or permissive the goals are for a particular measurand, based

on the BV inherent to the measurand. More recently the EFLM have suggested the use of minimally acceptable BV goals for laboratories unable to achieve the recommended quality levels (Aarsand AK). For measurands under tight physiological control and therefore with lower BV, the noise due to analytical imprecision may be a more likely contributor than BV to total error and not meeting BV-based goals (Petersen *et al.*, 2002).

BV based total analytical error can be calculated using the following formula.

Total allowable error = 1.65*(allowable imprecision) + allowable bias

Or

$$TE_a = 1.65*(X*CVi) + Y*\sqrt{(Cvi^2 = CVg^2)}$$

Allowable imprecision are X = 0.25, 0.5, 0.75 and allowable bias is Y = 0.125, 0.25, 0.375 for optimal, desirable and minimal acceptable levels respectively.

The observed total error (TE_{obs}) can be directly compared to total allowable error limits/guidelines published by ASVCP (Harr *et al.*, 2013; Westgard, 2013a).

Failure to include relevant quality goals, including biologic variation-based goals, failing to assess bias and/or imprecision to determine observed total error and use of inappropriate QC have been identified as problems in some method validation studies in the veterinary literature (Freeman and Flatland, 2017).

The literature demonstrates the importance of including quality goals such as TE_a, BV goals for imprecision (CV) and bias for focus (Flatland *et al.*, 2018a). And if biologic-variation based total error goals were considered, evaluating the contributions of bias and imprecision separately is still required. Published veterinary BV studies have been collated and curated and can be found at a free online database (VetBiologicalVariation.org).

1.1.3 Quality control validation

Statistical QC is used to provide peace of mind regarding the stable performance of a laboratory testing system, including the instruments, reagents, and operator. To detect significant disruptions in stable performance efficiently and effectively, recommendations are to employ statistical analysis capable of detecting errors with ≥ 0.90 probability and to limit false rejections/false error alerts to a probability ≤ 0.05 (Westgard, 2016a).

QC validation is a process where a statistical control rule or rules are selected for a QC results specific for each measurand. This requires a quality goal that enables the desired clinical interpretation of results. By validating the QC approach, it ensures that an appropriate procedure for high probability of error detection (P_{ed}) and low probability of false rejection (P_{fr}) is applied (Westgard, 2016a).

Laboratories using statistical QC often employ the Westgard rules which are denoted by a shorthand notation. For example, a 1-2s rule means that one

control measurement exceeds two control limits/standard deviations, the control limits for this rule are set as the mean plus or minus 2 SDs. A 1-3s refers to a rule that is set at plus or minus 3 SDs and the rule is violated when one measurement exceeds 3 SDs from the QCM mean. The numerical 2 placed in front of a rule such as 2.2s (or 2-2s) indicates the rule is violated when two consecutive control measurements or two single measurements across two QCMs are outside 2 SD control limits.

A multirule is a combination of the individual Westgard rules that are applied sequentially to accept or reject the QC run if one out of the set is violated. These rules are commonly used with a Levey-Jennings chart and when any rule is violated a trouble shooting process should take place. The multirule QC process is more complicated but may provide better P_{ed} for procedures that have performance that does not allow a high standard of QC with simple control rule. Multirules notation shows the individual rules employed with a slash in between, for example 1-3s/2-2s/R-4s/4-1s (Westgard *et al.*, 1981; Westgard and Barry, 1986; Westgard JO, 1996; Cembrowski GS, 1996). What rule to adopt depends on both the level of quality requirement for each measurand and its observed performance. Therefore, the rule selection process should consider:

1) Analytical performance of the measurand

2) The number of controls used

3) The quality goal used (Westgard, 2016a)

The chosen rule aims to have a practical goal for sensitivity ($P_{ed} \geq 0.90$) and specificity ($P_{fr} \leq 0.05$) and this may be predicted using power function curves.

Operating specification charts can then be used to review the candidate control rules applicable for the observed bias and CV for each measurand. This process can be enhanced using sigma metrics (see below), known as Westgard sigma rules. Measurands achieving sigma metrics of 6 or greater can be controlled with a single rule that can achieve acceptable P_{ed} and P_{fr} using 1-2 control measurements. Measurands with lower sigma metrics will require a multirule and/or larger numbers of control measurements to achieve acceptable P_{ed} and P_{fr} (Westgard and Barry, 2008). Commercially available software called Westgards EZRules which allows the user to select the rule(s) that fulfil $P_{ed} \geq 90\%$ and $P_{fr} \leq 5\%$ (Westgard, 2016a). However, it has been acknowledged that P_{ed} of 85% (Rishniw *et al.*, 2012; Flatland *et al.*, 2014a) is acceptable in veterinary medicine for point of care testing (POCT) or for haematology where use of a single QCM is traditional in the UK and for which $P_{ed} \geq 0.90$ cannot be achieved with a 1-3s rule; the P_{ed} that can be characterised with use of a 1-3s rule with a single QCM is ≥ 0.85.

Harmonisation of statistical quality control (SQC) practices between laboratories is lacking and the theory of quality systems for laboratory networks has not routinely been put into practice in human medicine (Jassam *et al.*, 2011). CLSI released a document in 2016 that describes a roadmap for the selection of SQC strategies (Parvin, 2017) but due to the complexity of the guidelines the uptake of SQC has not been widespread. Simplified guidelines are required to aid clinical laboratories; some headway has begun for single (Yago and Alcover, 2016) and multirule (Bayat, 2018) SQC procedures, but more work is required for veterinary medicine. More recently,

a report outlining the importance of QC validation and reasons for the lack of implementation amongst veterinary laboratories highlights the risk of misusing QC rules for a method without understanding the associated P_{ed} and P_{fr} (Freeman *et al.*, 2024).

1.1.4 Sigma metrics

Sigma metrics are a statistical tool that can be used for laboratory clinical performance management. They were first developed by Motorola in the 1980s, but were not applied in laboratory quality until 2001 (Nevalainen *et al.*, 2000). The sigma metric management system has been built on a foundation of previous work covering decades, including the earliest work that included bias and imprecision into analytical total error, systematic error, and allowable total analytical error (Westgard and Barry, 1986).

There are two methodologies for using sigma metrics:

1) Counting defects/errors and expressing them as defects per million (DPM) which is then converted to a 0-6 sigma metric where 6 is regarded as world class performance (3.4 defects per million) and 2-3 as an indication of unstable performance (Westgard, 2001).

2) A measurement of the variation of the measurement process using bias, imprecision, and total allowable error (TE_a) to reach 6-sigma (Hens *et al.*, 2014).

Simplistically, the six-sigma metric number can be thought of as the number of times the observed analytical variation (SD) fits into the allowed analytical variation.

There have been many studies addressing the development of sigma metrics in the clinical laboratory, such as their implementation into the chemistry laboratory as a performance indicator where actions were required for parameters < 3 sigma (Mao *et al.*, 2018) and evaluating the performance of non-verified third-party reagents which require strict monitoring (Cao and Qin, 2018). Diagnostic manufacturers are also incorporating sigma metrics into their designs to help choose product requirements that optimise certain assays (Petrides and Schneider, 2018). Bayet (2018) highlighted the calculation of sigma metrics and what should be observed for the medical laboratory for analytical/clinical context by using a tolerance limit or TE_a incorporated into the calculation (Bayat, 2018). The application of sigma metrics in veterinary medicine has been increasing (Farr and Freeman, 2008; Rishniw *et al.*, 2012; Rishniw and Pion, 2016; Flatland and Freeman, 2018b)

Sigma metrics can be calculated using the following formula.

$$\text{Sigma metric} = [(TE_a - \text{Bias}) / CV]$$

Where TE_a describes total allowable error, all terms are expressed in concentration units. Percentage units may also be used, and the equation is as follows.

$$\text{Sigma metric} = (TE_a\% - \text{Bias}\%)/CV\%$$

This calculation combines bias, imprecision, and total allowable error, making it a comprehensive analytical performance indicator (Westgard, 2013b). The

analytical sigma metric has been used in many applications, most notably on human chemistry methods, (Schoenmakers *et al.*, 2011) immunoassay (Marques-Garcia *et al.*, 2015) and enzyme methods (Zhang *et al.*, 2015) and for point of care equipment (Fei *et al.*, 2015). Jassum *et al.* (2011) used sigma metrics to assess performance in a cohort of UK laboratories. Sigma metrics was also used to monitor ST and XN series of Sysmex haematology analysers from hospital sites where they found their methods were consistently greater than 4 sigma (Harrison and Jones, 2017) which is also consistent with another report based on the Sysmex analysers (Seo *et al.*, 2015). There is, however, a lack of a harmonised approach to the use of the sigma calculation between clinical laboratories which results in the inability to directly compare their results. A network of clinical accredited laboratories in Egypt were used to address this issue. This study highlighted the importance of using the same TE_a rather than an individual TE_a for each laboratory (El Sharkawy *et al.*, 2018) to achieve comparability.

Because sigma metrics are an excellent quality tool it was included in this research as an important reflection of analytical capability which allows comparability across instruments that all use the same methods. It relates intuitively with the Westgards' sigma rules (Westgard, 2016a) and reflects on the degree of risk of error associated with an analytical process. As described by Westgard, methods that perform at 6 sigma or greater are easily controlled with a simple rule and they also utilise less QCM as a result; improving cost efficiency. Those with 6 sigma may be controlled with a simple rule (1-3s, 1-2.5s, with 1 or 2 measurements) while the methods at 4 to 5 sigma will require the addition of a 4^{th} rule as well as a multirule such as

($1_{3s}/2_{2s}/R_{4s}/4_{1s}$) which would require 4 control measurements and larger numbers of data points. Performance for methods at 3 to 4 sigma will require a more complex multirule and/or additional data of 4 measurements; these are less easily controlled and utilise more QCM (Westgard, 2016a). Sigma metrics, by definition, will change if performance changes, reflecting fluctuations in the observed CV and Bias over defined time intervals.

This literature supports the previous claims from Westgard (Westgard and Westgard, 2017; Westgard, 2001; Westgard, 2013b) who have been long-term champions of six sigma quality management approach to clinical laboratories. The above studies are mainly from human medicine, but the concepts can be adapted for veterinary medicine.

A comprehensive systematic review (Antony *et al.*, 2018) demonstrated the introduction and application of six Sigma in a healthcare setting as a quality improvement opportunity. This review gives an excellent retrospective analysis of the literature which covers 6 continents and demonstrates the top 5 benefits by applying six sigma as (1) increased patient satisfaction, (2) increased processing speed, (3) revenue enhancement, (4) cost savings and (5) defect reduction. This study also showed that these benefits vary from one continent to another. The various software and statistical tools are discussed as well as the biggest challenges, data availability, cultural differences, and resistance to change. Understanding sigma tools, techniques and top management involvement were the main elements identified for successful implementation. Overall, the growing interest in six sigma was noted but the lack of a roadmap is evident in the literature. *Quality Management and six sigma* (Coskun, 2010) is another excellent resource

towards six sigma total quality management, including a chapter in the context of developing countries that outlines some additional challenges to implementing six sigma where quality resources and infrastructure is primitive therefore parallels can be drawn for the veterinary industry and their framework can be easily applied.

1.1.5 Repeat patient testing vs commercial QCM

Quality control procedures are important in ensuring that quality of operations is maintained for haematology analysers to produce reliable results (Flatland and Weiser, 2014). The traditional method of ensuring quality controlled haematological data is with the use of one or more QCM which typically contain human derived blood products with added components including stabilisers and preservatives so that product integrity is maintained (Fraser and Peake, 1980; Franzini and Ceriotti, 1998; Hubert *et al.*, 2010). QCM is costly but up until now a necessary expense for the clinical laboratory. They also have a limited shelf life with a designated expiration date, after which they should not be used (Flatland and Vap, 2012).

The QCM and QC rules applied should achieve a high P_{ed} and low P_{fr} (Westgard, 2001). They should be commutable with the samples from diseased or healthy patients, (Vesper *et al.*, 2007) a condition that can be fulfilled by using patient samples of veterinary interest as QCM. Typical veterinary haematology patient samples are Ethylenediaminetetraacetic acid (EDTA)-anticoagulated whole blood. EDTA samples are inherently labile and can deteriorate over hours depending on storage conditions. Therefore QCM

is not a very comparable product due to the longer shelf life and suboptimal matrix effects. The inherent instability of the EDTA samples may be manipulated deliberately for use in error detection rather than the commercially available QCM. Storing EDTA at 4^0C and 24^0C has shown a degree of significant changes in haematology measurands, particularly for platelets but it was not statistically significant until 48 hours when analysed using an ADVIA 120 haematology analyser (Furlanello *et al.*, 2006). A comparison between the ADVIA 120 and a Sysmex XT-2000iV (similar model to the haematology instruments used in this study) reported the Sysmex analyser performed favourably over the ADVIA and stability for most measurands was satisfactory (Bauer *et al.*, 2012). This anticipated degree of change can be exploited in a process called repeat patient testing (RPT) which relies on the principle that under prescribed circumstances such as specific storage conditions and defined time intervals that there is an expected variance between a baseline result and repeated results for the same sample. A repeated measurement exceeding the expected threshold suggests instability within the analytical system (Westgard, 2016a). The use of retained patient specimens as a quality measure in haematology has been reported in the literature in 1966 on human specimens and this method of quality assurance was evaluated in the decades that followed, (Dutra, 1966; Riddick, 1972; Carstairs *et al.*, 1977; Orser, 1983; Barnett and Weisbrot, 1979) but its implementation varied greatly. Cembrowski *et al*., (1988) set out to try to establish some optimisation for this area of research (Cembrowski *et al.*, 1988) and was able to confirm previous reports (Brittin *et al.*, 1969) that most haematology measurands are stable up to 40 hours (when using a

Coulter S+IV haematology analyser) and validated that control limits (i.e. selecting a control rule) may be derived from commercial quality control.

Westgard (2016b) recommends that RPT begins with gathering a set of pilot data, over 20 days, that can be used to generate control limits. These limits can then be used to determine whether repeat patient testing – quality control (RPT-QC) results are “in-control” or “out of-control”. After this initial phase, the fresh RPT sample is measured on day one and again within a defined time interval and with defined storage conditions. It is also recommended to use patient samples within reference intervals as they represent a significant proportion of caseload and may avoid extreme variabilities between measurements that may occur at concentration outside of the reference interval (Westgard, 2016b). The difference in the repeated analysis is then calculated and plotted using power function charts similarly to other QC procedures (Cembrowski *et al.*, 1988). There have been some positive outcomes in the literature (Vanyo *et al.*, 2018; Flatland and Freeman, 2018b) indicating good potential for application in veterinary laboratories.
RPT-QC has become of increased interest since commutability of commercial QCM has not been investigated for haematology in veterinary medicine and there is not a wide selection/availability of veterinary quality control material. In addition to having an animal matrix, RPT-QC has also gained attention as a less expensive option compared to commercial QCM as it uses excess veterinary samples routinely available within the laboratory. RPT validation across a network of laboratories for haematology remains novel, a study looking at RPT as a quality control (RPT-QC) for biochemistry was published in 2020 (Flatland and Freeman, 2020) and is an excellent

template for implementing RPT-QC for biochemistry testing in a veterinary laboratory. This is important research as to date the unavailability of veterinary specific quality control material is problematic and forces veterinary laboratories to rely on human based QCM rather than a species-specific control material. This biochemistry study also demonstrated the technical and statistical procedure for RPT-QC development and the results showed that error detection was comparable with commercial quality control material and achievable using ASVCP quality goals.

There are some limitations regarding RPT-QC for veterinary haematology such as access to fresh canine EDTA samples daily or when the laboratory has been closed for multiple days (such as extended closure periods/holidays). These limitations must be considered in the design of RPT-QC prior to data collection for embarking with RPT-QC.

1.1.6 External quality assurance

External quality assurance (EQA) or proficiency testing (PT) is included as part of a laboratory quality management system (QMS) as a formal comparative testing program so that laboratories can assess and monitor performance when compared to a peer group using the same method (Camus *et al.*, 2015). This was introduced to the medical laboratory more than 60 years ago to show that different results were obtained from the same specimen by various laboratories. The results generated were utilised to standardise processes and ensure laboratories were using methods based

on manufacturers specifications and calibration recommendations (Miller *et al.*, 2011b).

It's expected that laboratories can generate uniform results and the emphasis of EQA is based on bias rather than imprecision. ISO/IEC Guide 17043:2023 (ISO/IEC, 2023) and CLSI GP27-A2 (CLSI, 2007) have defined EQA/PT programs as interlaboratory comparisons that assesses performance of not only the analytical equipment but the competence of the personnel.

As mentioned earlier, veterinary laboratories have the added complication that most EQA/PT schemes are based on a human-derived specimens and, therefore, may not give meaningful results due to non-commutability of the EQA specimen (Unsal *et al.*, 2010). This is in addition to the high cost associated with an EQA program that may not be affordable for veterinary laboratories. An informal comparative testing program may be a useful alternative for veterinary laboratories and in-clinic veterinary laboratories that opt to participate in a comparative testing program, as recommended by ASVCP (Flatland *et al.*, 2013). Camus *et al.*, (2015) describe an overview of alternative comparative testing for veterinary analysers including POCT analysers using common reference intervals should have comparable performance to a commercial, reference or university laboratory to verify and monitor individual analyser performance. There are few reports of comparison studies using veterinary specimens but two reports using canine samples comparing POCT veterinary haematology analysers to a reference analyser demonstrated poor performance and the need for ongoing performance evaluation (Becker *et al.*, 2009; Rishniw and Pion, 2016).

1.1.7 Reference intervals

Reference intervals (RIs) are a tool that allow the comparison of results between individual animals or a group of animals. Described first by Wootton et al in 1951, blood samples from 100 healthy human volunteers were analysed for 18 biochemistry measurands and if a normal distribution was not present then this was corrected by log transformations and the upper and lower limits were assessed. Results that were outside the 10% limits was regarded as suspicious while 1% outside the limits was deemed abnormal (Wootton *et al.*, 1951). This initial work was further developed as reference intervals as we know them today by using the limits of two standard deviations from the mean of a healthy populations and the central 95% of this group is called the reference interval (Gräsbeck and Fellman, 1968; Grasbeck, 1969). Reference intervals did not appear in veterinary medicine until 1978 (Lumsden and Mullen, 1978) and it wasn't until 2012 that the American Society of Veterinary Clinical Pathology (ASVCP) Quality Assurance Laboratory Standards (QALS) committee extended the recommendations to veterinary laboratories (Friedrichs *et al.*, 2012) from the 1970 standard, (updated in 2008) for production of human reference intervals from the International Federation of Clinical Chemistry (IFCC) (Horowitz, 2008).

There are certain factors to consider when establishing RIs and these conditions have been recommended to be carefully considered by ASVCP and CLSI. Some of these conditions include the population selection relevant

to the breeds seen in the laboratory's patients; patient signalment criteria that define breed, health status, age, sex, reproductive status, nutrition, demographics (biological, clinical & geographic); sample collection standardisation; and sample quality (absence of interfering substances). The development of RIs as recommended by the IFCC is complex and these complexities led to the lack of use or incorrect use of these guidelines (Ceriotti *et al.*, 2008). The revision in 2008 recognised the difficulties that individual laboratories faced in developing their own RIs and permitted the use of RIs established elsewhere to be adopted by transference and verification (CLSI, 2008). It's been recognised that RIs are "one of the most widely used medical decision-making tool" (Horn and Pesce, 2003) but misclassification as initially recognised by Schneider can occur due to the following causes: intra-individual variation, inter-individual variation, and analytical variation (Schneider, 1960).

As part of a harmonisation process for laboratories, the continued use of common RIs should be evaluated. It is not always feasible for individual laboratories to establish their own RIs and furthermore the use of common RIs is widely accepted and now an important part of giving standardised information from multiple laboratories (Jones *et al.*, 2004). Laboratories can implement RIs via transference studies which refers to a (Horn and Pesce, 2003) validation of RIs by comparing results from 20 healthy reference individuals from the population to determine the application of the existing RIs and if they are relevant to a new method and/or instrument (Horn and Pesce, 2003; Friedrichs *et al.*, 2012). Transference validation assumes the demographic is comparable and corresponds to the original RIs, (Geffré *et*

al., 2009) it should not be used as an initial calculation of RIs as de novo production of RIs is the still the preferred method (Solberg and Stamm, 1991). Following the collection of data and comparison to the RIs in use, they are valid only when one or two fall outside the reference limits. If three or more fall outside the reference limits, then transference must be rejected and re-evaluated with a second set of 20 results from healthy individuals before determining final rejection or acceptance of the candidate reference interval. Where zero values fall outside the reference limits then this suggests the RIs are too wide which may result in failure to detect abnormal levels. Evaluation of a scatter plot of the results from the 20-40 healthy individuals is recommended to help determine if the reference interval may appear to be too wide and lacking sensitivity in detection of abnormality at either low or high reference interval limits.

Where bias is observed between the new and existing method then regression statistics (if available) or other corrections may be used to adjust the Ris. Where excessive imprecision is evident, then the use of transference should be re-assessed (Horn and Pesce, 2003; Friedrichs *et al.*, 2012).

1.1.8 Aims of this research project

There are no known veterinary laboratory haematology systems with multiple laboratories that operate as an integrated harmonised laboratory system. There are publications in the human literature particularly in clinical chemistry (Jassam *et al.*, 2011; Theodorsson, 2016) that promote this approach for laboratories with multiple instruments and/or multiple laboratories, but publications regarding the successful implementation of laboratory harmonisation are few. There are multiple human Scandinavian and European laboratories that are using common reference intervals, presumably based on harmonisation, but the methodology used is not in the public domain and the degree of success in maintaining harmonisation is not reported. There are only three publications describing pilot projects for the implementation of repeat patient testing for quality control in the veterinary literature (Flatland and Freeman, 2018b; Flatland and Freeman, 2020; Vanyo *et al.*, 2018).

The purpose of the proposed investigation is to address several aspects of the development of a novel system for the five separate laboratories of the VPG into a single laboratory quality system for haematology.

These projects all involve novel studies, as there are few or no relevant peer-reviewed publications in the veterinary literature.

Objectives of this project

To harmonise a group of veterinary haematology analysers within an veterinary network so that samples could be distributed cross functionally with minimal risk of misclassification between the systems.

The specific aims of the project:

- Evaluation of the current performance of 6 Sysmex-XT 2000iV haematology analysers (one or two at each VPG laboratory) based on commercially available quality control materials and use of patient samples to assess instrument imprecision and bias.
- The results of evaluations are necessary to take action to achieve harmonisation, verify harmonisation and form the quality performance basis used for QC validation and choice of QC rules to achieve a high P_{ed} and low P_{fr} with commercial quality control materials. The reference haematology instrument at VPG/Cork is a basis for comparison of the other haematology analysers within the VPG system.
- Evaluation of a novel QC approach in veterinary laboratory medicine: RPT as a method for low-cost, veterinary-specific quality control for veterinary haematology.
- Compare RPT-QC with traditional commercial quality control results to determine if it provides advantages that may be beneficial in

veterinary testing and ensuring stable system performance for veterinary testing.

- Design, evaluate and implement an internal comparative testing scheme for haematology to ensure that all laboratories are harmonised regarding production of results across a range of clinically important results, as well as allowing the ongoing use of common reference intervals.
- Design and implement an internal EQA pilot program for comparative testing for in-clinic POC analysers and the reference laboratory as an offering for clinics using the services of the VPG. This will help clients electing to participate in the EQA scheme to achieve goals for hospital certification and provide ongoing education regarding laboratory quality leading to a stronger laboratory-client bond and ultimately improving the ongoing quality of patient care.

CHAPTER 2: EVALUATION OF SYSMEX XT-2000*iV* ANALYSER PERFORMANCE ACROSS A NETWORK OF FIVE VETERINARY LABORATORIES USING A COMMERCIALLY AVAILABLE QUALITY CONTROL MATERIAL

This is the first and fundamental stage of the research approach in which performance evaluation was conducted for each individual analyser in three separate stages to optimise and harmonise the haematology analysers within the laboratory network. Primary analysis and data collection was undertaken by S Daly with guidance from Dr Peter Graham and Dr Kathy Freeman as co-authors on the manuscript produced from this research and represents original work.

Adaptations: This chapter has been changed from its previously published version to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

Reference: This chapter has been published as Daly SC, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000*i*V analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021; 50: 568–578. doi:10.1111/vcp.13016

Received: 18 January 2021 | Revised: 12 March 2021 | Accepted: 17 March 2021

DOI: 10.1111/vcp.13016

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

ORIGINAL ARTICLE

Evaluation of Sysmex XT-2000*iV* analyzer performance across a network of five veterinary laboratories using a commercially available quality control material

Susan Daly[1] | Kathleen P. Freeman[2] | Peter A. Graham[3]

[1]SYNLAB-VPG/Cork, Cork, UK

[2]SYNLAB –VPG/Exeter, Exeter, UK

[3]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, Brushville House, Dosco Business Park, South Douglas Road, Cork T12XR6V, UK.
Email: susan.daly@synlab.co.uk

Abstract

Background: Laboratory and instrument harmonization is seldom reported in the veterinary literature despite its advantages to clinical interpretation, including the use of interchangeable results and common reference intervals within a system of laboratories.

Objectives: A three-step process was employed to evaluate and optimize performance and then assess the appropriateness of common reference intervals across a network of six Sysmex XT-2000*i*V hematology analyzers at 5 commercial veterinary laboratory sites. The aims were to discover if harmonization was feasible in veterinary hematology and which quality parameters would best identify performance deviations to ensure a harmonized status could be maintained.

Methods: The performance of 10 measurands of a commercially available quality control material (Level 2—Normal e-CHECK (XE)-Hematology Control) was evaluated during three 1-month time periods. Precision and bias were assessed with Six Sigma, American Society of Veterinary Clinical Pathology (ASVCP) total error quality goals and biologic variation (BV)-based quality goal approaches to performance measurement.

Results: Instrument adjustments were made to 1 analyzer twice and 3 analyzers once between evaluations to improve performance and achieve harmonization. Sigma metrics improved from 37/50 > 6 to 58/60 > 6 and to all >5 over the course of the harmonization project. BV-based quality goals for desirable bias and for laboratory systems of 0.33 × CV_I (within-subject biologic variation) were more sensitive and useful for assessing performance than the ASVCP total error goals.

Conclusions: Optimization and harmonization were achieved, and because BV-derived bias goals were achieved, common reference intervals could be implemented across the network of analyzers.

KEYWORDS
analytical performance, biologic variation, hematology, quality system, sigma metrics, Sysmex, veterinary

2.1 Introduction

Interlaboratory or intralaboratory comparisons across multiple instruments in human clinical laboratories are well documented (Jassam *et al.*, 2011) (Nevalainen *et al.*, 2000; Miller *et al.*, 2011a; Harrison and Jones, 2017) however similar evaluations in the veterinary literature are few (Farr and Freeman, 2008). This could be attributed a multitude of factors, including the fact that human laboratory regulations are not applied to veterinary laboratories (Farr and Freeman, 2008; Sacchini and Freeman, 2008); results of comparative external quality assurance (EQA) performance are not routinely available and/or shared; and a lack of reporting of efforts or possibly an absence of such efforts to harmonise results within or between veterinary laboratory businesses and networks. Harmonisation of methods across laboratories means the standardisation of results across methods and laboratories when no certified reference material is available. It is based on an agreed upon standard for performance amongst participating laboratories to ensure that results have sufficient uniformity to be used interchangeably across a laboratory network for an individual patient (Miller *et al.*, 2011a). Synlab Veterinary Pathology Group (Synlab-VPG) operates six Sysmex XT-2000*i*V analysers within their network of 5 laboratories located in the UK and Ireland providing an environment in which to explore a veterinary haematology harmonisation approach for the benefit of group clinical pathologists, client veterinarians and ultimately patients. Were a successful approach identified, it could provide a useful template to implement a harmonisation process for the veterinary haematology community.

At the core of the harmonisation process is a set of quality goals that are used to measure acceptable performance and progress towards harmonisation. Such measures include error goals for veterinary haematology provided by the American Society for Veterinary Clinical Pathology (ASVCP) (Nabity *et al.*, 2018) or from internal expert opinion based on QC validation; canine biologic variation goals for bias and CV, and the canine biologic variation goal for a laboratory system ($0.33CV_I$) (Fraser, 1992b). Sigma metrics is also a useful indicator of performance where methods >6 sigma are considered to have 'World Class' performance (Westgard, 2001).

The objectives of this present study were to:

1) Evaluate the performance of the six Sysmex haematology analysers within the five SYNLAB-VPG laboratories using a commercially available quality control material.
2) Optimise the performance of each analyser to ensure the total error budget is contained within ASVCP guidelines or those determined by internal expert opinion.
3) Determine if biologic variation goals for CV and bias and/or laboratory system performance would enable the use of common reference intervals.
4) Determine those quality goals that are the most useful for identifying performance deviations that need to be addressed to maintain the interchangeability of results across the laboratory system.

We hypothesised that the performance across the laboratories would be similar since the same instruments were used, as well as the same reagents and quality control materials. We additionally hypothesised that the quality goals or minimal deviation from them would be achievable.

2.2 Materials and Methods

2.2.1 Analysers

Six Sysmex XT-2000*i*V haematology analysers (Sysmex Corporation, Japan) using veterinary software (version 00-13) were compared amongst five veterinary sites within SYNLAB-VPG. These are designated as instruments 1 to 5 (site 4 has 2 instruments, designated as 4A and 4B). Ten measurands were selected for statistical quality performance analysis as we considered them the most clinically relevant in current veterinary practice: Red blood cell count (RBC), Haemoglobin (HGB), Haematocrit (HCT), Mean corpuscular volume (MCV), Mean corpuscular haemoglobin concentration (MCHC), White blood cell count (WBC), Platelet count (PLT), Red cell distribution width-coefficient of variation (RDW-CV), Plateletcrit (PCT) and Reticulocyte number (RETIC). Additional measurands, within the capability of the analyser and represented in the quality control material (QCM) package insert were not included in the performance evaluation. The methods used to determine these measurands are summarised in SEED 2012 (Development, 2012) and Sysmex -2000*i*V/XT -1800*i*V user manual.

The Sysmex XT-2000*i*V haematology analyser can be operated in open or closed mode; for this study open mode was used, i.e. the probe aspirates the required volume from an open uncapped tube.

2.2.2 Quality control material

A single level of QCM, (level 2 – Normal e-CHECK™ (XE)-Haematology Control) was analysed once per day at instrument start-up, prior to analysis of patient samples, by a fully trained technician according to a standard operating procedure. The QCM was refrigerated when not in use; and opened vials were used within 7-days. Performance was deemed acceptable for inclusion in the study data when the daily internal quality control (IQC) value was within the manufacturer's recommended acceptable range.
The manufacturer's target mean for the QCM were chosen for comparison across the group of analysers.

2.2.3 Data collection

Over 3 separate evaluation periods of November (Month 1), December 2019 (Month 2) and March 2020 (Month 5) each of the laboratories collected approximately 30 consecutive days of QCM result data from six Sysmex XT-2000*i*V analyser. Only 4 sites (5 analysers) could participate in the Month 1 period due to an inability to obtain the Level 2 QCM for one site. The data was exported to Microsoft Excel (Microsoft Excel for Mac 2011, Version 14.7.7 (170905) Last update 14.7.7). Each laboratory sent their individual

spreadsheets for analysis and the data were used to calculate quality evaluation metrics. Intervention recommendations such as optimal QCM handling and instrument adjustment and/or service were applied between each evaluation period based on the results obtained and the likely root causes for detected deviations in performance.

2.2.4 Quality evaluations

To evaluate the observed analytical imprecision and bias, the SD and CV (Jensen and Kjelgaard-Hansen, 2006), bias in units and absolute bias % (Westgard, 2016a) from the manufacturer's QCM target mean were calculated as described previously (Rishniw and Pion, 2016) for all ten measurands from each site/analyser for each evaluation period.
From the initial calculations for imprecision (CV%) and bias (absolute bias %) the total error observed (TE_{obs}) was calculated and expressed as a percentage to allow comparability to analytical quality requirements (Flatland *et al.*, 2018a) using the following formula;

$$TE_{obs}\% = |\text{Bias }\%| + 2 \times \text{CV}\% \text{ (Flatland } et\ al.\text{, 2018) (Nabity } et\ al.\text{, 2018)}$$

A Sigma metric approach was taken to estimate clinical performance for the ten measurands across all sites using the following formula (Westgard, 2001; Westgard, 2013b) and Total Error Allowed (TE_a) from Table 2.

$$\text{Sigma metric} = (TE_a\% - |\text{Bias}\%|) / \text{CV}\%$$

The sigma metric was chosen as an important reflection of analytical capability, allowing comparison across methods. It is considered to be a reflection of risk associated with an analytical process and correlates with the 'Westgard Sigma Rules' (Westgard and Barry, 2008; Westgard, 2016a). According to Westgard, measurands with 6 sigma or greater performance are easily controlled with a simple quality control (QC) rule and low numbers of QCM data, while those with 5-6 sigma can be controlled with a short multirule ($1_{3s}/2_{2s}/R_{4s}$) and low numbers of QCM data. Those methods with 4-5 sigma performance can be controlled with addition of a fourth rule to a multirule ($1_{3s}/2_{2s}/R_{4s}/4_{1s}$). Those methods with 3-4 sigma performance require an even more complex multirole with larger numbers of QCM data, but still may not achieve desirable levels of error detection (Westgard, 2016a). The sigma metric, by definition, will change if performance changes, reflecting fluctuations in the observed CV and Bias over defined time intervals.
For measurands with a sigma metric < 6, a quality goal index (QGI) was calculated using the following formula:

$$QGI = |BIAS\%| / (1.5 * CV\%)$$

This calculation allowed us to determine whether imprecision and/or inaccuracy affects the overall quality goals using the following interpretative guidance: a QGI < 0.8 is considered indicative of imprecision while a QGI > 1.2 is indicative of inaccuracy (bias). QGI results between these two

boundaries indicate both bias and imprecision contributing to < 6 sigma performance (Parry DM, 2009).

Biologic variation goals were applied as three additional performance indicators: biologic variation-based CV, biologic variation-based bias and a biologic variation-based laboratory system goal (Fraser, 1992b). Canine biologic variation goals were chosen because canine specimens are the most common submissions across all the 5 laboratories. We calculated optimal, desirable and minimal values for TE, CV and bias goals based on canine biologic variation data (Flatland *et al.*, 2018a) for applicable measurands and 0.33 X CV_I as the goal for multiple instruments within a laboratory system (Fraser, 1992b). The 0.33 x CV_I goal allowed us to detect differences between means of serial measurements for each measurand to ensure comparability by checking the allowable difference is < 0.33CV_I (Fraser, 1992b). The +/- 0.33 x CV_I goal was evaluated compared to the manufacturer's target mean for the QCM. PCT and RET did not have biologic variation data available. We used the following formulas (Fraser, 1992b).

<u>Biologic variation CV goals:</u>

Optimal CV: < 0.25(CV_I)

Desirable CV: < 0.5(CV_I)

Minimally acceptable CV: < 0.75(CV_I)

<u>Biologic variation bias goals:</u>

Optimal Bias : < 0.125 x $(CV_I^2 + CV_G^2)^{1/2}$

Desirable Bias: < 0.25 x $(CV_I^2 + CV_G^2)^{1/2}$

Minimally Acceptable Bias: $< 0.375 \times (CV_I^2 + CV_G^2)^{1/2}$

Within-subject biologic variation is denoted as $CV_{I,}$ while between subject variation is $CV_{G.}$

2.3 Results

2.3.1 Sigma metrics and QGI

2.3.1.1 Initial evaluation of 5 Sysmex XT-2000iV analysers Month 1

All the haematology measurands chosen for this study for analyser 1 and analyser 2 had sigma metrics >6. Of the remaining 3 analysers, Analyzer 5 had only 7 and Analysers 4A & 4B had only 5 measurands above six-sigma (Table 2.1).

Table 2.1: Initial results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 1 (November) evaluation.

Measurand	**Analysers**					
	1	**2**	**3**	**4a**	**4b**	**5**
WBC	6.54	10.64	N/A	9.76	9.38	8.35
RBC	15.21	13.00	N/A	3.88	5.51	7.10
HGB	7.24	23.00	N/A	6.00	2.82	5.57
HCT	9.38	12.70	N/A	3.48	4.92	4.08
MCV	11.50	12.20	N/A	12.34	12.40	8.15
MCHC	9.28	10.90	N/A	5.65	9.37	8.12
PLT	6.15	9.05	N/A	7.35	9.7	16.97
RDW-CV	18.00	25.36	N/A	0.47	0.596	2.03
PCT	6.06	8.56	N/A	6.50	7.75	7.00
RETIC No.	6.20	7.63	N/A	3.20	4.57	7.03
N σ >6	10	10	N/A	5	5	7
N σ 5.0-5.9	0	0	N/A	1	1	1
N σ 4.0-4.9	0	0	N/A	0	2	1
N σ 3.0-3.9	0	0	N/A	3	0	0

σ N < 3	0	0	N/A	1	2	1
Intervention	None	None	N/A	ISA	ISA	ISA
	Improved QCM handling					

N/A – not available due to inability to source QCM; ISA – Instrument service and Adjustment; N – number

2.3.1.2 QGI Month 1

The QGI interpretation for those measurands performing < 6 sigma is summarised in Table 2.4. Bias was more often a concern (14/24) than imprecision (8/24). Both bias and imprecision accounted for 2 occurrences. Adjustments with instrument service at analyser 4A and analyser 4B and analyser 5 were conducted to help eliminate bias. Improved control material handling was instituted across all the laboratories before additional evaluations were undertaken in Month 2.

2.3.1.3 Follow up Month 2 and Month 5 evaluation on 6 Sysmex XT-2000iV analysers

As demonstrated in Tables 2.1, 2.2 and 2.3, sigma metrics generally improved over the following 2 evaluation periods resulting in all but 2 measurands at one site (Analyser 4A) achieving sigma greater than 6. All sites and all measurands achieved sigma greater than 5. An instrument service visit was required for Analyzer 3 and an additional visit required for

Analyzer 5 to achieve this improvement in performance. QGI review of the remaining 2 measurands on Analyzer 4A identified imprecision for PLT and bias was implicated for RBC (Table 2.4).

Table 2.2: Follow up results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 2 (December) evaluation.

Measurand	**Analysers**					
	1	**2**	**3**	**4a**	**4b**	**5**
WBC	14.12	6.81	8.87	6.80	7.13	9.53
RBC	15.11	14.77	11	13.96	12.10	17.75
HGB	19.94	12.68	12.82	14.83	13.54	20.74
HCT	2.92	13.14	2.20	9.19	12.77	10.94
MCV	16.68	13.44	5.20	12.74	12.83	8.02
MCHC	14.13	12.05	5.71	10.15	9.53	11.32
PLT	8.07	7.62	4.94	9.73	6.29	6.94
RDW-CV	20.40	20.40	13.64	18.68	17.71	24.71
PCT	7.61	9.92	4.55	9.01	6.43	7.06
RETIC No.	5.59	6.48	7.83	5.59	7.10	5.04
N σ > 6	8	10	5	9	10	9
N σ 5.0-5.9	1	0	2	1	0	1
N σ 4.0-4.9	0	0	2	0	0	0
N σ 3.0-3.9	0	0	0	0	0	0
σ N < 3	1	0	1	0	0	0
Intervention	None	None	ISA*	None	None	ISA*

*ISA – Instrument service and Adjustment

Table 2.3: Follow up results and sigma metric summary for the selected measurands for 6 Sysmex XT-2000iV Haematology Analysers within a system of veterinary laboratories for month 5 (March) evaluation.

Measurand	**Analysers**					
	1	**2**	**3**	**4a**	**4b**	**5**
WBC	7.62	7.78	6.41	9.22	7.97	10.34
RBC	14.22	18.09	14.2	5.24	12.01	9.34
HGB	15.97	24.58	30.04	17.50	18.45	10.08
HCT	13.82	13.92	9.05	15.05	16.10	7.85
MCV	13.60	14.91	10.07	11.12	10.67	8.43
MCHC	13.50	13.12	8.51	10.57	12.55	10.70
PLT	7.60	9.05	6.80	5.12	7.30	7.65
RDW-CV	19.17	13.10	19.53	14.40	16.52	14.55
PCT	7.42	7.57	7.35	7.9	7.9	8.04
RETIC No.	6.90	8.00	8.50	7.60	6.00	8.35
N σ > 6	10	10	10	8	10	10
N σ 5.0-5.9	0	0	0	2	0	0
N σ 4.0-4.9	0	0	0	0	0	0
N σ 3.0-3.9	0	0	0	0	0	0
σ N < 3	0	0	0	0	0	0

Table 2.4: Summary of investigation of measurands with sigma metrics < 6 for 6 Sysmex XT-2000iV Hematology Analysers in a system of veterinary laboratories using a commercially available quality control material.

Site	Measurand	Sigma metric	QGI	Interpretation of QGI	CV-obs %	Achieved CV_A BV Goal	Bias-obs %	Achieved Bias BV Goal	Achieved 0.33 CV_I Goal
1	RETIC	5.59	0.69	Imprecision	6.04	N/A	6.253	N/A	N/A
	HCT	2.92	1.53	Bias	3.18	YES – Des	0.72	YES - Opt	YES
3	HCT	2.20	6.56	Bias	3.08	YES – Des	3.20	NO	YES
	MCV	5.20	0.16	Imprecision	1.31	YES – MA	0.19	YES - Opt	YES
	MCHC	5.71	2.21	Bias	1.30	Yes- MA	2.542	NO	NO
	PLT	4.94	2.597	Bias	3.84	YES – Des	1.01	YES - Opt	YES
	PCT	4.55	0.04	Imprecision	5.42	N/A	0.32	N/A	N/A
4a	PLT	5.12	0.17	Imprecision	3.72	YES – Des	0.93	YES - Opt	YES
	RBC (Month 1)	3.88	1.22	Bias	1.75	YES – Des	3.22	NO	NO
	HCT	3.48	1.32	Bias	1.70	YES – Des	4.06	NO	NO
	RDW-CV	0.47	1.52	Bias	3.61	NO	8.27	NO	NO
	RETIC	5.59	0.77	Imprecision	5.92	N/A	6.87	N/A	N/A

	RBC (Month 5)	5.24	6.67	Bias	1.60	YES - Des	1.60	YES - Des	NO
	MCHC	5.65	0.56	Imprecision	1.43	YES - MA	1.87	NO	YES
	RETIC	3.20	1.13	Both imprecision and bias	8.15	N/A	13.90	N/A	N/A
4b	RBC	5.51	0.55	Imprecision	1.57	YES - Des	1.32	YES - Des	YES
	HGB	2.82	1.66	Bias	1.88	YES - Des	4.69	NO	NO
	HCT	4.92	0.99	Both Imprecision and bias	1.56	NO	2.33	YES - MA	NO
	RDW-CV	0.60	1.75	Bias	3.10	NO	8.15	NO	NO
	RETIC	4.75	53.71	Bias	5.61	N/A	14.37	N/A	N/A
5	HGB	5.57	3.77	Bias	0.89	YES - Opt	5.04	NO	NO
	HCT	4.08	1.72	Bias	1.49	YES - Opt	3.88	NO	NO
	RETIC	5.04	0.29	Imprecision	7.29	N/A	3.217	N/A	N/A
	RDW - CV	2.03	27.75	Bias	1.61	YES- Des	6.71	NO	NO

Tools used for investigation included Quality Goal Index and Relationships of Observed Coefficient of Variation, Observed Bias, and Canine Biologic Variation-based Quality Goals.

RBC = Red blood cell count; HGB = Haemoglobin; HCT = Haematocrit; MCV = Mean Cell Volume; RET =Reticulocyte number; RDW-CV = Red Cell Distribution Width- Coefficient of variation.

QGI = Quality Goal Index (%) = Abs Bias (%)/1.5 CV(%), Interpretation: The QGI reflects the extent to which both bias and precision meet their respective quality goals without considering goals for CV and bias as separate entities. The QGI is interpreted as follows: QGI < 0.8 = imprecision and QGI > 1.2 = inaccuracy, whereas QGI of 0.8 – 1.2 = both imprecision and inaccuracy.**; CV-obs %= Observed analytical Coefficient of Variation; Bias-obs % = Observed analytical Absolute Bias (%); Opt = optimal; Des = desirable; MA = Minimally acceptable; BV = Biologic Variation; CV_I= Within Subject biologic variation coefficient of variation; CV_A = analytical coefficient of variation; N/A = not available for this measurand**

2.3.2 Biologic variation goals

2.3.2.1 Initial evaluation of BV goals on 5 Sysmex XT-2000iV analysers

Achievement against CV, bias and $0.33CV_I$ goals could only be assessed for those 8 measurands for which canine biologic variation (BV) data were available. No BV data were available for PCT and RETIC.

Analytical CVs achieved optimal or desirable CV goals for analyser 1 and analyser 2. Optimal or desirable CV goals were achieved for 7 of the 8 measurands for analyser 5, 6 of the 8 measurands for analyser 4A and analyser 4B. (Table 2.5)

Table 2.5: Summary of CV biologic variation goals that reached optimally or desirably acceptable levels for 8 measurands where biologic variation data is available, assessed during three evaluations.

Analyser number	CV BV goal		
	Month 1	Month 2	Month 5
1	8	8	8
2	8	8	8
3	N/A	6	8
4A	6	8	8
4B	6	8	8
5	7	8	8

Not applicable (N/A) is for an analyser that was not using correct quality control material during the first evaluation.

Optimal or desirable bias goals based on biologic variation were achieved for 7 measurands for Analyser 1. Analyser 2 was able to achieve optimal or desirable bias goals based on biologic variation for 6 measurands. Analyser 4A and Analyser 4B were able to achieve optimal or desirable bias goals for 2 and 4 measurands respectively. Analyser 5 was able to achieve optimal or desirable bias goals for 3 measurands but failed to achieve minimally acceptable bias goals for all other measurands. (see Table 2.6)

Table 2.6: Summary of bias biologic variation goals that reached optimally or desirably acceptable for 8 measurands where biologic variation data is available assessed during three evaluations.

Analyser number	Bias BV goal		
	Month 1	Month 2	Month 5
1	7	7	7
2	6	8	6
3	N/A	5	8
4A	2	7	7
4B	4	8	6
5	3	7	6

Not applicable (N/A) is for an analyser that was not using correct quality control material during the first evaluation.

For Analyser 1, the 0.33 CV_I goal for laboratory system performance for QCM was achieved for 7 measurands for which BV information available. For Analyser 2, 0.33 CV_I goal for lab system was achieved for all measurands except PLT. For Analyser 4A, the goal was achieved for 3 of the 8 measurands (WBC, MCV, MCHC) but not for the remaining measurands. For Analyser 4B, the 0.33 CV_I goal was achieved for 5 measurands but not for

HGB, HCT, RDW-CV. For Analyser 5, the 0.33 CV_I goal was achieved for 2 out of 8 measurands (MCV, MCHC) (see Table 2.7).

Table 2.7: Summary of 0.33 x CV_I biologic variation goal for multiple instruments within a laboratory system that reached optimal or desirable acceptable goals for 8 measurands where biologic variation data is available, assessed during three evaluations.

Analyser number	0.33 CV_I BV goal		
	Month 1	Month 2	Month 5
1	7	6	8
2	7	6	6
3	N/A	5	6
4A	3	7	8
4B	5	6	7
5	2	6	7

Not applicable (N/A) is for an analyser that was not using the correct quality control material during the first evaluation.

2.3.2.2 Follow up evaluation of BV goals on 6 Sysmex XT-2000iV analysers in Months 2 and 5

Analyser 3, now included in month 2, achieved optimal or desirable CV goals based on biologic variation for 6 of the 8 measurands and for all 8 measurands in month 5. The other sites achieved optimal or desirable CV goals for the 8 measurands at both follow up evaluations (Table 2.5). Analyser 1 achieved optimal or desirable bias goals for the same 7 measurands for month 2 and month 5. Analyser 2 achieved optimal or desirable bias goals 8 measurands for month 2 and 6 measurands for month 5. Analyser 3 achieved optimal or desirable bias goals for 5 measurands for Month 2 and all 8 measurands for Month 5.

Analysers 4A and 4B improved during Month 2, Analyser 4A met optimal or desirable bias goals for 7 measurands where biologic variation data was available for month 2 and month 5. Analyser 4B achieved optimal or desirable bias goal for all 8 measurands in month 2 but only achieved 6 of 8 measurands in month 5.

Analyser 5, in contrast to Month 1, could achieve optimal or desirable bias goal for 7 measurands for month 2 and 6 measurands for month 5 (Table 2.6).

The 0.33 CV_I goal was achieved by analyser 1 for 6 measurands, it did not achieve this goal for 2 measurands in Month 2 but this resolved by Month 5 where the goal was achieved for 8 measurands. Analyser 2 could not achieve goal for 2 measurands in Month 2 and in Month 5 also achieved this goal for 6 of 8 measurands but not for HCT and MCHC.

Analyser 3 achieved this goal for 5 of 8 measurands in Month 2. This increased to 6 of 8 measurands in Month 5.

Analyser 4A achieved this goal for 7 of 8 measurands in Month 2 but by Month 5 achieved goals for all 8 measurands. Analyser 4B achieved this goal for 6 measurands in Month 2 and achieved goal for 7 of 8 measurands in Month 5 except HCT.

Analyser 5 achieved this goal for 6 of 8 measurands in Month 2. In Month 5, 7 measurands achieved this goal, apart from MCHC (Table 2.7).

2.3.2.3 Comparison of performance measures

Tables 2.5, 2.6 and 2.7 summarise the biologic variation goals for bias, CV and 0.33 CV_I goal for laboratory system performance for QCM that reached optimal or desirable acceptable goals for 8 measurands where biologic variation data is available, assessed during three evaluations. Biologic variation goals for bias and 0.33 CV_I goal for laboratory system proved discriminatory when compared to ASVCP or expert opinion where sigma metrics was < 6 sigma.

2.4 Discussion

This is the first report comparing six Sysmex analysers in a network of five veterinary laboratories in the veterinary peer-reviewed literature. We hypothesised that performance across the laboratories would be similar and would have minimal deviations from quality goals. This was true for some sites, but this study has also demonstrated that 4 of the 6 analysers initially exhibited differences in performance, despite initial harmonisation efforts, i.e. using the same instruments, reagents, QCM, processing frequency of QCM and the same technical training. This study demonstrated that at certain times, Sysmex haematology analysers may fail to meet quality requirements (Rishniw and Pion, 2016), particularly for the erythroid measurands.
To monitor any degree of individuality for a network of Sysmex analysers we found sigma metrics to be a valuable indicator of performance related issues. We decided to use this approach for several reasons; it is a universal

benchmark that lends itself well to multiple site evaluation and continuous comparison. It incorporates both bias and imprecision and as sigma metric values decrease the QC rule failure increases (Litten, 2017). We also used quality goals based on BV as we wanted to minimise analytical noise in serial results and ensure that bias and/or imprecision would not mask physiological changes (Fraser, 1992b; Petersen *et al.*, 2002). This would also be useful for ongoing patient monitoring and for the continued use of common RI within this network of laboratories.

We calculated 170 sigma metrics from ten measurands (Table 2.1, 2.2 and 2.3) and from those, 24 (14%) did not achieve > 6 sigma for 'world class' performance. For these, the QGI was useful to analyse potential problems with those measurands by predicting whether the underperformance was due to imprecision or bias or both and informing choice of intervention and adjustment remedies.

Bias was more likely to be a contributing factor than imprecision based on the QGI (58 % of < 6 sigma occurrences). The high degree of bias in comparison to imprecision was not unexpected as managing bias is a regular occurrence in a clinical laboratory but we needed to firstly ascertain how the range of bias affected the sigma metric value. Where bias was problematic, more than half of the occurrences did not meet the minimally acceptable biologic variation bias goals and/or the $0.33CV_I$ goal for QCM. However, the inability to achieve the minimally acceptable BV bias goal was not unique to measurands with < 6 sigma. We found examples where BV bias goals were not achieved but sigma metrics were > 6. This was because the bias was much lower than the total allowable error goal (based on ASVCP or expert

opinion). When bias was a larger proportion of the TE_a, the sigma value was < 6; for these measurands the bias ranged from 3.22 – 8.27%. Our evaluation suggests that bias > 3% on the Sysmex XT-2000*i*V analyser indicates an increased likelihood that the sigma metric value will be below 'world class' performance and an increased likelihood to fail at least 1 or, in most cases, 2 of the quality goals (Desirable BV bias and 0.33 x CV_I goals) of the 3 biologic variation-based goals. For these occurrences the imprecision varied from 0.89 to 3.61 (CV). Additional performance evaluations are needed to confirm the suggestion from these results that observed bias results > 3.0% could be used to determine emerging or established problems and whether instrument servicing is indicated when these occur.

Imprecision accounted for a third of occurrences where sigma metrics did not achieve > 6 sigma and of those, 1 did not meet the minimally acceptable BV-CV goal. We found the CV_{obs} to be significantly higher for this occurrence, 7.21, in comparison to better performing sites (Analysers 1 & 2) where performance was better with CV_{obs} < 1 during that same evaluation. When the CV_{obs} was < 2 and bias was not excessive, the sigma metric was consistently > 5. We did note some higher CV values (> 2%) that achieved either optimal, desirable, or minimal CV goals and were still achieving between 5-6 sigma in comparison to the better performing sites so this alone is not a discriminatory performance indicator. Combined imprecision and bias accounted for 2 measurand < 6 sigma, BV was only available for one of these measurands which achieved only one of the three quality goals, and this was at a minimally acceptable level.

The 0.33 x CV_I goal was not achieved for 11 out of 18 occurrences with sigma metric < 6; 9 of these were during the initial phase, 1 occurrence was during the second phase and 1 during third phase. Ten occurrences were attributed to bias, and 1 was due to both imprecision and bias. The ongoing attention to QC, QCM handling, and emphasis on QCM performance, as well as instrument servicing when excessive bias and/or low sigma metrics (< 5) was identified likely contributed to the increasing ability to achieve the BV-based laboratory systems goal over the 3 periods of evaluation. Continued monitoring of performance is needed to determine if this improved level of performance will be sustained. The 0.33 x CV_I goal proved to be a good discriminatory goal for poor performance based on comparison with the QCM target.

From the initial evaluation, it was clear that analysers 4A & 4B and analyser 5 required attention from Sysmex technical support as nearly half of the measurands had sigma metrics < 6 sigma during this phase and for the first assessment for analyser 3 in Month 2 where half of their measurands were < 6 sigma. This was particularly seen in the erythroid measurands and following a service from Sysmex we saw improvements for those measurands in terms of sigma metrics which were now > 5 sigma and many > 6 sigma on subsequent evaluations. The performance variations and need for correction of bias was discussed with the service representative prior to their adjustments being done.

We standardised QCM handling amongst the sites, ensured routine maintenance procedures were adhered to and documented in the analyser

maintenance log and improved the technicians' understanding of QC evaluation for identifying shifts, trends and out of control results. Following these approaches, we observed improved sigma metrics across the three phases of QC data collection.

From this study, sigma metrics < 5, whether attributed to bias or imprecision, did not meet all quality goals that were set. Sigma metrics > 6 had better capabilities for reaching the BV goals and for continued improvement as these measurands had fewer QCM failures. Sigma metrics < 4 is also unacceptable as these measurands could only achieve 1 out of the 3 BV quality goals and from our study required technical attention.

This raises the question as to whether there is a hierarchy of quality goals that should be utilised to achieve harmonisation? We found the choice of BV goals used in this study to be sufficient for challenging the performance of the measurands. We also found that the TE or expert opinion goals were not discriminatory since all analysers were able to achieve these goals, even when observed performance and sigma varied widely. Collected bias and imprecision data were more usefully applied to check achievement of the BV goals rather than TE or expert opinion goals that were not major contributors to the improvement process. Biologic variation goals were superior in that once desirable biologic-variation-based CV goals were met we could be confident that patient result would not be misclassified clinically and imprecision would not mask physiological change, which is reassuring for ongoing patient monitoring. We could also be assured we could continue to use common reference intervals as long as we could meet the minimal acceptable BV bias goal (Fraser, 1992b; Fraser *et al.*, 1997).

A limitation to using biologic variation is that it varies with measurand and consequently how permissive or stringent the BV quality goals will vary according to measurand (Flatland *et al.*, 2018); however, a single numerical quality goal is not suitable across all measurands, which have different interpretations and clinical utility, as well as differences in analytical performance. Despite this limitation, our study demonstrated that we could achieve the optimal or desirable BV quality goals for most measurands and meet minimally acceptable for some measurands within the accumulation of data from the three evaluations. From our final evaluation (Month 5) after the recommended improvements from the first two data sets, we could meet these BV goals most of the time. Where we did not meet BV goals (25 occurrences across all evaluations < 6 sigma) the most serious concern is that we are at risk of masking within-patient change due to analytical variation. Ongoing evaluations are needed to determine if sigma metrics alone or in combination with absolute observed bias and BV-based goals (especially 0.33 x CV_I goal) can be reliably used on an ongoing basis for monitoring performance.

This study demonstrates that these quality specifications are achievable and that harmonisation is possible. However, fluctuations in the ability to achieve these goals and in the sigma metrics were seen throughout the three evaluations for some sites while others remained relatively consistent. This may be due to inherent analytical variability between the analysers even though these are of the same manufacturer and model. It may also be contributed to the capability and training of the operators, and environmental conditions at each site. Perhaps the degree of quality management

engagement and varying throughput of the analysers contributes. A publication from Sysmex (Development, 2012) also made reference to this by quoting differences even with analysers from the same manufacturer will still have some disparity and also indirectly highlighted the need to further scrutinise measurands that are calculated. MCHC is derived from HGB and HCT values (Development, 2012) and so its validation must be compared to their performance; for example, when MCHC was < 6 (Table 2.2, Analyzer 3) the corresponding Sigma metrics for HCT and HGB were 2.20 and 12.82 respectively indicating that the poor MCHC performance could likely be attributed to poor HCT performance.

From our experience to date, in conjunction with BV bias and 0.33 x CV_I goal we recommend using a sigma-based quality management system as a method to monitor and compare performance in a network of laboratories; however, this should be carried out on a continuous and routine basis to detect any changes in performance. When the application of sigma metrics to assess and compare instruments, particularly for chemistry and immunoassay has been conducted in human laboratories, they determined that sigma metric < 4 results in QC rule failures 20 times more frequently than six sigma measurands (Litten, 2017). Additionally for the better performing methods, the cost of supplies and reagents is significantly lower (Litten, 2017; Harrison and Jones, 2017). More specifically, a study of sigma metrics in human laboratories using the Sysmex XN and non- XN series (not our exact model) from a network of laboratories found sigma metrics consistently > 4 with the exception of HCT which had a number of occurrences of < 4 sigma (Harrison and Jones, 2017). On review of our

sigma metrics, HCT also had the most occurrences of < 4 sigma. This was difficult to compare to other studies as HCT is measured on the Sysmex analysers and calculated on most other haematology analysers.

2.5 Conclusion

Overall, the results from this study show that we were able to evaluate the performance of the Sysmex XT-2000*i*V across a network of laboratories and achieved harmonisation and optimisation of procedures. We determined that the total error budget is indeed contained within ASVCP guidelines or internal expert opinion goals. This harmonisation process demonstrated that sigma metrics are a useful monitor of performance across a network of laboratories capable of highlighting measurands that require attention (and specifically sigma metrics < 5 warrants investigation and attention (from Sysmex technical support). Keeping sigma metrics consistently above 5 ensures an efficient and effective quality management system. We intend to use sigma metrics < 5 as the threshold for action to be taken to improve performance for haematology in our laboratories. We intend to further evaluate the use of observed bias < 3% and failure to achieve the biologic-variation-based desirables - Bias and 0.33 CV_I goals as indicators of unacceptable performance.

We saw continuous improvements in sigma metrics and most quality specifications during the three performance evaluation phases and have demonstrated that they are a good exercise in risk assessment and quality improvement. Further investigation is needed to determine whether applying

more stringent QC rules based on the observed performance may be successful in improving performance.

CHAPTER 3: QUALITY CONTROL VALIDATION FOR A VETERINARY LABORATORY NETWORK OF SIX SYSMEX XT-2000*i*V HAEMATOLOGY ANALYSERS

This is the second stage of the research where a quality control validation approach was carried out to maintain the harmonisation status achieved in chapter 2 and support the transition to harmonisation to ensure suitable clinical interpretation across the laboratory network. This research was led by S Daly and in collaboration with Dr Peter Graham and Dr Kathy Freeman as co-authors on the manuscript produced from this research and represents original work.

Adaptations: This chapter has been changed from its previously published version to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

Reference: This chapter has been published as Daly SC, Freeman KP, Graham PA. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000*i*V hematology analyzers. *Vet Clin Pathol*. 2022; 51: 565-576. doi: 10.1111/vcp.13163

Received: 7 September 2021 | Revised: 15 April 2022 | Accepted: 10 May 2022

DOI: 10.1111/vcp.13163

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

ORIGINAL ARTICLE

Quality control validation for a veterinary laboratory network of six Sysmex XT-2000*iV* hematology analyzers

Susan Daly[1,2] | **Peter A. Graham**[2] | **Kathleen P. Freeman**[3]

[1]SYNLAB-VPG/Cork, South Cork Industrial Estate, Cork, Ireland

[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

[3]SYNLAB-VPG/Exeter, Unit 1b, Science Park, Exeter, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, South Cork Industrial Estate, Vicar's Road, Pouladuff, Cork T12 DNW4, Ireland.
Email: susan.daly@synlab.co.uk

Abstract

Background: Quality control (QC) validation is an important step in the laboratory harmonization process. This includes the application of statistical QC requirements, procedures, and control rules to identify and maintain ongoing stable analytical performance. This provides confidence in the production of patient results that are suitable for clinical interpretation across a network of veterinary laboratories.

Objectives: To determine that a higher probability of error detection (P_{ed}) and lower probability of false rejection (P_{fr}) using a simple control rule and one level of quality control material (QCM) could be achieved using observed analytical performance than by using the manufacturer's acceptable ranges for QCM on the Sysmex XT-2000*iV* hematology analyzers for veterinary use. We also determined whether Westgard Sigma Rules could be sufficient to monitor and maintain a sufficiently high level of analytical performance to support harmonization.

Methods: EZRules3 was used to investigate candidate QC rules and determine the P_{ed} and P_{fr} of manufacturer's acceptable limits and also analyzer-specific observed analytical performance for each of the six Sysmex analyzers within our laboratory system using the American Society of Veterinary Clinical Pathology (ASVCP)-recommended or internal expert opinion quality goals (expressed as total allowable error, TE_a) as the quality requirement. The internal expert quality goals were generated by consensus of the Quality, Education, Planning, and Implementation (QEPI) group comprised of five clinical pathologists and seven laboratory technicians and managers. Sigma metrics, which are a useful monitoring tool and can be used in conjunction with Westgard Sigma Rules, were also calculated.

Results: The QC validation using the manufacturer's acceptable limits for analyzer 1 showed only 3/10 measurands reached acceptable P_{ed} for veterinary laboratories (>0.85). For QC validation based on observed analyzer performance, the P_{ed} was >0.94 using a 1-2.5s QC rule for the majority of observations (57/60) across the group of analyzers at the recommended TE_a. We found little variation in P_{fr} between manufacturer acceptable limits and individual analyzer observed performance as this is a characteristic of the rule used, not the analyzer performance.

3.1 Introduction

Quality goals have been established for clinical laboratory procedures based on biologic variation data or expert opinion (Fraser, 1992a; Fraser, 1992b) (Nabity *et al.*, 2018; Harr *et al.*, 2013; Flatland *et al.*, 2018a) and analytical equipment that achieves these quality goals will generate results that are suitable for clinical decision making (Westgard, 2016a). Once analysers have been evaluated and optimised to meet quality goals, statistical quality control (QC) aims to maintain performance within those goals. Statistical QC relies on one or more control rules applied to the results generated from regular analyses of quality control materials (QCM). For statistical QC to be effective in achieving its purpose, the selected rule(s) must be sensitive and specific for the identification of deteriorating analytical performance. That is, statistical control validation is a necessary step in the design and implementation of statistical QC rule(s).

In a network of veterinary laboratory haematology analysers that have undergone a harmonisation process outlined in chapter 2, statistical QC performs an additional role of maintaining harmonisation and allowing for continued interchangeability of results and common reference intervals within the veterinary network. In the same way that QC validation supports the effectiveness of QC to ensure the stability of a single analyser, QC validation is a necessary step in the design and implementation of QC rules for a network of harmonised analysers.

The sensitivity and specificity of QC are reflected by the probability of error detection (P_{ed}) and probability of false rejection (P_{fr}), respectively. The P_{ed} is

a measure of the frequency with which a control rule would cause analytical runs to be rejected when results contain errors beyond the inherent imprecision. Ideally error detection should be set to 100% for medically significant errors; however error detection at ≥ 90% is considered sufficient (Westgard, 2016a). Conversely, P_{fr} is a measure of the frequency with which analytical runs are rejected when there is no apparent reason or issue. The goal is to have the highest possible P_{ed} to ensure that medically important errors are not missed and a low P_{fr} (≤ 5%) (Westgard, 2016a) to reduce the waste cost and efficiency impact on patient sample volume, reagent use and result turnaround time. This optimises the efficiency and capability of statistical QC as a tool for demonstration of stable laboratory system performance. The analytical performance capabilities of the analyser informs the choice of the number of control materials and statistical rules selected, which then determines the P_{ed} and P_{fr} achieved. This step by step process results in a quality control validation.

Laboratories using statistical QC often employ the Westgard rules which are denoted by a shorthand notation. For example, a 1-2s rule means that control limits encompass two standard deviations (SD) on each side of the observed mean. A 1-3s refers to a rule that is set at ±3 SDs and the rule is violated when one measurements exceeds ±3 SDs from the QCM mean. The numeral 2 placed in front of a rule such as 2-2s indicates the rule is violated when two consecutive control measurements or two single measurements across two QCMs are outside 2SD control limits (Westgard, 2016a).

Six Sigma process-improvement methodology has been applied to clinical laboratory analyses such that performance capability (bias and imprecision)

relative to TE_a can be represented on a numeric scale as a "sigma metric" (σ). Sigma metrics and QC rules have been combined in Westgard's Sigma Rules (Westgard, 2016a); such that sigma metrics can, in turn, determine whether a simple single QC rule or a more complicated collection of rules is required. A sigma metric ≥ 6 for an analytical method indicates < 3.4 errors/defects per million results[10], defined as world class performance, and the implied low bias and imprecision means these methods are easily controlled with simple QC rules and low numbers of QCM measurements. It would take a large deterioration in performance to cause a clinical error and such a large shift could be readily detected with a simple QC rule. In fact, for measurands with sigma metrics ≥ 6.0, acceptable P_{ed} and P_{fr} can be achieved using 1 or 2 control measurements. Conversely, measurands with lower sigma metrics will require a multirule and/or larger numbers of control measurements and may or may not be able to achieve acceptable P_{ed} and P_{fr} (Westgard and Barry, 2008). Commercially available software called EZRules3© allows the user to explore and select QC rules based on the total error quality goal and observed analytical performance, as well as the P_{ed} and P_{fr} that can be achieved (Westgard and Westgard, 2016).

It has been acknowledged that P_{ed} of ≥ 0.85 can be used (Rishniw and Pion, 2016) for point of care testing (POCT) for veterinary medicine or for haematology where use of a single QCM is preferred. The use of a single haematology control material is traditional in the UK but P_{ed} ≥ 0.90 cannot be achieved with a 1-3s rule and single QCM data point can only achieve P_{ed} ≥ 0.85 (Westgard, 2001; Lester *et al.*, 2013).

This paper describes a network of 6 harmonized Sysmex XT-2000*i*V haematology analysers across 5 locations which required a QC approach that could ensure the maintenance of individual analyser performance and additionally support continued harmonisation. In deciding which QC approaches to evaluate, the following preferences and constraints had to be taken into consideration. First, the use of a simple control rule (e.g., 1-2.5s, 1-3s), rather than a multirule, and a single level of QCM was preferred due to simplicity of training and evaluation and based on traditional laboratory practices and economics. Second, the choice of QCM was limited because a third party control material is not available for the Sysmex analyser, only a manufacturer-provided QCM. It was felt that the potential disadvantages of a single level of QCM could be mitigated since all blood smears undergo additional non-statistical quality control via microscopy by a fully trained technician to validate the automated results prior to pathologist's review of laboratory data and correlation with clinical findings before results are reported.

The authors set out to discover whether an effective QC approach supporting network harmonisation could be successfully implemented given these constraints and preferences.

This study addressed the following objectives:

1. Determine that a higher P_{ed} can be achieved using QC rules validated for each individual analyser than that achieved using the manufacturer's acceptable limits for the control limits.

2. Determine that a P_{ed} (≥ 0.85) and P_{fr} (≤ 0.05) can be achieved using a single QCM (n =1) with a simple single control rule (1-2.5s or 1-3s).
3. Assess the use of sigma metric evaluation as part of the QC approach in conjunction with the validated QC rules appropriate for each analyser, with a goal of using it to monitor when instrument servicing is needed in order to maintain a high level of instrument performance.
4. Assess whether analyser performance criteria established in a previous study (chapter 2) (Bias < 3%, achievement of desirable biologic variation-based goals for CV and bias, and 0.33CV_I goals, with sigma metrics > 5) would be confirmed in this study as useful contributions to an overall QC approach.

3.2 Material and Methods

3.2.1 Haematology analysers

The data from six Sysmex XT-2000*i*V analysers (Sysmex corporation, Kobe, Japan) located in five laboratory locations in the UK (n=4) and Ireland (n=1) were evaluated. Analysers were designated Analyser 1 , 2, 3, 4a, 4b and 5. This network of analysers had previously undergone optimisation and successful harmonisation (Daly *et al.*, 2021). Analyser 1 was designated the reference analyser.

3.2.2 Quality control material

The QCM used for these evaluations was Level 2 - Normal e-CHECK™ (XE)-Hematology Control (Sysmex Corporation) provided in an 8 vial kit with one new vial used per week. Typically the lot number changed for each new kit with slight changes in manufacturer's target means and acceptable limits. A single lot number was used for all analysers during the study. The haematologic measurands evaluated were white blood cell count (WBC), red blood cell count (RBC), haemoglobin concentration (HGB), haematocrit (HCT), mean cell volume (MCV), mean cell haemoglobin concentration (MCHC), reticulocyte count (RETIC), platelet count (PLT), plateletcrit (PCT) and red cell distribution width-coefficient of variation (RDW-CV). The QCM was analysed by a fully trained technician according to standard operating procedures; when not in use, the QCM vial was refrigerated.

3.2.3 Evaluation of manufacturer's acceptable limits as control rules using EZRules3©

The manufacturer's acceptable limits for the haematologic measurands were evaluated using data from 1 month of QCM results from Analyser 1. The width of the manufacturer's acceptable limits for a measurand was divided by the standard deviation (SD) for the reference lab (Analyser 1) to determine the number of SDs contained within this range and that number divided by 2

to determine the SDs on each side of the target mean. This was rounded to the nearest QC rule available within EZRules3© to evaluate the performance of the control rule most closely representative of the manufacturer's acceptable limit. EZRules3© allows manual selection for simple control rules for 1-2s, 1-2.5s, 1-3s, 1-3.5s, 1-4s, 1-5s and 1-6s. If the manufacturer's range control rule was equivalent to < +/- 2SD's, a 1-2s rule was used. If the manufacturer's range control rule was > 6SD's, a 1-6s rule was used.
The control rule was then assessed using EZRules3© to determine the P_{ed} and P_{fr} that are possible using the startup QC design which is a program that allows the user to follow a series of prompts, the manufacturer's target mean as the decision level concentration, and the ASVCP recommendations (Nabity *et al.*, 2018) or internal expert opinion for total allowable error (TE_a) as the chosen quality requirement and number of QCMs set to n=1. For those measurands for which no ASVCP recommendation for TE_a was available (PCT and RETIC number), expert opinion goals from an internal working group composed of pathologists and technicians were used (6 and 5 persons, respectively). The expected instability setting in the software was set to off.

3.2.4 QC validation – evaluation of control rules based on observed individual analyser performance

The QC validation for each of the 6 Sysmex instruments, customised for observed performance, was determined by using approximately 1 month of QC data (March 2020) from chapter 2. We used EZRules3© to plot the

observed imprecision and observed bias (calculated from the manufacturer's target mean) generated from daily repeated QC measurements for each analyser and measurand to determine the individual analyser's QC rules and associated P_{ed}, P_{fr} and sigma metrics.

The candidate rules evaluated using the manual selection option were 1-2.5s, n=1 and 1-3s, n =1. A Ped ≥ 0.90 can theoretically be achieved with a 1-2.5s rule, and single QCM, if performance is sufficiently good.

3.2.5 Review of criteria for measurands with poor performance

Findings from chapter 2 indicated that sigma metrics > 5, bias < 3% and measurands achieving desirable biologic variation goals for CV and Bias gave confidence in stable analytical performance. These criteria were reviewed and compared to the findings in the study to confirm that harmonisation was maintained during the course of this study.

3.3 Results

3.3.1 QC validation for manufacturer's acceptable limits

The results of QC validation for the manufacturer's acceptable limits for the QCM using the observed SD for Analyser 1 are summarised in Table 3.1. The acceptable criteria for P_{ed} (> 0.85) was met by only 3 measurands (HGB, PCT and WBC). Acceptable P_{fr} (≤ 0.05) was achieved for all 10 measurands

using the available control rules that most closely represented the manufacturer's acceptance limits.

Table 3.1: Summary of QC validation for Sysmex quality control Level 2, Manufacturer's acceptable quality control material (QCM) ranges using one month of quality control data from the system reference laboratory Sysmex XT 2000iV haematology analyser (analyser 1).

Measurand (units)	**RBC ($x10^6$/uL)**	**HBG (g/dl)**	**HCT (%)**	**MCV (fL)**	**MCHC (g/dL)**	**RDW-CV (%)**	**RETIC ($x10^6$/uL)**	**WBC ($x10^3$/uL)**	**PLT ($x10^3$/uL)**	**PCT (%)**
Manufacturers acceptable lower limit	4.05	11.3	32.7	76.7	31.7	13.8	63.8	6.81	188	15
Manufacturer's acceptable upper limit	4.47	11.9	36.1	84.7	35.7	18.6	132.6	7.69	254	27
Width of acceptable range	0.42	0.6	3.4	8.0	4.0	4.8	68.8	0.88	66	12
Observed standard deviation (SD) for Analyser 1	0.03	0.15	0.35	0.44	0.32	0.09	3.71	0.16	6.50	0.79
Number of observed SD within acceptable range	14.00	4.00	9.70	18.18	12.50	53.33	18.54	5.50	10.15	1.27
Number of observed SD on each side of the target mean	7.00	2.00	4.85	9.09	6.25	26.66	9.27	2.75	5.07	0.63
Closest available control rule	1-6s	1-2s	1-5s	1-6s	1-6s	1-6s	1-6s	1-3s	1-5s	1-2s
Total allowable error (TEa)% **(ASVCP except for PCT and RETIC #)**	10	10	10	7	10	10	**40**	15	20	**25**
(CV_{obs}) % (Analyser 1)	0.66	1.32	1.00	0.54	0.96	0.54	5.81	2.12	2.87	3.65
Observed Bias % (Analyser 1)	0.008	0.44	0.66	0.75	1.05	0.19	4.00	1.16	2.33	2.89

Probability of error detection (P_{ed})	>0.01	>0.98	>0.13	>0.01	>0.01	>0.01	>0.01	>0.85	>0.13	>0.98
Probability of false rejection (P_{fr})	0.0	0.05	0	0	0	0	0	0	0	0.05
Sigma metric	15.14	7.24	9.34	11.57	9.32	18.17	6.20	6.53	6.16	6.06

RBC = Red Blood Cell count; HBG = Haemoglobin; HCT = Haematocrit; MCV = Mean Corpuscular Volume; MCHC = Mean Corpuscular Haemoglobin Concentration; RDW-CV = coefficient of variation of the Red Cell Distribution Width; RETIC = reticulocyte count; WBC = White Blood Cell count; PLT = Platelet count; PCT = Plateletcrit, CV_{obs} = Observed Coefficient of Variation.

Number of observed standard deviations in the manufacturer's acceptable range was divided by 2 to determine the number of SD on each side of the manufacturer's target mean. This was used to determine the closest control rule available for selection in the EZRules3 program for evaluation. If the number of SD within the acceptable range/2 is much less than the closest available control rule, the width of the acceptable range will be reduced compared to the manufacturer' acceptable range, resulting overestimation of the P_{ed} (PCT). If the SD within the acceptable range/2 is much greater than the closest available control rule, the width of the acceptable range will be greater than the manufacturer's acceptable range, resulting in underestimation of the P_{ed} (RDW-CV).

3.3.2 QC validation based on observed analyser performance for 6 Sysmex analysers

The results of the QC validation for 2 control rules (1-2.5s and 1-3s) based on observed analyser performance (March 2020) for the selected measurands on 6 Sysmex analysers (supplementary tables is available in Appendix E, Table E.1, E.2 and E.3). Only 3 measurands did not achieve $P_{ed} \geq 0.85$: RBC and PLT on analyser 4a using a 1-3s rule (n=1), PCT for both QC rules, 1-2.5s and 1-3s (n=1). All other measurands across the 6 analysers achieved $P_{ed} \geq 0.85$ with either control rule.

$P_{fr} \leq 0.05$ was achieved for all 10 measurands for each of the 6 Sysmex analysers and both candidate control rules.

Sigma metrics were 6 for 56/60 observations. Where sigma metrics were < 5.5, 3 of these measurands did not achieve P_{ed} 0.85 (as above for analyser 4a), the other observation was for RETIC ($\sigma = 5.98$) for analyser 4b which did achieve acceptable P_{ed} (0.85).

3.3.3 Evaluation of QC rules

The 1-2.5s rule offered the highest P_{ed} for all measurands for the group of Sysmex analysers based on observed analyser performance and yielded P_{fr} of only 1% (supplementary tables is available in Appendix E, Table E.1, E.2 and E.3). Manufacturer's QC limits achieved acceptable P_{ed} for only 3 measurands (see above, Table 3.1).

3.3.4 Sigma metric monitoring

A monthly review of sigma metrics showed 3 measurands that performed < 5.5 sigma failed to achieve quality goals for P_{ed} for one (RBC & PLT) or both QC rules (PCT). On further investigation, use of the quality goal index (QGI) (Parry DM, 2009) demonstrated that observed imprecision was implicated for 2 measurands while both imprecision and bias were implicated for one measurand. The same analyser (4a) accounted for all 3 measurands with poor performance in comparison to the other analysers where performance was > 5.5 sigma. (supplementary material is available in Appendix E, Table E3.1, E3.2 and E3.3).

3.3.5 Review and optimisation of criteria for measurands with poor performance

Criteria established in chapter 2 for stable analytical performance indicated that measurands with bias < 3%, achievement of desirable biologic variation-based goals for CV and bias, and 0.33CVi goals, with sigma metrics > 5 were stable performers supporting harmonisation. In this study, we noted sigma metric > 5.5 rather than 5 demonstrated analytical stability. P_{ed} was achieved when observed CV and Bias achieved desirable biologic variation goals as reported in chapter 2.

Imprecision limits in order to achieve acceptable P_{ed} (> 0.85) for RBC (see supplementary table available in Appendix E, Table E3.1), analyser 4a using a 1-2.5s rule, if desirable bias based on biologic variation applied (1.761%) then

the observed CV % must be ≤ 1.45 in order to achieve P_{ed} > 0.94. Similarly, for PLT for analyser 4a (see supplementary table available in Appendix E, Table E3.2). Sigma metric was 5.13 and for the 1-2.5s rule we could determine that if desirable bias was achieved based on biologic variation (5.17%) then the observed CV % must be ≤ 2.60 % (3.72%) therefore the 1-3s rule was not satisfactory; the 1-2.5s rule was satisfactory at 0.86 P_{ed} and the other analysers achieved > 0.94 P_{ed}.

3.4 Discussion

To the authors' knowledge this is the first reported QC validation for a network of Sysmex analysers utilised for veterinary haematology and for ongoing harmonisation within a laboratory system. The quality goal used to determine the P_{ed} and P_{fr} was total allowable error (TE_a) provided by the ASVCP (Nabity *et al.*, 2018) or from internal expert opinion (RETIC and PCT). The QC validation using manufacturer's acceptable limits and observed analytical performance for Analyser 1, the reference analyser for this group showed only 3/10 measurands (HGB, WBC, PCT) could achieve an acceptable P_{ed} with recommended TE_a or expert opinion in the case of PCT (Table 3.1). Consequently, manufacturer's acceptable limits could not be relied upon to identify clinically important equipment malfunction for most measurands. The wide manufacturer's limits were not unexpected as this has been highlighted previously (Farr and Freeman, 2008; Manzocchi *et al.*, 2017) and these limits are generally derived from groups of instruments rather than individual instruments. Manufacturer's limits that failed to achieve acceptable P_{ed}, error

detection was as low as > 0.01 using the closest available QC rule (Westgard, 2016a; Westgard and Westgard, 2016; Manzocchi *et al.*, 2017) of 1-5s or 1-6s compared to the QC validation using the observed performance of individual analysers for all 10 measurands achieving P_{ed} > 0.94 using a 1-2.5s QC rule and > 0.85 using a 1-3s QC rule for observed performance for analyser 1, resulting in much higher error detection.

We determined that a higher P_{ed} could be achieved using QC rules validated for each individual analyser rather than the use of the manufacturer's acceptable limits. We determined that we could achieve an acceptable P_{ed}, > 0.94 for 95% of observations and P_{fr} ≤ 0.05 for all observations for our network of Sysmex analysers. We failed to meet the P_{ed} criteria (> 0.85, n=1) for 3/60 (5%) of observations (Analyser 4a for RBC, PLT & PCT) but were able to achieve the required P_{ed} > 0.85 using either 1-2.5s and/or 1-3s QC rule, n=1, for all other measurands. It is noteworthy that the performance of Analyser 4a improved following instrument service and was able to achieve the quality goals. This is in comparison to the manufacturers' P_{ed}, where only 2 measurands achieved > 0.94 (Table 3.1). We are satisfied that QC validation from observed analytical performance offers the most consistent probability of error detection ≥ 0.85 for the largest number of measurands and, therefore, greater confidence in quality control and likelihood of network stability. The use of the manufacturer's acceptable limits does not provide sufficient P_{ed} for most measurands to have peace of mind regarding detection of unstable instrument performance. The probability of false rejection was within P_{fr} criteria (≤ 0.05) for analyser 1 using both manufacturer's limits and observed analytical performance. Overall, the individual analyser validation

provided us with the ability to achieve error detection ≥ 0.85 and low levels of false rejection resulting in more satisfactory QC using the 1-2.5s rule.

The 1-2.5s rule proved to be the optimal solution for one level of control material, since the P_{ed} was > 0.94 for the majority of observations whereas a P_{ed} of 0.88 was the best achieved using the 1-3s rule at the recommended TE_a. There with little variation in P_{fr} between both rules assessed for the individual analyser validation as P_{fr} is a function of the rule used and number of QCM and is not based on analyser performance.

For the measurand (PCT) that failed to meet acceptable criteria for P_{ed} using both QC rules (1-2.5s & 1-3s), it was considered whether the TE_a based on expert opinion was too stringent, however as the remaining analysers performed optimally this was more likely to be declining performance. And, on closer evaluation, the large CV (5.31%) and sigma metric < 6 ($\sigma = 4.56$) suggest that for this measurand on this particular instrument, performance was deteriorating and required attention and this is reflected in both QC rule failures.

Sigma metrics were applied as an additional performance monitoring tool and in conjunction with QC validation and the use of sigma metrics as a performance indicator is well documented (Harrison and Jones, 2017; Westgard and Westgard, 2015; Litten, 2017). Sigma metrics were > 6 for all measurands for Analyser 1 when based on observed analytical performance. For measurands on other analysers that performed < 6 sigma (RBC &PLT for analyser 4a), the bias and/or imprecision was distinct from the other measurands across the instruments that performed at > 6 sigma.

On closer evaluation for these measurands, we could determine limits for observed imprecision when applying a 1-2.5s rule in order to achieve acceptable P_{ed}. In review of the literature we could not conclusively draw relevant comparisons due to limited research in veterinary publications regarding imprecision goals. However, we think the following information is of interest to others performing QC validation using haematology analysers; biologic variation data is not available for PCT; if bias was set to 0%, then an observed CV ≤ 4.30% is required to achieve acceptable P_{ed} and > 6 sigma; the observed CV for analyser 4a was 5.31%. RETIC for analyser 4b had a sigma metric of 5.98 but *was* able to achieve the acceptable P_{ed} and P_{fr} for 1-2.5s and 1-3s rules; however, unlike the other measurands, the observed CV was not distinct from the other analysers. Biologic variation data is not available for this measurand but from our data, if bias was 0%, an observed CV ≤ 7.0% is required to meet quality goal (40%).

These observations demonstrate that, depending on the error budget, a low % CV with a high bias may still achieve the quality goal and vice versa but must at least meet desirable goals based on biologic variation. The use of sigma metrics in conjunction with the QC rules (Westgard Sigma rules) highlighted analytical instability for analyser 4a compared to the other 5 analysers and that technical attention was required for those measurands performing < 5.5.

The measurands that were easily controlled were > 5.5 sigma, had results within desirable bias and desirable CV based on biologic variation and could achieve high P_{ed} and low P_{fr} at the specified TE_a. By meeting these criteria, these measurands were controlled using a simple QC rule. This is more cost effective (Harrison and Jones, 2017; Litten, 2017) and less labour intensive

for the technician and offers a degree of confidence to the clinical pathologist interpreting the results in a multi-site, multi-analyser environment. For these measurands a 1-2.5s rule was considered the best candidate rule for ongoing use. This validation study is a much needed step in a harmonisation process that ensures that our network of analysers are comparable. We know that analytical variability exists between instruments of the same model (Development, 2012) (Daly *et al.*, 2021) therefore, by implementing the validated QC rule and monitoring analytical performance using Westgard's Sigma Rules (Westgard, 2016a) we are controlling this aspect of variability. The benefits gained from harmonisation through maximally efficient high P_{ed} and low P_{fr} control approaches include a reduction of network costs, uniformity of standard operating procedures, unified quality management policy, unified training and proficiency, less re-run waste and increased uniformity of turnaround times across the laboratories.

Measurands more difficult to control were < 5.5 sigma, had lower error detection (< 0.85 P_{ed}), a high observed CV (> 3.5%) in comparison to better performing measurands and in most cases failed to meet desirable CV based on biologic variation (supplementary material available in Appendix E, Tables E3.1, E3.2 and E3.3). In review of the criteria for identifying a need for analytical technical attention, chapter 2 suggested that a bias > 3%, failure to meet desirable biologic variation goals and sigma metric < 5 were immediate triggers for poor performance requiring servicing. However, this study suggested some modification to those criteria, because some measurements showed suboptimal performance (PLT, RBC & PCT – Analyser 4A) when sigma metrics were < 5.5 rather than < 5. For those example measurands for which there is biologic variation data available (RBC, HGB, HCT, WBC &

PLT), those that failed to meet desirable biologic variation goals and have a sigma metric < 5.5 warrant further investigation as they were poorly controlled with a 1-2.5s and/or 1-3s rule.

A more complicated multirule could be adopted in these cases but in the interest of keeping QC simple and not too laborious for the technicians and network quality Managers, our recommended approach is to monitor the sigma metric, observed bias % and observed CV %. We are confident that using 1) the 1-2.5s QC rule, 2) ensuring desirable biologic variation-based quality requirements for bias and CV are met and 3) maintaining sigma metrics > 5.5 sigma are excellent indicators of analytical stability sufficient to maintain harmonisation. In addition, we continue to use non-statistical measures such as haematologist's blood film microscopic evaluation, clinical pathologist's consideration of clinical findings, patient history and comparison/correlation of previous results. These measures are particularly important and recommended for measurands for which statistical QC is not performed or does not provide high P_{ed}. When statistical QC may not be sufficient to provide the necessary quality monitoring, the addition of non-statistical QC is recommended to help ensure that accurate and reliable results are reported (Westgard, 2016a).

Some academic work has been published on QC validation, but this has predominantly been for human medicine (Jassam *et al.*, 2011; Westgard, 2016d). In veterinary medicine, biochemistry has been a more common focus (Farr and Freeman, 2008; Manzocchi *et al.*, 2017). Early work (Freeman and Gruenwaldt, 1999) as well as some comparative work (Flatland *et al.*, 2014b; Cook *et al.*, 2016; Freeman and Flatland, 2017) has created a good foundation for our harmonisation study and emphasises QC validation as a

requirement for veterinary laboratories which is further supported from this study.

3.5 Conclusion

We recommend QC validation based on individual analyser performance in order to achieve high P_{ed} and low P_{fr} rather than the use of the manufacturer's acceptable QC limits for haematology in veterinary laboratories. We were able to validate 57/60 observations using a 1-2.5s and/or 1-3s QC rule; however, for optimal P_{ed}, we will apply a 1-2.5s QC rule. This simple rule used with a single QCM is adequate for individual analysers and for a harmonised network of analysers if sigma metrics remain > 5.5 and measurands meet desirable biologic variation goals for CV and bias. These standards will be applied across the group of analysers and will aid the internal quality team, pathologists and technicians to recognise when analytical issues arise.

CHAPTER 4: REPEAT PATIENT TESTING-QUALITY CONTROL SHOWS PROMISE AS AN ALTERNATIVE TO COMMERCIAL QUALITY CONTROL MATERIAL FOR A NETWORK OF 4 SYSMEX XT2000*iV* HAEMATOLOGY ANALYSERS

This chapter is a first report of validating RPT-QC for a network of haematology analysers as an alternative to commercial quality control material. This research was led by S Daly and in collaboration with Dr Peter Graham and Dr Kathy Freeman as co-authors on the manuscript produced from this research and represents original work.

Adaptations: This chapter has been changed from its previously published version to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

Reference: This chapter has been published as Daly SC, Graham PA, Freeman KP. Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000*i*V hematology analyzers. *Vet Clin Pathol*. 2023; 00: 1-9. doi:10.1111/vcp.13211

Received: 21 May 2022 | Revised: 8 September 2022 | Accepted: 22 October 2022

DOI: 10.1111/vcp.13211

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

ORIGINAL ARTICLE

Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000*i*V hematology analyzers

Susan Daly[1,2] | Peter A. Graham[2] | Kathleen P. Freeman[3]

[1]SYNLAB-VPG/Cork, South Cork Industrial Estate, Cork, Ireland

[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

[3]SYNLAB-VPG/Exeter, Exeter, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, South Cork Industrial Estate, Vicar's Road, Pouladuff, Cork T12 DNW4, Ireland.
Email: susan.daly@thevpg.ie

Abstract

Background: Repeat patient testing-quality control (RPT-QC) uses retained patient samples as an alternative to commercial quality control material (QCM). We elected to calculate and validate RPT-QC limits for red blood cell count (RBC), hemoglobin (HBG), hematocrit (HCT), and white blood cell count (WBC).

Objectives: (1) To validate RPT-QC across a network of four harmonized Sysmex XT-2000*i*V hematology analyzers and determine the total error that can be controlled with RPT-QC. (2) To generate quality control (QC) limits using the standard deviation (SD) of the duplicate measurement differences and determine a suitable simple QC rule with a probability of error detection >0.85 and probability of false rejection <0.05. (3) Monitor RPT-QC using sigma metrics as a performance indicator and (4) to challenge RPT-QC to ensure acceptable sensitivity.

Methods: Fresh adult canine EDTA samples with results within reference intervals were selected and run again on days 2, 3, and 4. QC limits were generated from the SD of the duplicate measurement differences. The QC limits were challenged using interventions designed to promote unstable system performance. The total error detectable by RPT-QC was determined using EZRULES 3 software.

Results: In all, 20-40 data points were needed for RPT-QC calculations and validated using 20 additional data points. The calculated limits differed among the network of analyzers. The total error that could be controlled was the same or better than that of the manufacturer's commercially available quality control material using the same analyzer for all measurands except hematocrit, which required a higher total error goal than that proposed by ASVCP guidelines to achieve an acceptable probability of error detection. The challenges designed to mimic unstable system performance were successfully detected as out-of-control QC.

Conclusions: The challenges for RPT-QC resulted in acceptable detection of potential unstable system performance. This initial study demonstrates that RPT-QC limits differ among the network of Sysmex XT-2000*i*V analyzers, indicating a requirement

4.1 Introduction

Quality control procedures for haematology analysers are necessary to ensure that the quality of operations is maintained and that reliable results are produced (Flatland and Weiser, 2014). The traditional method of ensuring quality controlled haematological data uses one or more commercial quality control material (QCM). Commercial QCM used in veterinary medicine is typically derived from human blood and synthetic/artificial components; therefore, may not comprise a commutable matrix for veterinary specimens. A preferable matrix would be from the animal origin, which includes components comparable to the patient throughput (Unsal *et al.*, 2010). There is not an available supply of veterinary-specific commercial QCM for haematology; therefore, veterinary laboratories have little choice but to use a QCM with a potentially suboptimal matrix. Commercial QCM is a costly but necessary expense and has a relatively short shelf life (Flatland and Vap, 2012). Most importantly, it does not provide appropriate assurances as to how the veterinary specimens may behave on the analyser and, therefore, is substandard in its application to a veterinary laboratory.

Repeat patient testing-quality control (RPT-QC) offers an alternative to commercial QCM by using retained patient samples to determine analytical stability. It relies on the principle that under specific storage conditions and time intervals, a labile EDTA specimen will deteriorate in a predictable way

from the baseline measurement to the later repeated measurement of the same specimen. A repeated measurement exceeding thresholds of expected deterioration suggests instability within the analytical system. The concept of RPT-QC has been well documented in human medicine; however, its application has varied (Dutra, 1966; Carstairs *et al.*, 1977; Orser, 1983; Cembrowski *et al.*, 1988). Recommendations for an RPT-QC protocol have been proposed, (Westgard, 2016b) in which a pilot set of data is gathered over a minimum of 20 days to generate control limits. These limits can then be used to determine whether RPT-QC results are “in-control” or “out-of-control.” After this initial limit-derivation phase, a fresh RPT sample would be measured on Day 1 and again within a defined time interval and defined storage conditions. It has been recommended to use patient samples yielding within reference interval results as they represent a significant proportion of caseload and may avoid extreme variabilities between measurements (Westgard, 2016b). The difference between original and repeated analyses can then be calculated and plotted similarly to other QC procedures (Cembrowski *et al.*, 1988). There have been some positive reports in the literature (Vanyo *et al.*, 2018; Flatland and Freeman, 2018b; Flatland and Freeman, 2020) that indicate a good potential for application in veterinary laboratories.

This study addresses the following hypotheses.

1. That RPT-QC limits could be generated and validated across a network of 4 Sysmex-XT2000*i*V haematology analysers.

2. That RPT-QC performance can achieve the ASVCP-recommended total allowable error (TE_a) goals using a simple QC rule ie, 1-2.5s or 1-3s, a probability of error detection (P_{ed}) of ≥ 0.85, and a probability of false rejection (P_{fr}) of ≤ 0.05.
3. That Sigma metrics is a satisfactory monitoring tool for RPT-QC, and sigma metrics > 6 (reflecting world class performance) can be achieved.
4. That deliberately created QC 'challenges' mimic unstable performance that is detectable by RPT-QC.

4.2 Materials and Methods

4.2.1 Analysers and measurands

RPT-QC data from 4 Sysmex XT-2000*i*V analysers (Sysmex Corporation, Kobe, Japan) were evaluated for the following measurands: red blood cell count (RBC), haemoglobin (HGB), haematocrit (HCT), and white blood cell count (WBC). These analysers were previously optimised and harmonised in chapter 2 (Daly *et al.*, 2021) to ensure stable performance prior to this study. Analysers were designated Analyser 1 (located in the first author's laboratory), 3, 4a, and 4b. The laboratory with analyser 2 was unable to participate and is not included in this report. A single level of commercial QCM (level 2 - Normal e-CHECK™ (XE)-Haematology Control), which is composed of stabilised human RBCs, human WBCs, platelets, and nucleated

RBC components in a preservative medium, was analysed once per day at analyser start-up, prior to analysis of patient samples, by a fully trained technician according to a standard operating procedure and was within accepted limits before assessing patient samples. These measurands were chosen as those most useful in our network of laboratories for statistical QC (Daly *et al.*, 2022). Differential cell counts and platelet counts were considered to be more adequately assessed by nonstatistical QC by technician evaluations of blood smears (Daly *et al.*, 2022).

4.2.2 RPT-QC limits generation

The process of RPT-QC data collection and calculations were conducted as recommended by Westgard (Westgard, 2016b). Canine EDTA samples that were received by the laboratory for routine testing were selected by a trained haematology technician for RPT-QC based on time of sampling (fresh/same-day submissions), surplus volume, and results that were within or close to reference intervals used by the laboratory. The day of accession was denoted “Day1”. The selected sample was refrigerated at 4^{o} to 8^{o}C and re-analyzed on days 2, 3, and 4 following acclimation to room temperature, thorough mixing, and rechecking for clots at each time point prior to analysis. The Day1 minus Day2 (or Day1−Day2) change was used to define control intervals that could be used for QC on successive days during the working week (Monday through Friday or Monday through Saturday). The Day1 minus Day3 (or Day1−Day3) change was used to define control intervals that could be used for QC on Mondays for those laboratories that worked Monday

- Saturday. The Day1 minus Day4 (or Day1−Day4) control intervals were for QC on Mondays for those laboratories that worked Monday - Friday. This process was carried out across the network of 4 Sysmex analysers, and the data was exported to Microsoft Excel (Microsoft Excel for Mac 2011, Version 14.7.7 (170905) Last update 14.7.7). Each analyser accumulated 50-60 RPT-QC samples comprised of a dataset of up to 40 samples for Day1−Day2, Day1−Day3, and Day1−Day 4 intervals to generate control limits.

A scatter plot of the differences of the repeated measurements was produced to identify outliers by visual examination. A maximum of 3 outliers were excluded from each control interval dataset following a visual review of a scatter plot of repeated measurements with except for one dataset, Day1−Day4 HGB, from which 4 were identified and removed. More complicated statistical evaluation for outliers was not undertaken to keep the procedure simple. If outliers were not readily apparent by visual examination, all data points were retained and included in the calculation of the RPT-QC intervals. The mean of the differences for the repeated measurements for each time interval and the standard deviation of the duplicate (SD_{dup}) measurements were calculated for each time interval (Day−Day2, Day1−Day3, Day1−Day4). The coefficient of variation (CV) and the difference squared (Diff Sq) were calculated from the mean of the repeated measurements, as described previously (Flatland and Freeman, 2018b) (Vanyo *et al.*, 2018; Westgard, 2016b). Supplementary material is available in Appendix F, Figure F4.1 for example of initial calculations.

The control limit range was calculated using the SD_{dup} and 1-2.5s and 1-3s control rules.

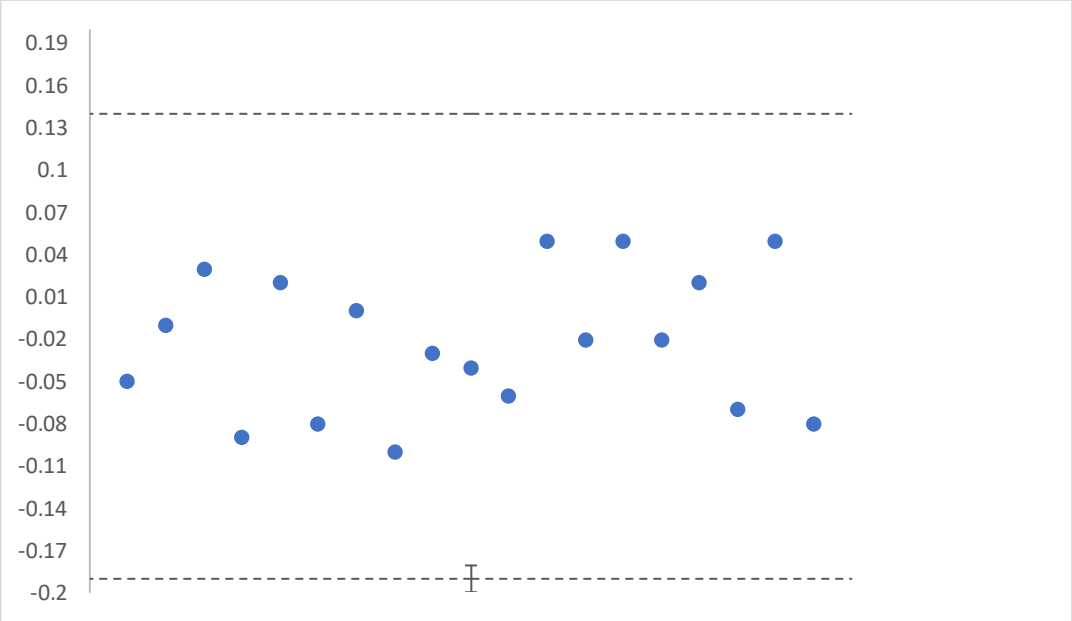
The following formula was used to calculate the SD_{dup};

$$\mathrm{SDdup} = \sqrt{\frac{\sum difference^2}{2n}}$$

Where n is the number of observations (Westgard, 2016c).

The control rule chosen for use was based on evaluation of the range of QC data compared to the width of the control limit interval (Figure 4.1) and whether or not validation was achieved by inclusion of at least 17 or more additional data points within the chosen control limits.

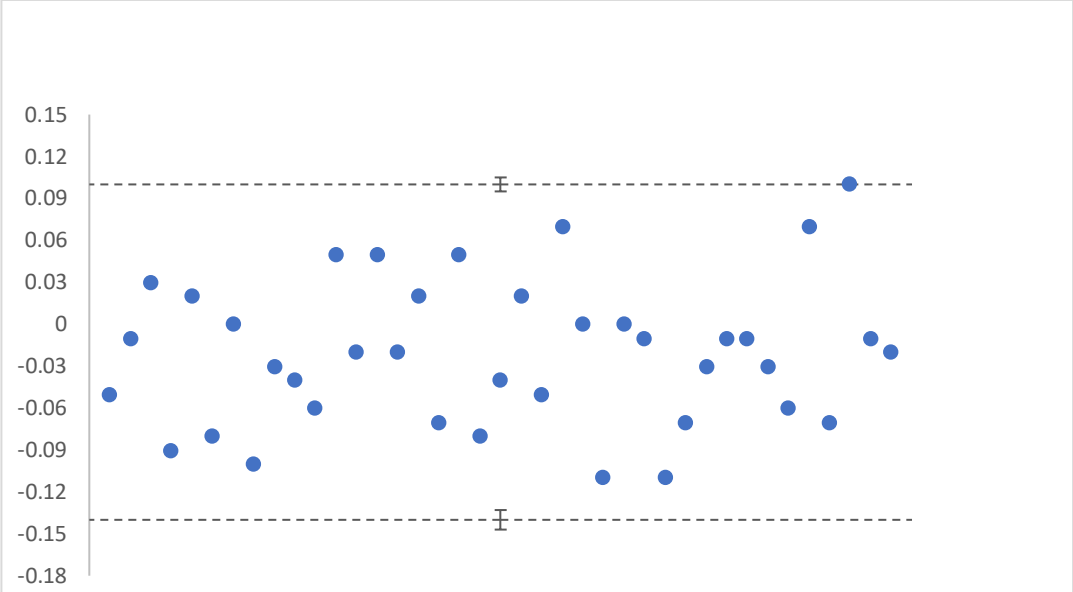
Figure 4.1: Initial scatterplot for RBC using 19 data points (one outlier removed) from the difference of Day1-Day2 samples used to calculate control limits for 1-3s rule.



The dashed lines represent the upper and lower control limits (0.14 and -0.19, width 0.33). Graph shows the data does not adequately fill the control limits which results in poor sensitivity in detection of error.

If the QC data was not all included within the calculated limits or the data did not appear to adequately fill the calculated control limit interval, then an additional 10-20 QC data points were combined with the first data set and control limit intervals were recalculated using the 30-40 total QC data points, as described above (Figure 4.2).

Figure 4.2: Subsequent scatterplot for RBC using 39 data points (one outlier removed) from the difference of day1-day2 samples used to calculate control limits for 1-3s rule.



The dashed lines represent the upper and lower control limits (0.1 and -0.14, width 0.24). Graph shows the data adequately fills the control limits, resulting in greater sensitivity in detecting out of control events with the smaller width control limits.

4.2.3 QC validation for calculated RPT-QC control limit interval

QC validation for the calculated RPT-QC control limits intervals for each measurand, time period and laboratory analyser were undertaken using the following specifications (ASVCP recommendations for TE_a for haematology for allowable total error, number of QCM = n =1, $P_{ed} \geq 0.85$, $P_{fr} \leq 0.05$, and simple rules used [1-2.5s or 1-3s]). The choice of $P_{ed} \geq 0.85$ was made because this is the minimum P_{ed} that can be achieved with a simple 1-3s rule, and a single level control result based on the models provided in the

EZRules3 program. EZRules3 was then used to determine P_{ed} and P_{fr} that could be achieved using the recommended TE_a (ASVCP) for each of the measurands. EZRules3 was also used to determine the lowest TE_a that could be achieved using the control rules chosen based on a simple QC rule and CV of the mean difference for the measurand while achieving a $P_{ed} > 0.85$ and $P_{fr} < 0.05$.

4.2.4 RPT-QC and sigma metrics

The sigma metric for each measurand and RPT-QC limit were calculated according to the following formula, using the ASVCP recommendations for allowable total error, the CV from each RPT-QC interval, and absolute bias = 0%:

$$\text{Sigma} = (TE_a\% - \text{absolute bias}\%)/CV\% = TE_a\%/CV\%$$

The SD, CV, mean difference, and control limit ranges for 1-3s, and 1-2.5s rules were calculated for each measurand and data set. The CV was calculated using the mean of the averages of the Day1 −Day2, Day1 −Day3 or Day1−Day4 (see supplementary material available in Appendix F, Figure F4.1) results as follows:

$$CV = (SD_{dup} / \text{mean of averages of the Day1 and repeated results}) \times 100$$

4.2.5 RPT challenges

RPT-QC challenges were performed at a single laboratory (analyser 1) over one week and were run as a single evaluation as per procedure for analysing RPT-QC. Challenges were set up to create intentional sample alterations to test the control limits for RPT-QC. The sample alterations included using an aged sample past days 1-4, freezing a sample, and a low volume sample ('short collection' not coming up to the fill line) with excess EDTA, resulting in an elevated EDTA:blood ratio sample. Samples were also manipulated by spiking the sample with saline creating a haemodiluted sample. Haemoconcentrated samples were created by removing plasma. The addition of native canine lipids obtained from a lipaemic serum sample to an EDTA blood sample created a lipaemic sample.

4.3 Results

4.3.1 RPT-QC limit generation

Table 4.1-4.4 summarise the information pertinent for the generation of the RPT-QC intervals for each measurand, time period, and analyser, including upper and lower control limits that were generated from the 4 Sysmex analysers for WBC, RBC, HGB, and HCT for either 20, 30 or 40 QC data for Day1−Day2, Day1−Day3, and Day1−Day4. Most datasets had at least one

outlier; only 4 datasets were identified with no outliers. Forty data points were required for the generation of RPT-QC intervals for most measurands. Only 10 of the 48 data sets across the measurands validated the initial calculations with 20 data points for either the 1-2.5s or 1-3s QC rules. Calculated RPT-QC intervals that were obtained from calculations performed with 40 data points were subsequently validated with a further 20 data points. Figure 4.1 (see above) illustrates the use of a scatterplot from the differences obtained for Day1−Day2 for RBC using 19 data points (analyser 1). One outlier was removed. Figure 4.2 (see above) illustrates a scatterplot for 39 datapoints as one outlier was removed for RBC Day1−Day2 with RPT-QC limits indicated by dashed lines. The scatter plots show that the 40 data points more completely fill the range specified for the RPT-QC limits.

4.3.2 QC Validation for RPT-QC intervals

ASVCP recommended TE_a quality goals of 10% total error were achievable across all measurands and time periods across all analysers for RBC and HGB using the specified criteria of a simple rule, single quality control assessment, and P_{ed} and P_{fr} described above. Analyser 4b could not achieve the ASVCP recommended quality goal for WBC for Day1−Day2 & Day1−Day3, although the other 3 analysers could. HCT could not achieve ASVCP recommended TE_a across all analysers and time periods.

4.3.3 RPT-QC and sigma metrics

Sigma metrics were > 6 for each analyser for all measurands except for HCT, which was < 3 sigma across all analysers and measurands using the ASVCP recommended total allowable error quality goal of +/- 10%.

Tables 4.1-4.4 summarise the QC validation, total allowable error that could be achieved and sigma metric for each measurand.

Table 4.1: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for White Blood Cell Count (WBC).

Analyser	**1**			**3**			**4a**			**4b**		
Day (time interval)	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4
SD	0.159	0.169	0.227	0.21	0.175	0.234	0.15	0.15	0.19	0.278	0.238	0.220
CV %	1.29	1.51	1.85	2.04	1.76	2.47	1.69	1.75	2.12	2.88	2.73	2.35
Mean difference	-0.111	-0.040	-0.556	-0.06	0.098	0.022	0.009	-0.071	-0.018	0.155	-0.214	0.094
Control limits range	0.796	0.843	1.42	1.04	1.10	1.17	0.906	0.928	1.123	1.65	1.42	1.32
ASVCP TE_a(%)	15	15	15	15	15	15	15	15	15	15	15	15
Lowest achievable TE_a (%)	8	10	12	11	10	14	11	11	13	18	17	15
Sigma metric based on ASVCP recommended TE_a	11.63	9.93	8.11	7.35	8.52	6.07	8.88	8.57	7.11	6.25	6.23	6.47
QC rule	1-2.5s	1-2.5s	1-3s	1-2.5s	1-3s	1-3s	1-3s	1-3s	1-3s	1-3s	1-2.5s	1-3s

CV = coefficient of variation; ASVCP = American Society of Veterinary Clinical Pathology; TE_a= total allowable error

Table 4.2: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Red blood cell count (RBC).

Analyser	**1**			**3**			**4a**			**4b**		
Day (time interval)	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4
SD	0.039	0.042	0.062	0.061	0.104	0.084	0.033	0.049	0.039	0.082	0.076	0.065
CV %	0.56	0.60	0.90	0.86	1.47	1.18	0.50	0.75	0.58	1.21	1.11	0.96
Mean difference	0.021	-0.033	-0.015	0.005	0.007	0.076	0.008	-0.013	-0.025	-0.056	-0.034	-0.017
Control limits range	0.240	0.210	0.370	0.360	0.520	0.500	0.199	0.297	0.234	0.412	0.455	0.325
ASVCP TE_a (%)	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TE_a (%)	4	4	6	5	8	7	3	5	4	8	7	6
Sigma metric based on ASVCP recommended TE_a	17.86	16.78	11.17	11.64	6.8	8.43	20.12	13.33	17.33	8.26	9.01	10.42
QC rule	1-3s	1-2.5s	1-3s	1-3s	1-2.5s	1-3s	1-3s	1-3s	1-3s	1-2.5s	1-3s	1-2.5s

CV = coefficient of variation; ASVCP = American Society of Veterinary Clinical Pathology; TE_a= total allowable error

Table 4.3: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Haemoglobin concentration (HGB).

Analyser	**1**			**3**			**4a**			**4b**		
Day (time interval)	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4
SD	0.085	0.049	0.083	0.081	0.061	0.063	0.134	0.152	0.168	0.117	0.127	0.164
CV %	0.50	0.48	0.50	0.57	0.73	0.39	0.85	0.97	1.07	0.72	0.13	0.16
Mean difference	-0.039	-0.020	-0.003	-0.034	-0.029	-0.042	-0.113	-0.108	-0.154	-0.089	-0.144	-0.127
Control limits range	0.510	0.481	0.498	0.450	0.366	0.366	0.804	0.760	0.840	0.701	0.635	0.820
ASVCP TE_a (%)	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TEa (%)	4	4	4	3	5	3	6	6	7	5	5	7
Sigma metric based on ASVCP recommended TEa	19.88	20.83	20	17.67	13.61	7.77	11.81	10.36	9.34	13.99	12.61	10
QC rule	1-3s	1-3s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-2.5s	1-3s	1-2.5s	1-2.5s

CV = coefficient of variation; ASVCP = American Society of Veterinary Clinical Pathology; TE_a= total allowable error

Table 4.4: Repeat Patient Testing-Quality Control: The Standard Deviation (SD), Coefficient of Variation (CV), Mean difference, Control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics and QC rule obtained for Haematocrit (HCT).

Analyser	**1**			**3**			**4a**			**4b**		
Day (time interval)	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4	Day 1-2	Day 1-3	Day 1-4
SD	1.67	2.67	4.14	2.25	3.23	4.86	1.96	3.23	4.23	2.21	3	4.6
CV %	3.47	5.41	8.24	4.47	6.21	8.34	3.94	6.37	8.23	4.34	5.84	8.78
Mean difference	-2.250	-3.980	-5.710	-3.150	-4.760	-6.000	-2.660	-4.420	-5.810	-2.890	-4.120	-6.220
Control limits range	8.375	13.33	20.70	11.28	16.12	15.81	9.89	16.15	21.15	11.05	15.00	23.00
ASVCP TE_a (%)	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TE_a (%)	22	34	50	25	35	45	25	39	50	27	37	54
Sigma metric based on ASVCP recommended TE_a	2.88	1.85	1.21	2.24	1.61	1.2	2.54	1.57	1.22	2.3	1.71	1.14
QC rule	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s

CV = coefficient of variation; ASVCP = American Society of Veterinary Clinical Pathology; TE_a= total allowable error

4.3.4 RPT-QC challenges

The intentional sample alterations using saline-diluted and frozen samples were detected as out-of-control by the RPT-QC for all measurands. Spiking samples with native canine lipids, excessive EDTA, and incorrect storage were successful in causing errors in some but not all measurands. WBC was out-of-control with added lipid. RBC counts and HGB concentrations were out-of-control in the specimens with excessive EDTA. RBC, HGB, and HCT were out-of-control when the specimen was incorrectly stored (extended storage beyond 4 days).

4.4 Discussion

We found that RPT-QC limits could be generated using a single protocol across our network of laboratories using 4 Sysmex haematology analysers. Additional information regarding the stability of specimens under defined storage and handling conditions is also possible from assessment of this data but beyond the scope of this study and paper. The collection of RPT data across 4 Sysmex analysers was the most challenging process of this study. The provision of a spreadsheet to technicians for entering data was essential for managing the RPT samples and data daily. Initial assessments across analysers and time periods for RPT-QC demonstrated that more than 20 data points are often required to produce RPT-QC limits that will validate with a

subsequent dataset of 20 results. This differed from prior recommendations to collect 20 data points (Westgard, 2016b) and prior studies that were able to validate limits determined using 20 data points (Flatland and Freeman, 2020; Flatland and Freeman, 2018b; Vanyo *et al.*, 2018). Following the RPT-QC limit generation, the validation process required calculation of limits using 30-40 data points for more than half of measurands/time interval evaluations. We found that by only using 20 results, the data was either not included within the control limits or did not adequately fill the range but when an additional 10-20 data points were used; then, we saw a better distribution across the QC data range. Except for HCT, successful validation with 20 data points was more likely for Day1-Day2 limits than Day1−Day3, and Day1−Day4. The calculations using 30-40 data points were more often required for Day1−Day3, and Day1−Day4 RPT-QC intervals, likely due to more variability in the data due to increased aging. We found the use of scatter plots highly valuable to visualise the distribution of the data and for outlier identification. Intuitively you would think the SD of the duplicates would provide RPT-QC limits that were not too wide or too narrow. However, when assessed visually, some RPT-QC limits were too narrow for the data set, and others appeared too wide, with the possibility of decreased sensitivity in error detection (Figure 4.1), which necessitated the use of additional data points to calculate the control limits.

As the sample aged, the trend was for the QC range to become wider, with higher SDs and CVs. This is not unexpected since more variation is expected with increased sample aging. However, it was interesting that the differences with increased aging were not uniformly similar within or across the

laboratories. We found this particularly evident for HCT, where we saw the biggest increases in range, SD, and CV as the sample aged. Interestingly, in chapter 2, we identified one of our analysers performing sub-optimally in comparison to the others in the network; this was also reflected in the RPT-QC performance for that analyser (Analyser 4b), demonstrating reduced analytical sensitivity.

The total error goals represented by ASVCP recommendations for haematology were achievable for the specified conditions ($P_{ed} \geq 0.85$, $P_{fr} \leq 0.05$, simple QC rule [1-2.5s or 1-3s] and n =1) for RBC and HBG for all 4 analysers and for 3 of the 4 analysers for WBC. The ASVCP recommended TE_a goals could not be achieved for HCT on any of the analyzers with RPT-QC. This was not an unexpected finding, as described in our study, that HCT is affected by MCV which has been previously found to increase with sample aging (Vanyo *et al.*, 2018; Flatland and Freeman, 2018b) and, therefore, was not included in this study. The inclusion of HCT was important despite the increased TE associated with differences used for RPT-QC. Although the total error that could be controlled with RPT-QC was unacceptable for interpretation of patient results, knowledge of the expected difference for the various time intervals was considered important for QC and optimising our QC procedures for RPT-QC.

The sigma metrics using ASVCP TE_a goal for the RPT-QC for RBC and HGB are clearly reflective of world-class performance. Suboptimal sigma metrics (< 3 sigma) using the ASVCP quality goal of 10% was seen for HCT for all time intervals and measurands for each analyser; the recommended TE_a of 10% could not be achieved (Table 4.4) and high CVs were seen. Analyser 3

and 4b had lower sigma metrics compared with the other analysers and was due to the high CVs generated. The lower sigma metrics for analyser 4b likely reflect suboptimal performance and the need for servicing to determine if better performance can be obtained from this analyser. It is interesting that analyser 4b has consistently been identified as a poor performer within the group based on traditional QC performance evaluation and QC validation. This illustrates the fact that some analysers may not perform as well as others. The fact that Analysers 4a and 4b were within the same laboratory, with the same environmental conditions, maintenance, reagents, control materials, and technicians, indicates that sources of variation other than those inherent to the analyser itself were unlikely. This analyser required more frequent servicing (up to 2 times the servicing normally expected) to maintain performance based on our internal criteria for performance monitoring (sigma metrics < 5.5). The 5.5 sigma metric goal is considered useful for HCT as a basis for needing for service. For other measurands (RBC, HGB, WBC), a sigma metric < 5.5 indicates a need for service.

The RPT-QC challenges were sufficient to demonstrate that specimens falling outside of our selection criteria (fresh, canine EDTA samples filled to the fill line of the collection tube) caused out-of-control events using the QC limits for the reference analyser, demonstrating that the QC limits were sufficient to detect the induced changes in sample quality.

4.4.1 Implications for our network of analysers using RPT-QC

We can accommodate a 20% TE_a goal cut-off for WBC counts that is within reference intervals for dogs rather than the recommended TE_a of 15% by the ASVCP Quality Assurance and Laboratory Standards Committee, (Nabity *et al.*, 2018) based on the expert opinions of pathologists involved in our Quality Education, Performance, and Implementation Group (QEPI, comprised of 5 clinical pathologists and 7 laboratory technicians, managers, and directors with a special interest in quality). The use of 20% as a quality goal for TE_a for haematology does not result in significant changes in the interpretation of results, as indicated by the average difference of WBCs between Day1−Day2, Day1−Day3 and Day1−Day4 of 0.155, -0.214, and 0.094, respectively.

Continued monitoring to determine if all analysers can consistently achieve 20% total allowable error quality goals for WBC with RPT-QC is needed. The lowest achievable TE_a for HCT ranged from 22 - 27% across the analysers for Day1−Day2, from 24 - 39% for Day1−Day3, and between 45 - 54% for Day1−Day4 compared with the ASVCP recommendation of 10% (Tables 4.1-4.4). The TE_a for HCT that could be achieved was considered acceptable for RPT-QC because RBC and HGB were also assessed and able to achieve their respective ASVCP total error goals. In addition, a further control measure for HCT using a spun PCV (considered the gold standard for evaluation of the erythron in veterinary medicine) is generated whenever HBG x 3 differs from the automated HCT by 5 or more units.

An alternative of Day1−Day4 could be used for measurand measurements at least 6 hours apart to use on Mondays for laboratories operating on a Monday-Friday schedule. However, further research will be required to determine acceptable RPT-QC ranges for this option.

Two QC rules (1-2.5s and 1-3s) were assessed based on their ability to achieve acceptable P_{ed} (≥ 85%) and P_{fr} (≤ 5%) when n=1. The preferred QC rule is 1-3s, where one violation exceeding ±3 SDs from the mean deems the QC run a failure, this simple approach can be easily adopted by the laboratory and pathology personnel. We found not all measurands for each time interval could be validated using the 1-3s rule, in fact only 45% could be validated. We have tailored RPT-QC monitoring spreadsheets to include both 1-3s or 1.25s QC rules specific to each analyser; however, educating/training the user on the differences in QC rule violations is imperative for successful implementation.

Sigma metrics were employed as a quality performance indicator, where sigma > 6 is indicative of world-class performance (Westgard, 2016a). Our previous findings (Daly *et al.*, 2021) showed that a sigma metric < 5.5 using commercially available QCM indicated instability within the analytical system with a need for further investigations and troubleshooting procedures. Additional studies are needed to determine an acceptable sigma metric cut-off for monitoring the performance of RPT-QC. For the measurands at time intervals with high CVs and unacceptable TE_a compared with the reference analyzer, we found this was reflected in the sigma metric. For example, Day1−Day2 and Day1−Day3 WBC intervals or analyzer 4b had a sigma metric of 6.25 and 6.23, respectively, compared with the reference analyser,

where sigma metrics were 11.63 and 9.93, respectively. The poor performance of HCT was also reflected in the sigma metrics for all analysers in the network, demonstrating the utility of sigma metrics as a performance indicator for RPT-QC.

4.5 Conclusion

We found that RPT-QC intervals could be generated for a network of Sysmex XT-haematology analysers but that the use of 40 data points was frequently required to achieve the subsequent validation of the RPT-QC by a dataset of 20 points. The TE_a goals provided by the ASVCP for haematology can be achieved for RBC and HBG. Alteration in acceptable TE_a for WBC measurands from 15% to 20% was considered acceptable based on internal expert opinion, while larger TE_a goals for HCT (ranging from 22 to 54%, depending on the time interval) were considered acceptable based on the acceptable TE_a performance for RBC and HBG. Sigma metrics provide a unitless measurement of performance capacity and, as demonstrated in previous studies using commercially available QCM, provide a reliable indication of sub-optimal performance. Continued evaluation of RPT-QC in parallel with commercial QCM is needed to determine if an RPT-QC cut-off of < 5.5 for the sigma metrics of various measurands (RBC, HGB, and WBC) indicates the need for analyser service to maintain and improve analyser performance. The challenges produced for RPT-QC to mimic unstable system performance demonstrated that RPT-QC was sensitive to alterations in system performance. Continued monitoring in parallel with commercial

QCM is required to determine if RPT-QC is reliable for daily internal quality control for haematology.

CHAPTER 5: REPEAT PATIENT TESTING-QUALITY CONTROL COMPARED TO COMMERCIAL QUALITY CONTROL MATERIAL FOR THE SYSMEX XT-2000*iV* HAEMATOLOGY ANALYSER IN A MULTI-SITE VETERINARY LABORATORY

This chapter is a follow up to chapter 4 outlining the implementation process for RPT-QC and evaluates RPT-QC in parallel with commercial quality control material. This research was led by S Daly and in collaboration with Dr Peter Graham and Dr Kathy Freeman as co-authors on the manuscript produced from this research and represents original work. Dr Mark Rishniw was consulted for VB coding when automating the spreadsheets.

Adaptations: This chapter has been changed from its previously published version to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

Reference: This chapter has been published as Daly S, Rishniw M, Graham PA, Freeman KP. Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT-2000*i*V hematology analyzer in a multi-site veterinary laboratory. *Vet Clin Pathol*. 2024; 53(Suppl. 1): 48-59. doi:10.1111/vcp.13332

Received: 29 March 2023 | Revised: 28 December 2023 | Accepted: 12 January 2024

DOI: 10.1111/vcp.13332

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

ORIGINAL ARTICLE

Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT-2000*i*V hematology analyzer in a multi-site veterinary laboratory

S. Daly[1,2] | M. Rishniw[3] | P. A. Graham[2] | K. P. Freeman[3]

[1]Veterinary Pathology Group, Brushville House, Dosco Business Park, Cork, Ireland
[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK
[3]Veterinary Information Network, Davis, California, USA

Correspondence
S. Daly, Veterinary Pathology Group, Brushville House, Dosco Business Park, South Douglas Road, Cork T12 XR6V, Ireland.
Email: susan.daly@thevpg.ie

Abstract

Background: Quality control material (QCM) for hematology in veterinary laboratories is limited, and repeat patient testing quality control (RPT-QC) is an alternative method using excess matrix-specific samples.

Objectives: This study aimed to determine if median differences between RPT-QC analyses for each time interval for RBC, HGB, HCT, and WBC were the same, determine if unified RPT-QC limits can be applied to a network of veterinary laboratories, compare the performance of RPT-QC to commercial QCM for the reference analyzer and evaluate the experience over a 4 month period and design, improve and implement an automated spreadsheet for RPT-QC data management.

Methods: The potential to unify individual analyzer RPT-QC limits for red blood cells (RBC), hematocrit (HCT), hemoglobin (HGB), and white blood cells (WBC) on multi-site Sysmex XT-2000-*i*V analyzers was explored by a difference of means test and confidence interval determination for the median difference for each network analyzer in comparison to the network reference analyzer. User experience of an automated RPT-QC data management Excel spreadsheet was collected by user feedback during monthly meetings. Numbers of out-of-control results and the root causes for these for RPT-QC were compared against those of a commercial QCM over a 4-month period.

Results: Differences between individual analyzer RPT-QC limits were too large to allow for unification of network limits. The automated spreadsheet successfully highlighted out-of-control events for RPT-QC. Trends or shifts were more frequent for commercial QCM based on observed performance and a 1–2.5s QC rule than for RPT-QC. Following routine troubleshooting, RPT-QC out-of-control events were resolved with an alternative RPT-QC sample indicating random error associated with excessive deterioration. Use of an automated spreadsheet for recording RPT-QC, documentation and troubleshooting of out-of-control events, and collating monthly summary calculations were considered an asset in laboratory quality management.

5.1 Introduction

Quality control material (QCM) for haematology analysers in veterinary clinical laboratories is limited and relies on commercially available QCM which is often of human origin and therefore is not a representative matrix for veterinary specimens (Flatland and Vap, 2012). This is problematic for veterinary laboratories that want to generate results of the highest standard (Unsal *et al.*, 2010). Previous reports (Daly *et al.*, 2021) indicate that a commercial QCM using the manufacturer's acceptable ranges ('bottle values') does not reliably demonstrate out-of-control events due to wide manufacturer's acceptable ranges (Manzocchi *et al.*, 2017). Quality control based on observed individual analyser performance is optimal (Daly *et al.*, 2022) for ensuring veterinary laboratories are maintaining standards according to quality benchmarks designed to highlight unstable system performance. The study carried out in chapter 3 for this network of veterinary haematology analysers that were subject to QC validation using a 1-2.5s QC rule showed a greater sensitivity for impactful error detection than using manufacturer's ranges.

The only commercially available option for veterinary laboratories is to use human-derived QCM, which is costly and has a relatively short shelf life (Flatland and Vap, 2012). However, there is an alternative called repeat patient testing quality control (RPT-QC). RPT-QC utilises retained excess patient samples representative of the laboratory throughput and relies on the principle that under defined time intervals and storage conditions, an EDTA haematology specimen will vary within a defined range from the baseline

measurement to the repeated measurement for the same specimen. If the repeated measurement exceeds these expected limits, then instability may be present within the analytical system (Westgard, 2016a). RPT-QC has been reported in the human literature for some time (Dutra, 1966; Carstairs *et al.*, 1977; Orser, 1983; Riddick, 1972) but was not utilised widely until further research emerged following a protocol proposed by Westgard (Westgard, 2016b). In veterinary laboratories, its application has been limited until recently (Flatland and Freeman, 2018b; Vanyo *et al.*, 2018).

RPT-QC was previously validated within a network of Sysmex XT-2000*i*V veterinary haematology analysers and then challenged to demonstrate that out-of-control events reflecting unstable system performance was detected in chapter 4. The network of harmonised haematology analysers (4 analysers participated across 3 separate sites) generated upper and lower RPT-QC limits for each analyser (Daly *et al.*, 2023). An RPT-QC system was validated and achieved total error equivalent to, or less than, total allowable error (TE_a) recommended by the American Society for Veterinary Clinical Pathology (ASVCP) (Nabity *et al.*, 2018) for red blood cells (RBC), haemoglobin (HGB), white blood cells (WBC) but not for haematocrit (HCT). However, HCT was still considered useful to include in statistical quality control; despite its poor performance, (as previously identified) understanding the expected difference for the time intervals was relevant information for quality control.

This study evaluated the implementation and suitability of RPT-QC in a harmonised network with the following objectives:

1. Determine if median differences between RPT-QC analyses for each time interval for RBC, HGB, HCT, and WBC were the same.

2. Determine whether unified RPT-QC limits could be applied across all network analysers or if RPT-QC limits need to be developed based on the performance of individual analysers.
3. Compare the performance of RPT-QC with that of a commercial quality control material for the reference analyser.
4. Develop, improve, and implement an automated data management spreadsheet to simplify RPT-QC within the laboratory.
5. Evaluate and summarise the experience of implementing RPT-QC within a commercial laboratory network using RPT-QC and commercially available QCM in parallel over 4 months, including technicians' feedback, challenges, and improvements made over the trial period.

5.2 Materials and Methods

For this study, measurands evaluated using RPT-QC were RBC, HGB, HCT, and WBC, and for each of the following repeat time intervals: Day1-Day2, Day1-Day3 and Day1-Day4. Repeat intervals were selected to provide QC limits applicable to daily (Day1-Day2), weekend (Day1-Day3), and public holidays (Day1-Day4).

An alternative RPT-QC protocol was also generated for longer closure periods (>Day1-Day4 interval) or if a suitable RPT-QC sample is not available. This consisted of running a freshly collected haematology specimen without visual haemolysis or lipaemia that was not more than 10%

above the upper or below the lower reference limit i.e. as close to or within reference limits and using this as the Day 1 reading. This sample was analysed again in 3-4 hours and results used as Day 2 reading. The Day1-Day2 RPT-QC limits for acceptable performance were used.

Technicians were fully trained, having undergone the required training according to the standard operating procedure (SOP) for haematology within the laboratory network, and passed competency evaluations prior to being entrusted with this task. Analysers were maintained according to the manufacturer's instructions and a standard operating procedure (SOP). The SOP for RPT-QC was developed in the first author's laboratory and then shared across the analytical network. The analysers used in the study were Sysmex XT-2000*i*V analyzers (Sysmex Corporation, Kobe, Japan) designated by numerical identifiers (1 [reference analyser], 3, 4a and 4b) according to network locations. Canine EDTA samples selected by each individual laboratory site were used for RPT-QC, and a commercially available control material (Level 2 – Normal e-CHECK (XE)-Haematology Control, Sysmex Corporation, Kobe, Japan) was used for the daily QCM.

5.2.1 Determine if differences between RPT-QC repeat analyses for each of the time intervals differed across the network.

The differences between individual analyser replicate measurements for each time interval (Day1-Day2, Day1-Day3 and Day1-Day4) for RBC, HGB, HCT, and WBC were compared across the network using the Kruskall-Wallis statistical test.

5.2.2 Determine if unified RPT-QC limits can be applied across the network of analysers.

RPT-QC limits were calculated in chapter 3, and the application of a single set of RPT-QC limits across the network of analysers requires that the differences in replicate analysis used to determine the SD_{dup} for each interval (Day1-Day2, Day1-Day3 and Day1-Day4) do not differ between analysers. The disparity between individual analyser replicate differences was explored by a difference of means test and confidence interval determination for the median difference (Wheelan, 2013).

For the differences of the means test, individual analyser replicate differences [the difference between the initial result (Day 1) and subsequent replicates (Day2, Day3, Day4)] were compared to those of the reference analyser (analyser 1). The absence of a difference between mean replicate-differences was reflected by overlap in the confidence intervals for the mean replicate-difference. If no overlap in the confidence intervals for the mean difference were present, these mean differences were considered to be different from one another.

5.2.3 Comparison of RPT-QC with a commercial quality control material

In the reference laboratory (analyser 1) and analyser 3, RPT-QC and a commercially available QCM (Level 2 – Normal e-CHECK (XE)-Haematology Control, Sysmex) were assayed at analyser start-up, prior to analysis of

patient samples, using an SOP for 4 months. A single RPT-QC specimen was chosen daily as previously described for RPT-QC, labeled with a sticker on the stopper, and analysed for RPT-QC the next working day.

Commercial QCM and RPT-QC results were evaluated to determine the number of out-of-control (OOC) events, the number of repeated assays, and whether the OOC events were the result of random error (no underlying problem detected) or reflected true instability of the system. This was counted on a daily basis and comments noted on the QC spreadsheets.

As an additional comparative evaluation tool, the sigma metric for each measurand on each analyser was calculated using the ASVCP recommendations for TE_a (Nabity *et al.*, 2018), the CV from each RPT-QC interval and absolute bias = 0% according to the following formula.

$$\text{Sigma} = (TE_a\% - \text{absolute bias}\%)/CV$$

The CV for each RPT-QC interval was determined by the following formula:

$$\text{CV for RPT-QC interval} = (\text{SD of duplicate measurements}/\text{mean of the average of the duplicate measurements}) \times 100$$

Sigma metrics for the commercial QCM were calculated daily whereas the RPT-QC automated spreadsheet provided monthly sigma metrics for each measurand as the sigma metrics were consistently high and monthly frequency was considered sufficient.

5.2.4 Automated spreadsheet development

To increase the efficiency and effectiveness of statistical QC, RPT-QC daily data management spreadsheets were developed using Microsoft® Excel® for Microsoft 365 MSO (Version 2208 Build 16.0.15601.20072), 64-bit Macros were coded within Visual Basic using the Developer Tab in Excel in order to generate more automated applications/operations. The spreadsheet included each measurand data entry, hard-coded RPT-QC limits generated in chapter 4 and automated out-of-control alerts for the technician operating the Sysmex haematology analyser. A separate section, including a troubleshooting checklist, was developed for recording of out-of-control events and actions taken to bring the system back into control. A monthly Levey-Jennings (L-J) chart was automatedly generated and reviewed each month or generated more frequently, if the technician desired. Monthly refinement of the RPT-QC control limits was conducted using the pooled SD, which is a cumulative SD (Westgard, 2016b) to allow for ongoing QC limit updates. The automated spreadsheet was developed in conjunction with a veterinary specialist in spreadsheet programming (second author), and first used in the first author's laboratory and then shared across the network.

5.2.5 Technician experience and feedback

Technicians were consulted frequently during the design and trial phase through a series of regular meetings and provided ongoing feedback

regarding improvements in the spreadsheet at the first author's site. The spreadsheets were checked for any inconsistencies by senior technicians which provided upskilling and improved competence.

5.2.6 Statistical analysis

The Kruskal-Wallis test (Ostertagová *et al.*, 2014) was performed on RBC, HGB, HCT, and WBC replicates for each time interval (Day1-Day2, Day1-Day3 and Day1-Day4) for all analysers using IBM SPSS statistical software (version 28). This test determines whether there is a significant difference between independent groups by testing the null hypothesis that the medians of the groups are the same; $P < 0.05$ indicated a significant difference.

5.3 Results

The data used for further evaluation and statistics in this study is available from supplementary material, Appendix E, Table E5.1.
Supplementary material is available in appendix E, Table E5.2 which presents example calculations for RBC Day1-Day4, n=20.
For RBC, there is a difference between the median differences for Day1-Day2 but not for Day1-Day3 and Day1-Day4 across the analysers. For HGB there was no difference for Day1-Day2, but Day1-Day3 and Day1-Day4 showed a difference between the median differences across the analysers.

There is a difference between the median differences for HCT for Day1-Day2 and Day1-Day3 but no difference for Day1-Day4 across the analysers.
For WBC, there was no difference between the analyzers for Day1-Day2 and Day1-Day4, but Day1-Day3 showed differences in medians across the analysers, (Table 5.1) presents data on Kruskal-Wallis test.

Table 5.1: Kruskal-Wallis Test comparing median differences across the analyser network, significance set to $P < 0.05$.

Measurand	Time Interval	Test Statistic	P	Df	N
RBC	Day1Day2	13.019	0.005	3	133
	Day1Day3	4.111	0.250	3	131
	Day1Day4	3.862	0.277	3	129
HGB	Day1Day2	5.851	0.119	3	154
	Day1Day3	22.551	0.000	3	130
	Day1Day4	21.162	0.000	3	129
HCT	Day1Day2	14.117	0.003	3	113
	Day1Day3	9.215	0.027	3	129
	Day1Day4	3.461	0.326	3	153
WBC	Day1Day2	6.938	0.074	3	134
	Day1Day3	13.082	0.004	3	112
	Day1Day4	4.628	0.201	3	151

RBC, red blood cell; HGB, haemoglobin; HCT, haematocrit; WBC, white blood cell

5.3.1 Evaluation to determine if unified RPT-QC limits could be used across the network of analysers.

There was no difference between the network analysers and the reference analyser for the mean replicate difference for WBC, RBC, HBG, and HCT across each time interval (Day1-Day2, Day1-Day3, Day1-Day4). This suggests that common RPT-QC limits could be applied across the analysers in the network if the SD of the duplicates were similar across the group of analysers (Table 5.2).

Table 5.2: Difference of means test showing standard error for the network analysers (3, 4a, 4b) when compared to the reference analyser (1) for each time interval and measurand (RBC, HGB, HCT, WBC).

	Day1-Day2	Day1-Day3	Day1-Day4
RBC			
3	189.890	-120.827	-187.930
4a	130.653	-111.229	-134.462
4b	356.568	5.040	-155.905
HGB			
3	-14.153	27.149	98.382
4a	115.438	84.946	159.115
4b	90.810	97.556	135.800
HCT			
3	2.139	1.010	0.266
4a	1.705	0.927	0.110
4b	2.388	0.321	0.513
WBC			
3	-17.658	-87.144	-180.096
4a	-95.223	16.103	-252.724
4b	-48.559	46.494	-238.689

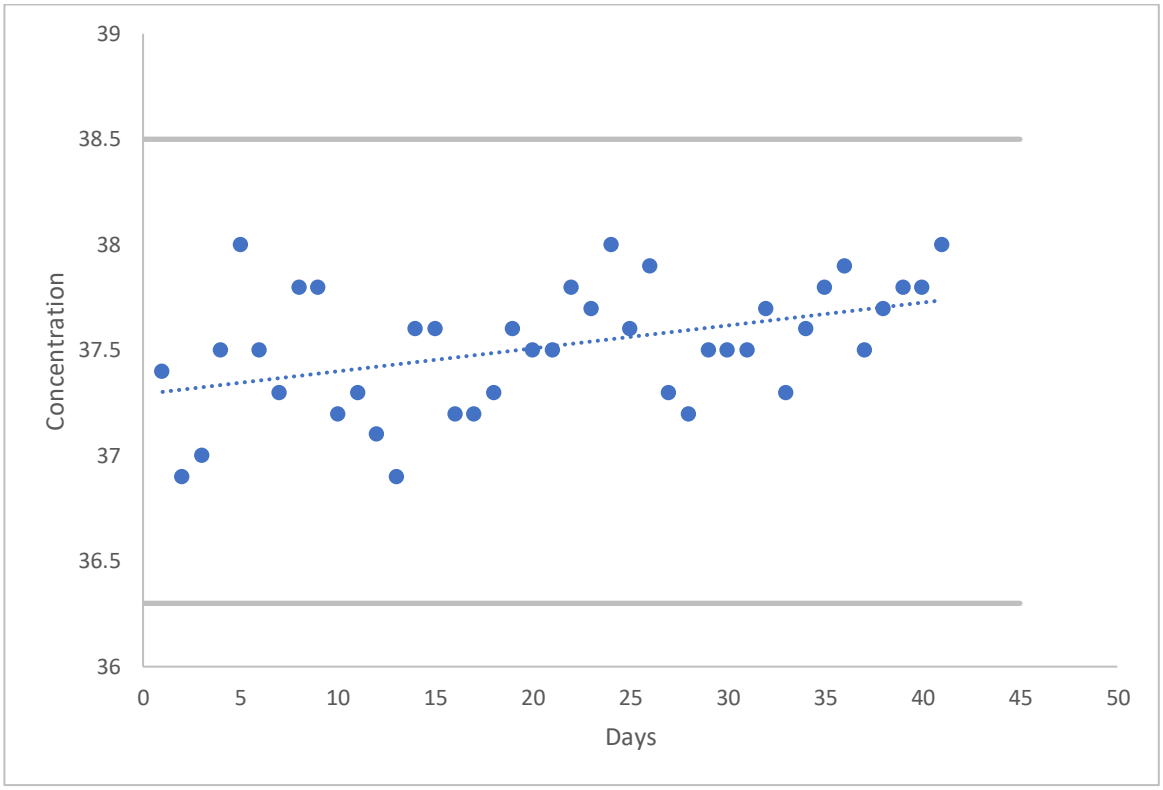
Abbreviations: RBC, red blood cell; HGB, haemoglobin; HCT, haematocrit; WBC, white blood cell

The control limits across the network had overlap for the majority of the confidence intervals for the mean differences for the network analysers compared to the reference analyser for each time interval when the same QC rule was used as the reference analyser. However, for RBC, HCT, and HGB, there were exceptions where no overlap occurred, indicating that common RPT-QC limits were not suitable for application across all of the analysers. Supplementary material available in Appendix E, Tables E5.3 – E5.6 presents confidence intervals for the mean difference, and upper and lower control limits for each measurand.

The performance of RPT-QC and commercial QCM over a 4-month time period for the reference analyser is compared and available in supplementary material, Appendix E, Table E5.7 and includes evaluations for RBC, HGB, HCT, and WBC. There were more OOC events overall with RPT-QC than with the commercial QCM.

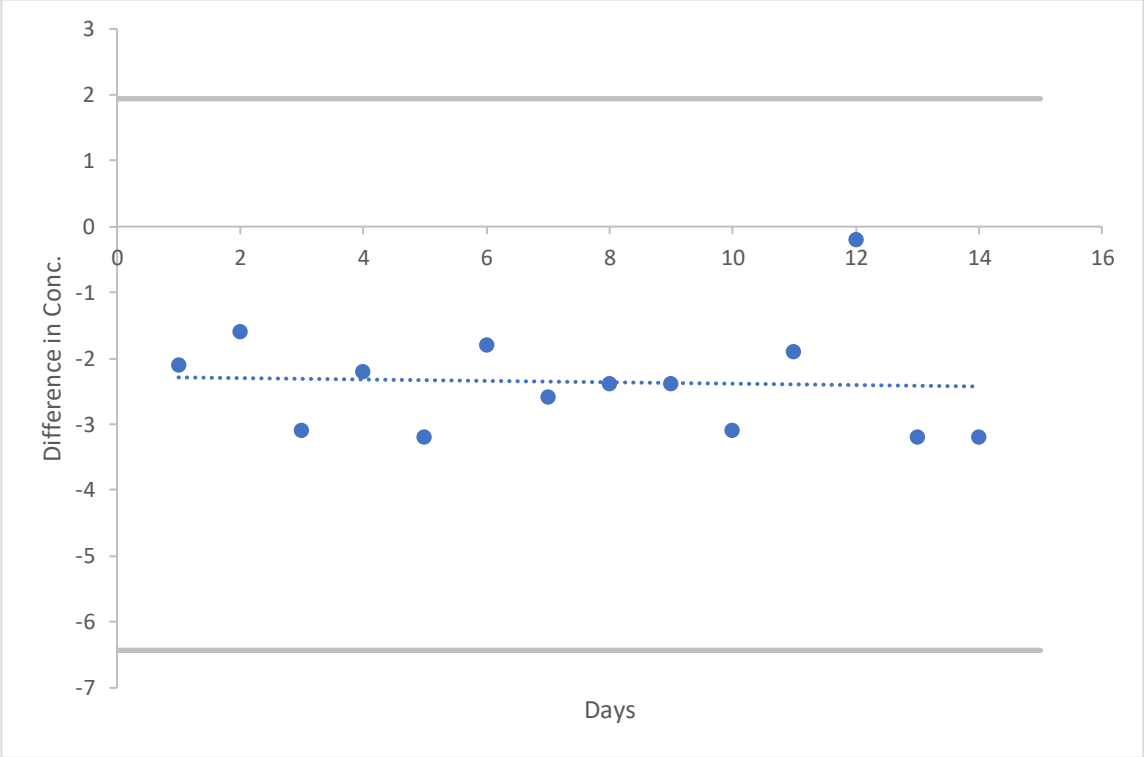
For events using a 1-2.5s QC rule, this was within the predicted P_{fr} (0.03), but for events using a 1-3s QC rule, this was greater than the predicted P_{fr} (0.00), which was unexpected. However, control charts for RPT-QC generally showed a more stable performance than that of commercial QCM (Figure 5.1, 5.2).

Figure 5.1: Levey Jenning (L-J) chart for commercial quality control material for haematocrit using a 1-2.5s QC rule.



The x-axis represents days and the y-axis is concentration for HCT. The grey lines represent the upper and lower reference intervals. The trendline was added to show rise in concentration over time.

Figure: 5.2 L-J chart for RPT-QC Day1-Day2 for HCT using a 1-2.5s QC rule.



The x-axis represents days and the y-axis is difference in concentration for HCT. The grey lines represent the upper and lower RPT-QC limits. The trendline is the mean difference and shows stable concentration over time.

Commercial QCM was within control limits (within 1-2.5s QC rule), and sigma metrics were > 5.5 except for one event where the sigma metric was < 5.5 but was > 5.0 (RBC) for the reference analyser.

RPT-QC was within control limits during this same time periods, except for 4 occasions for the reference analyser. These resolved with new patient samples subsequent to troubleshooting, comparable to using a 'new' commercial control for such incidences. Sigma metrics for RPT-QC were >

5.5 for all measurands except for HCT (< 3 using the ASVCP TE_a quality goal), based on monthly sigma evaluations.

5.3.2 Automated spreadsheet development.

Clear instructions for the use of the RPT-QC data management spreadsheet were developed (Figure 5.3 (A)), and the data collection summary tab records the initial calculations for each measurand (Figure 5.3 (B)). A separate tab for recording OOC data, troubleshooting, resolution, and corrective actions is included (figure 5.3 (C)).

Example images of functionality that the author incorporated into their spreadsheet design with the assistance of a visual basic coder are presented in Figure 5.3 (A-D). The development phase aimed to produce a document that was easy to use, but contained clear instructions and minimised the possibility of mistakes in data entry.

Figure 5.3 (A-D): Images of the automated data management spreadsheet.

(A) Image of the automated data management spreadsheet user instructions and tool bar for RBC measurand. **(B)** Initial calculations summary for RBC for each time interval Day1-Day2, Day1-Day3, Day1-Day4 and the pooled monthly SD and refined control limits. **(C)** Data entry for RBC showing one out of control event (Lab#76673) that alerts the technician by highlighting the datapoint required for further troubleshooting. **(D)** Troubleshooting page where actions taken can be logged which is a permanent record detailing the reason, action, outcome and responsible technician.

(A) HOW TO USE THIS SPREADSHEET

Select Measurand

DAILY USE

1. After reading instructions, click "**Select Measurand**" above
2. **Select First Measurand** from Dropdown List This will open 3 worksheets:
   A. Measurand Data Entry
   B. Measurand Actions Monday
   C. Measurand Actions Week
3. You should be on "**Measurand Data Entry**" page
4. **Select MONTH** (or ensure correct month is selected)
5. **Enter duplicate values** in cells corresponding to the correct Day/Date
6. **Click "Check OOC"** button
   Any runs that are out of control (OOC) will be highlighted by a yellow box
7. If a run has a highlighted OOC value, **click "Transfer OOC"** button
   This will transfer the OOC value to the appropriate "Measurand Actions " page
8. If no OOC value appears, **click "Select New Mesaurand"** This will return you to the Data Entry page

INSTRUCTIONS | Data Entry | RBC data entry | RBC Day Actions | RBC Week Actions | Dropdown List data

(B)

Starting values	Week	Day1-2	Day1-3	Day1-4
Data points	39	39	38	37
Mean Diff	0.021	0.021	-0.033	-0.015
Week SD_{dup}	0.039	0.039	0.042	0.062
Control Rule	1-3s	1-3s	1-2.5s	1-3s
Lower Limit	-0.096	-0.10	-0.138	-0.201
Upper Limit	0.138	0.14	0.072	0.171
Quality Goal	10	10	10	10
CV	0.560	0.5600	0.600	0.900
Sigma	17.9	17.9	16.7	11.1

(C)

MONTH: June | OOC ID | OOC Transfer

DAY	DATE	Lab #	Day Rpt	RBC Value 1	RBC Value 2	Difference
June 1, 2022	1	76654	Day 1-2	7.4	7.4	0
June 2, 2022	2	76655	Day 1-3	7.1	7.1	0
June 3, 2022	3	76659	Day 1-4	7.2	7.2	0
June 4, 2022	4	76672	Day 1-2	6.9	6.8	0.1
June 5, 2022	5	76671	Day 1-3	6.5	6.4	0.1
June 6, 2022	6	76675	Day 1-4	6.7	6.6	0.1
June 7, 2022	7	76677	Day 1-2	6.4	6.3	0.1
June 8, 2022	8	76679	Day 1-3	6.9	6.9	0
June 9, 2022	9					
June 10, 2022	10					
June 11, 2022	11					
June 12, 2022	12					
June 13, 2022	13					
June 14, 2022	14					
June 15, 2022	15					

... | Data Entry | RBC data entry | RBC Day Actions | RBC Week Actions | Dropdown List data

(D)

Back to RBC Data Entry

System Change, unusual result or out-of-control result documentation:

Date	Item	Reason	Action	Outcome	Technician
05-Jun	76673	Out of control	Re run on analyzer	New RPT-QC required	KF

5.3.3 Technician feedback

In the authors' opinions, technician feedback was positive, particularly once the automated data management spreadsheet was finalised and an RPT-QC standard operating procedure (SOP) and training was instituted. Overall, the technicians (5 technicians in total across 3 laboratories) were satisfied with the spreadsheet with a notable appreciation of the separate troubleshooting and recording section, which highlighted any issues to other technicians using the analyser.

Using RPT-QC resulted in fewer investigations of unusual results or trends when compared to 1-2.5s rule for the commercial QCM customised for individual analyser performance even though the commercial QCM wasn't out of control (OOC) i.e. results outside the specified limits for the measurand. The technicians also highlighted the advantage of having implemented RPT-QC when, on two occasions, the delivery of the commercial QCM was delayed, and on another occasion, the manufacturer's package inserts were not enclosed, precluding the determination of the assayed target value. RPT-QC was the only QC in use until the delivery issues were resolved (1 week).

One challenge to implementing RPT-QC was the reliance on getting fresh patient samples following longer closure periods; however, the inclusion of the alternative RPT-QC procedure was a useful aide to cope with these extended closure periods of which there were two occurrences, and this procedure was included in the SOP.

5.4 Discussion

RPT-QC was validated in chapter 4 for several Sysmex XT-2000*i*V veterinary haematology analysers,and the designed challenges indicated that RPT-QC could detect unstable system performance.

Overall, we were satisfied that RPT-QC could detect unstable system performance and that it was a less expensive, user-friendly, and readily available alternative to commercial QCM. The statistical evaluation was crucial in understanding the RPT-QC control limits and demonstrated that the application of unified (the same) QC limits across the network is not possible. The lack of overlap in confidence intervals for the various RPT-QC limits and detection of differences in median differences by the Kruskal-Wallis nonparametric evaluation (null hypothesis rejected) indicated that a common RPT-QC limit could not be applied across all laboratories. In addition, the different QC rules used meant the use of common QC limits would be problematic as each individual analyser requires a statistical QC validation to determine the appropriate QC rule. It is noteworthy that for future RPT-QC validations, it may work best to use 40 data points for all calculations (rather than 20) to avoid difficulty in subsequent assessments and to use a 1-2.5s rule to achieve P_{ed} > 0.90 instead of a 1-3s rule that is capable of > 0.85 P_{ed} (Daly *et al.*, 2023).

We did see some overlap in the confidence intervals for the upper and lower RPT-QC limits for most measurands for Day1-Day2 and Day1-Day3 time intervals using the Kruskal-Wallis test. Day1-Day4 showed no difference, and this was thought to be due to the non-proportionate deterioration of the

specimen across all time intervals as indicated by differing width of the control ranges (see supplementary material, Appendix E, Table E5.1), and by Day1-Day4, we expect the RPT-QC specimens to be approaching the end of their useful life for RPT-QC.

For RBC, no difference in median differences was seen for Day1-Day3 indicating a varying pattern of deterioration for this measurand that was not as uniform as the other measurands. This nonlinear variability in deterioration also indicates that individual analyser RPT-QC validation for QC limits is needed. This study supports the need for individual analyser RPT-QC validation and QC limit generation for each site. The apparent variability of deterioration at each site may reflect different laboratory conditions such as temperature, humidity, maintenance, and operator impact on performance.

This study found that unusual results/trends were more frequent for commercial QCM based on observed performance using a 1-2.5s QC rule than compared to RPT-QC (see supplementary material, Appendix E, Table E5.4). However, no OOC results were identified in commercial QCM data over the 4-month period in the reference laboratory. Trends in commercial QCM data were seen towards the end of the open vial shelf life, usually by day 4 and/or day 5, indicating deterioration, although the degree was not considered to be of clinical significance. This resulted in increased time spent on evaluating and rerunning the QCM to determine if any trends were likely to be of significance, even if not outside of the 1-2.5s range.

RPT-QC had 4 OOC (RBC, HGB) events over 4 months duration for the reference analyser highlighted by the automated spreadsheet. The use of an RPT-QC troubleshooting checklist failed to demonstrate an underlying cause

for these events, and a new RPT-QC specimen was then used to resolve each out-of-control event. Since a new RPT-QC specimen produced in-control results, it is likely that the original OOC result was due to random error associated with an unexpected level of specimen variation, which resulted in an increase in the width of the QC range, which was observed with increasing storage time; instability in the system was not evident when the commercial QCM was run.

Sigma metrics were flagged on two occasions for the reference analyser using the commercial QCM despite the QC being deemed in-control, highlighting poor performance as per criteria established in a chapter 3 for the network of analysers using a 5.5 sigma quality strategy. This indicates that commercial QCM control limits, even when validated for observed performance, are not reliable as the sole indicator of system performance, and a combined quality strategy is recommended (Westgard, 2016c). Sigma metrics for RPT-QC were generated monthly and identified poor performance for HCT using the recommended ASVCP TE_a quality goal. The 'reverse approach' to the determination (Korchia and Freeman, 2022) of achievable total error based on the use of a 1-2.5s rule, with one level of control material, indicated that total error goals for HCT that can be controlled should be in the range of 22-54% depending on the interval between patient sample evaluations. Use of these as expert opinion goals resulted in sigma metrics > 6 for HCT.

The use of an automated data management spreadsheet for RPT-QC meant less time spent on data entry and analysis than when using manual Excel spreadsheets that were originally used for QC recording. They also facilitated

the central electronic archiving of records and sharing of data amongst the laboratories. Manual Excel spreadsheets represent an alternative for point-of-care testing or reference laboratories with a single haematology analyser. The automated spreadsheet was developed with a laboratory network in mind, but it worked well for the individual site (reference laboratory) where it was initially trialed and refined for any coding issues. The instructions on the spreadsheet prevented erroneous use and worked well as a quick guide to technicians not using RPT-QC frequently. The automated spreadsheet was helpful in ensuring accurate identification of OOC events as these were flagged on the spreadsheet and then transferred to another section where the troubleshooting process was documented. This was useful as it meant only the in-control data were retained for calculating the monthly pooled SD and RPT-QC limits for the subsequent month. The revised RPT-QC limits were usually only slightly altered from the previous month, indicating the validity of the mean replicate-difference and calculated RPT-QC limits, but the spreadsheet allowed refinement with increasing data for new RPT-QC limits for the next month. The level of automation freed up the technician from manual calculations and/or coding errors when copying/pasting formulas. The monthly L-J chart and monthly sigma metric were useful in monitoring the ongoing performance of the 5.5 sigma quality assurance strategy as described in chapter 3. Overall, the spreadsheet was user-friendly as all data collection, OOC, unusual or scheduled events, and troubleshooting were in one place.

The experience of using RPT-QC in parallel with a commercial QCM was positive. The feedback from the technical team was that the use of the

automated data management spreadsheet was helpful. The involvement of all technicians in the development of the automated spreadsheet was also helpful in engaging them in the process and developing and using the improvements suggested. It had all the required capabilities needed to carry out RPT-QC monitoring. A disadvantage is that the spreadsheet design involved coding that was outside the capabilities of the laboratory technicians and, therefore, any retrospective or prospective changes would need to be carried out by personnel proficient in this area. An outside consultant (second author) was required for proficiency in advanced automation of spreadsheets based on the authors'/technicians' feedback.

As with any procedure in the laboratory, a standard operating procedure (SOP) was drafted. This aided the technical team with training and was used in the roll out of RPT-QC to the laboratory network.

Supply chain issues with commercial QCM were noted during COVID and following Brexit, further supporting the need for alternative QC methods. The availability and elimination of the cost of commercial QCM are major advantages of RPT-QC. RPT-QC does not incur costs for the laboratory as it uses excess patient samples. Any disadvantage for RPT-QC resulting from extended closures for holidays beyond the Day1-Day4 time interval is mitigated by having an alternative protocol that may be used during periods of long closure. The spreadsheet was also designed to allow the use of alternative time intervals where Day1-Day2 specimens are not widely available.

In addition, the canine RPT-QC specimens are a suitable commutable matrix for our veterinary laboratory network and a representative matrix for

veterinary patients (Unsal *et al.*, 2010). Canine specimens are rarely unavailable, and various levels of clinical interest could also be used. Controlling costs in the commercial laboratory is a consideration and reducing cost by approximately €1000 per analyzer (for a single level of commercial QCM) per year by implementing RPT-QC results in substantial savings for a laboratory. The cost savings are an added benefit following all of the above. In summary, the major advantages of using RPT-QC are cost savings, a QC matrix commutable with patient samples, the ability to refine control limits based on the use of a pooled SD_{dup} over time, and the use of an automated spreadsheet.

We chose to evaluate RBC, HGB, HCT, and WBC, in particular for RPT-QC, to streamline the process of RPT-QC and considered these adequate due to the high emphasis on nonstatistical QC for haematology, including technical review of blood smears. We elected not to include the calculated RBC indices (MCV, MCH, MCHC) due to the effects of specimen aging. RPT-QC should also work for other haematology measurements, but further review of the effects of specimen aging on platelets, reticulocytes, and WBC differential counts is needed. The technical review of the blood smear was considered adequate to QC these methods, with consideration of the correlation of the microscopic findings with the automated haematology patient results.

5.5 Conclusions

This chapter/study showed that RPT-QC can be used within single veterinary laboratories and across a laboratory network to function as a commutable and cost-effective method of quality control. We demonstrated statistically

that RPT-QC requires individual QC limit generation for each analyser because some median differences in the repeated results are different across the time intervals and the network instruments. RPT-QC does not have excessive repeats, shifts, or trends in data compared with QC based on observed performance using a commercial QCM. The positive feedback from technicians using the automated data management spreadsheet meant all data entry and monitoring was in one place, and L-J charts were a great visual aid and summary of results, necessitating the need for spreadsheet automation. The refined SD_{dup} will require further investigation, but early analyses indicated the QC limit generation from pooled monthly SDs in normal RPT-QC will result in ongoing refinements of RPT-QC limits, leading to improved application of quality control.

CHAPTER 6: MAINTAINING HARMONISATION OF SYSMEX XT-2000/V ANALYSERS IN A NETWORK OF VETERINARY LABORATORIES USING AN INTERLABORATORY COMPARATIVE TESTING PROGRAM

This stage of the research involved the design, evaluation and implementation of an internal laboratory comparative (ILC) testing program in order to maintain the harmonisation achieved in chapters 2 and 3 by comparing the network analysers to the designated reference analyser for the laboratory network. This chapter utilises canine samples which addresses the lack of available QCM that is of animal origin and demonstrates an appropriate specimen for maintaining harmonisation. This process included a performance re-evaluation in order to generate samples of various levels of clinical interest and to determine SDs and CVs for the comparative program. Data collected and analysis was carried out by S Daly and in collaboration with Dr Peter Graham and Dr Kathy Freeman.

Adaptations: This chapter has been changed from its previous version for publishing to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

This chapter has been submitted to the journal of *Veterinary Clinical Pathology* and is currently under-review.

6.1 Introduction

The implementation of laboratory harmonisation has been shown to encourage maintenance and operation of analytical equipment to a high standard, which, in turn, improves the total quality output (Daly *et al.*, 2021; Daly *et al.*, 2022; Theodorsson, 2016). This is not an easy undertaking but worthwhile as studies have shown continuous improvements following harmonisation which is demonstrated in chapter 2. Harmonisation as part of a total quality management system is well described in the literature for human laboratory medicine (Theodorsson, 2016; Unsal *et al.*, 2010; Manzocchi *et al.*, 2017) but adoption by veterinary laboratories, if present, is lacking in the veterinary literature. Veterinary laboratories face considerable challenges compared to human medicine for internal and external quality control materials which are commercially available but typically of human origin and, therefore, do not share the same components or matrix as the veterinary specimens (Unsal *et al.*, 2010). The matrix of the specimen refers to all the components besides the measurand and differences between the matrix of the QC material and patient specimens may contribute to bias, also called the ‘matrix effect’. Ideally the QC material should be the same matrix as the patient to minimise bias. This causes some difficulty for veterinary laboratories due to the lack of commercial species-specific QCM. For a

network of laboratories maintaining a harmonised status, QC material and lot numbers should be consistent across the network. However, chapter 2 found that even this may be unreliable in identifying unstable system performance if manufacturers recommended acceptance ranges ('bottle values') are used, (Manzocchi *et al.*, 2017) as these are based on results obtained across multiple laboratories and instruments. Additionally, commercial QCM can incur a substantial economic burden on laboratory networks that wish to achieve harmonisation between their analysers.

Previous harmonisation studies in veterinary laboratories have outlined an initial process of evaluating analyser performance with traditional QCM, (Daly *et al.*, 2021) QC validation using traditional QCM (Daly *et al.*, 2022) and to address species matrix concerns, the development, validation and implementation of repeat patient testing quality control (RPT-QC) limits contributes to maintaining harmonisation by providing species-specific internal quality control and assessing imprecision across a commercial laboratory network.

One approach to laboratory harmonisation uses common quality metrics to assess analyser performance by comparing results from network analysers to those from a designated EQA-monitored reference in-network analyser. The evaluated quality control metrics include measuring the observed total error (TE_{obs}) – the sum of the bias and imprecision in a measurement – and estimating sigma metrics across a range of clinically relevant control materials. The TE_{obs} is then compared against the total allowable error (TE_a), as recommended by the American Society for Veterinary Clinical Pathology

(ASVCP) (Nabity *et al.*, 2018) or by expert opinion, desirable biologic variation-based goals for bias and CV and the biologic variation goal for laboratory systems ($0.33CV_I$) (Fraser, 1992b). Similarly, a sigma (σ) metric result ≥ 6 for an analytical method indicates 'world class' performance, reflecting ≤ 3.4 defects (erroneous results) per million results (Westgard and Westgard, 2015). A 5.5 sigma standard was previously chosen for veterinary haematology analytical performance, (Daly *et al.*, 2022) becoming a continuous monitoring tool throughout the development of a total quality management system for this network and a useful tool for maintaining ongoing harmonisation.

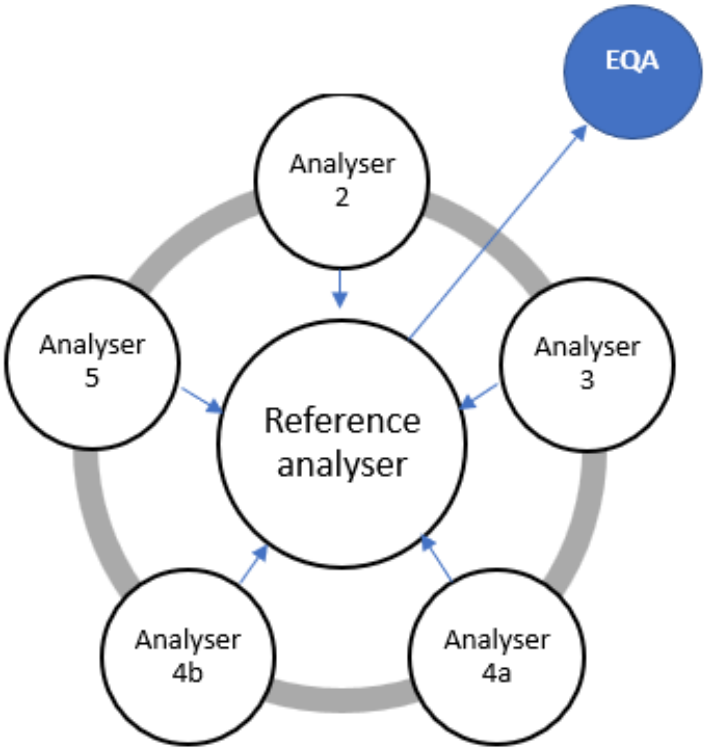
Whether laboratories can achieve reasonable harmonisation through such an interlaboratory comparison testing program using a species-specific (matrix-appropriate) material remains unresolved. Therefore, we explored the development and use of a low cost and matrix-appropriate (Unsal *et al.*, 2010) interlaboratory comparison (ILC) testing program in place of multiple site subscriptions to a traditional EQA program and its role in maintaining harmonisation of a network of laboratory analysers.

6.2 Materials and methods

In keeping with good practice for establishing harmonisation (Greenberg, 2014), performance evaluation (Daly *et al.*, 2021) and QC validation (Daly *et al.*, 2022) were prerequisite steps carried out on all haematology analysers within the network to ensure optimal performance across the network, to

identify analysers requiring attention and to ensure the ability to use common reference intervals. The laboratory network consisted of 6 Sysmex XT-2000*i*V analysers, (Sysmex corporation, Kobe, Japan), in 5 laboratory locations in the UK (n=4) and Ireland (n=1) using one haematology analyser in the network (located at the first author's site) as the reference analyser. We chose to include the following measurands as they are routinely reported as part of a CBC: red blood cell count (RBC), haemoglobin (HGB), haematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), white blood cell count (WBC), neutrophil (NEUT) count, lymphocyte (LYMPH) count, monocyte (MONO) count, eosinophil (EOSO) count, basophil count (BASO), platelet count (PLT), Red cell distribution width-coefficient of variation (RDW-CV), Plateletcrit (PCT) and Reticulocyte count (RETIC). The reference analyser was chosen based on multiple performance evaluations demonstrating excellent ongoing performance and the best performing analyser within the network. The reference analyser (Analyser 1) continued to subscribe to an EQA program (RIQAS monthly program) (Randox) and the network analysers (Analyzer 2, 3, 4a, 4b & 5) were each evaluated against the reference analyser using the interlaboratory comparison testing program (ILC). See Figure 6.1.

Figure 6. 1: Schematic for a harmonised network of analysers showing the network analysers linked to the reference analyser which remains subscribed to a EQA program.



6.2.1 Multi-level performance evaluation

To characterise performance across a range of results of clinical significance for canine blood and to adequately maintain harmonisation for a network of analysers we chose a multi-level approach which in turn required us to understand analyser performance across a range of concentrations. Five categories/levels (‘bins’) of canine samples with varying concentrations corresponding to 1) marked anaemia and leukopenia, 2) slight to moderate anaemia and leukopenia, 3) within reference interval for haematocrit and WBC, 4) slight to moderate polycythaemia, and 5) marked polycythaemia

were evaluated by 10 replicate analyses (multi-analysis event) on each analyser in order to characterise the instrument performance and to generate SDs and CVs which were then compared to desirable CV goals for the ILC testing program. Such limits could then be used to ensure there was not unacceptable imprecision or bias for any analyser within the network. The CVs generated were also used to calculate sigma metrics to monitor the > 5.5 sigma strategy used to indicate acceptable performance across the network for the single analysis events (Daly *et al.*, 2022) and for observed total allowable error (TE_{obs}) using the following formulas:

$$\text{Sigma metric} = (TE_a\% - |\text{Bias}\%|)/CV\% \text{ (Westgard, 2001)}$$

$$TE_{obs}\% = |\text{Bias }\%| + 2 \times CV\% \text{ (Nabity } et\ al.\text{, 2018)}$$

Where fresh samples containing the required concentration for the analysis were not available, samples were manipulated by the addition or removal of plasma until the desired concentrations were achieved. The easiest way to generate these levels was to aim for desirable haematocrit (HCT) percentages with a corresponding dilution or concentration of haemoglobin (HGB), red blood cells (RBC) and white blood cells (WBC). See Table 6.1 for desirable percentages of HCT of clinical interest for the multi-level event.

Table 6. 1: Five levels of haematology results of clinical interest generated based on haematocrit (HCT)(%) created by addition or removal of canine plasma.

Category	HCT target (%)
Anaemia +++	10-15%
Anaemia ++	20-25%
Normal	40-60%
Polychythaemia +	60-65%
Polychythaemia ++	70-75%

The results of the repeated analyses across the various ‘bins’ were recorded for each analyser using spreadsheets (Microsoft® Excel® for Microsoft 365 MSO (Version 2208 Build 16.0.15601.20072) to generate mean, SD and CVs for each category that could be compared to the reference analyser in the network and to the biologic variation-based desirable CV goal.

6.2.2 Single analysis events

Once SDs and CVs were generated from the multi-level performance evaluation, fresh specimens to be used for each single-analysis ILC event were manipulated by adding or removing plasma to cover the range of HCT concentration targets and corresponding HGB, RBC and WBC. A single ILC sample was scheduled every 6-8 weeks and sent from the reference

laboratory by overnight courier to each laboratory within the network. Each sample assessed performance using various levels (bins) across the range of clinically important medical decision limits for HCT and corresponding dilution or concentration of leukocytes and platelets. An email alert reminding the laboratories of the next-day delivery to all sites were sent prior to each event. Sample handling instructions were enclosed with each ILC sample. All sites (including the reference analyser) analysed the sample once on the same day. Results were returned by each laboratory via an automated site-specific spreadsheet with encoded CVs (generated from the multi-level performance evaluation for each bin which were then used for comparison to the reference analyser (ILC target result)).

6.2.3 Performance goals and ranking

Performance was monitored through use of quality goals, comprising total allowable error or TE_a as recommended by ASVCP (Nabity *et al.*, 2018) or internal expert opinion as recommended by an internal expert group of clinical pathologists and laboratory technicians, known as the Quality Education, Planning and Implementation Group (Daly *et al.* 2022). Desirable biologic variation-based goals for CV and bias (Fraser, 1992b; Fraser, 1992a; Flatland *et al.*, 2018) as well biologic variation goal for laboratory systems ($0.33CV_I$) (Fraser, 1992b) and sigma metrics (Westgard and Westgard, 2016) were also assessed. Sigma metric was used as a global indicator of performance, where measurands performing > 6 sigma are considered to demonstrate world class performance (Westgard, 2001; Westgard *et al.*,

2018). Previous optimisation demonstrated a 5.5 sigma cut off as a critical indicator of performance within a network of analysers (Daly *et al.*, 2021) (Daly *et al.*, 2022) and this cut-off was selected for this study.

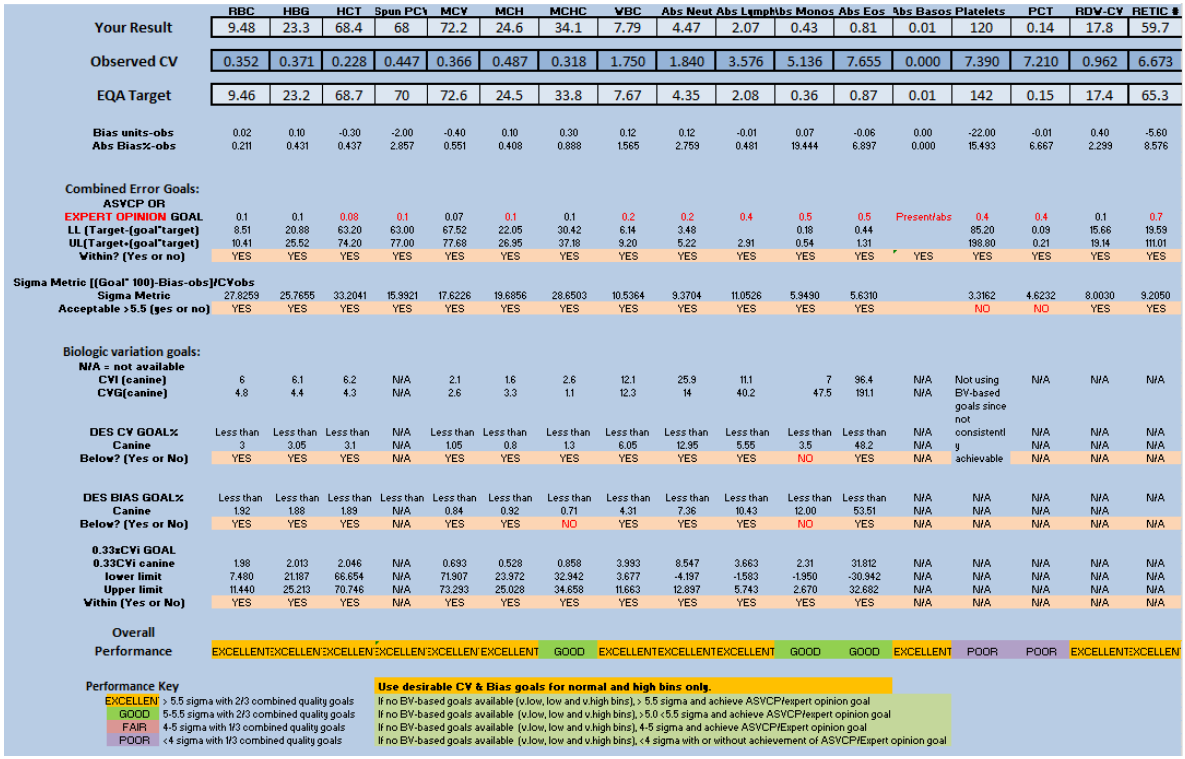
The ASVCP and/or expert opinion goals along with sigma metric were considered the 'combined quality goals' for our network of analysers and achievement of one or two goals for the level/bin being assessed was our "adequate performance" criterion in an ILC event. For the biologic variation-based goals for Bias, CV and laboratory systems goal (0.33 X CV_I), (Fraser, 1992b) we used 'desirable' biologic variation-based goals (Flatland *et al.*, 2018a) as the minimum level criteria for our laboratories. We used canine biologic variation-based data as canine samples are the most common species analysed in the laboratory network and the source of the patient samples used for ILC events. Biologic variation-based goals were not used for certain bins (very low, low, and very high HCT) as they are based on studies of healthy animals; instead for these circumstances, the measurand needed to meet criteria for the ASVCP or internal expert opinion goals and achieve a specified sigma metric value (Table 6.2). Clinical interpretation of results also was used to determine if "fair" or "poor" performance was likely to be of clinical significance and justified an action plan for improvement.

Table 6. 2: Performance criteria key for Internal Laboratory Comparative Testing assessment and RIQAS EQA assessment for the reference analyser.

Performance Key for Internal Laboratory Comparative Testing Events		
Performance Rank	**Rank Criteria**	**Additional Instructions**
Excellent	➢ 5.5 sigma and at least 2 of the 3 biologic variation-based quality goals	If no BV-based information available, > 5.5 sigma, with achievement of ASVCP/expert opinion goal
Good	➢ 5.0 < 5.5 sigma and at least 2 of the 3 biologic variation-based quality goals	If no BV-based information available, > 5.0 <5.5 sigma with achievement of ASVCP/expert opinion goal
Fair	4-5 sigma and at least 1 of the 3 biologic variation-based quality goals	If no BV-based information available, 4.0-5.0 sigma with achievement of ASVCP/expert opinion goal
Poor	< 4 sigma, with or without achievement of any of the 3 biologic variation-based goals	If no BV-based information available, < 4.0 sigma, with or without achievement of ASVCP/expert opinion goal

The performance criteria were encoded in an automated spreadsheet (Figure 6.2). Technicians recorded results directly into the spreadsheet (Figure 6.3) and could immediately see any performance issues that needed troubleshooting or escalation to the laboratory manager. This helped improve standardisation of technical quality comments and monitoring across the network. Any “poor” or “fair” performance following the spreadsheet assessment was reviewed by a pathologist (Figure 6.4). The quality goal index (QGI) was used to confirm performance issues due to bias and/or imprecision (Parry DM, 2009) following the return of the results.

Figure 6. 2: Automated spreadsheet showing result entry, quality goals and performance grading criteria.



	RBC	HBG	HCT	Spun PCV	MCV	MCH	MCHC	WBC	Abs Neut	Abs Lymph	Abs Monos	Abs Eos	Abs Basos	Platelets	PCT	RDW-CV	RETIC #
Your Result	9.48	23.3	68.4	68	72.2	24.6	34.1	7.79	4.47	2.07	0.43	0.81	0.01	120	0.14	17.8	59.7
Observed CV	0.352	0.371	0.228	0.447	0.366	0.487	0.318	1.750	1.840	3.576	5.136	7.655	0.000	7.390	7.210	0.962	6.673
EQA Target	9.46	23.2	68.7	70	72.6	24.5	33.8	7.67	4.35	2.08	0.36	0.87	0.01	142	0.15	17.4	65.3
Bias units-obs	0.02	0.10	-0.30	-2.00	-0.40	0.10	0.30	0.12	0.12	-0.01	0.07	-0.06	0.00	-22.00	-0.01	0.40	-5.60
Abs Bias%-obs	0.211	0.431	0.437	2.857	0.551	0.408	0.888	1.565	2.759	0.481	19.444	6.897	0.000	15.493	6.667	2.299	8.576
Combined Error Goals: ASVCP OR																	
EXPERT OPINION GOAL	0.1	0.1	0.09	0.1	0.07	0.1	0.1	0.2	0.2	0.4	0.5	0.5	Present/abs	0.4	0.4	0.1	0.7
LL (Target-(goal*target)	8.51	20.88	63.20	63.00	67.52	22.05	30.42	6.14	3.48		0.18	0.44		85.20	0.09	15.66	19.59
UL(Target+(goal*target)	10.41	25.52	74.20	77.00	77.68	26.95	37.18	9.20	5.22	2.91	0.54	1.31		198.80	0.21	19.14	111.01
Within? (Yes or no)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sigma Metric [(Goal* 100)-Bias-obs]/CVobs																	
Sigma Metric	27.8259	25.7655	33.2041	15.9921	17.6226	19.6856	28.6503	10.5364	9.3704	11.0526	5.9490	5.6310		3.3162	4.6232	8.0030	9.2050
Acceptable >5.5 (yes or no)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES		NO	NO	YES	YES
Biologic variation goals: N/A = not available																	
CVI (canine)	6	6.1	6.2	N/A	2.1	1.6	2.6	12.1	25.9	11.1	7	96.4	N/A	Not using	N/A	N/A	N/A
CVG(canine)	4.8	4.4	4.3	N/A	2.6	3.3	1.1	12.3	14	40.2	47.5	191.1	N/A	BV-based goals since not			
DES CV GOAL%	Less than	Less than	Less than	N/A	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	N/A	consistentl	N/A	N/A	N/A
Canine	3	3.05	3.1	N/A	1.05	0.8	1.3	6.05	12.95	5.55	3.5	48.2	N/A	y	N/A	N/A	N/A
Below? (Yes or No)	YES	YES	YES	N/A	YES	YES	YES	YES	YES	YES	NO	YES	N/A	achievable	N/A	N/A	N/A
DES BIAS GOAL%	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	Less than	N/A	N/A	N/A	N/A	N/A
Canine	1.92	1.88	1.89	N/A	0.84	0.92	0.71	4.31	7.36	10.43	12.00	53.51	N/A	N/A	N/A	N/A	
Below? (Yes or No)	YES	YES	YES	N/A	YES	YES	NO	YES	YES	YES	NO	YES	N/A	N/A	N/A	N/A	N/A
0.33xCVi GOAL																	
0.33CVi canine	1.98	2.013	2.046	N/A	0.693	0.528	0.858	3.993	8.547	3.663	2.31	31.812	N/A	N/A	N/A	N/A	N/A
lower limit	7.480	21.187	66.654	N/A	71.907	23.972	32.942	3.677	-4.197	-1.583	-1.950	-30.942	N/A	N/A	N/A	N/A	N/A
Upper limit	11.440	25.213	70.746	N/A	73.293	25.028	34.658	11.663	12.897	5.743	2.670	32.682	N/A	N/A	N/A	N/A	N/A
Within (Yes or No)	YES	YES	YES	N/A	YES	YES	YES	YES	YES	YES	YES	YES	N/A	N/A	N/A	N/A	N/A
Overall Performance	EXCELLENT	EXCELLEN	EXCELLEN	EXCELLEN	EXCELLEN	EXCELLENT	GOOD	EXCELLENT	EXCELLENT	EXCELLENT	GOOD	GOOD	EXCELLENT	POOR	POOR	EXCELLENT	EXCELLEN

Performance Key	
EXCELLEN	> 5.5 sigma with 2/3 combined quality goals
GOOD	5-5.5 sigma with 2/3 combined quality goals
FAIR	4-5 sigma with 1/3 combined quality goals
POOR	<4 sigma with 1/3 combined quality goals

Use desirable CV & Bias goals for normal and high bins only.
If no BV-based goals available (v.low, low and v.high bins), > 5.5 sigma and achieve ASVCP/expert opinion goal
If no BV-based goals available (v.low, low and v.high bins), > 5.0 <5.5 sigma and achieve ASVCP/expert opinion goal
If no BV-based goals available (v.low, low and v.high bins), 4-5 sigma and achieve ASVCP/Expert opinion goal
If no BV-based goals available (v.low, low and v.high bins), <4 sigma with or without achievement of ASVCP/Expert opinion goal

Figure 6. 3: Automated spreadsheet showing area on spreadsheet for technical comments.

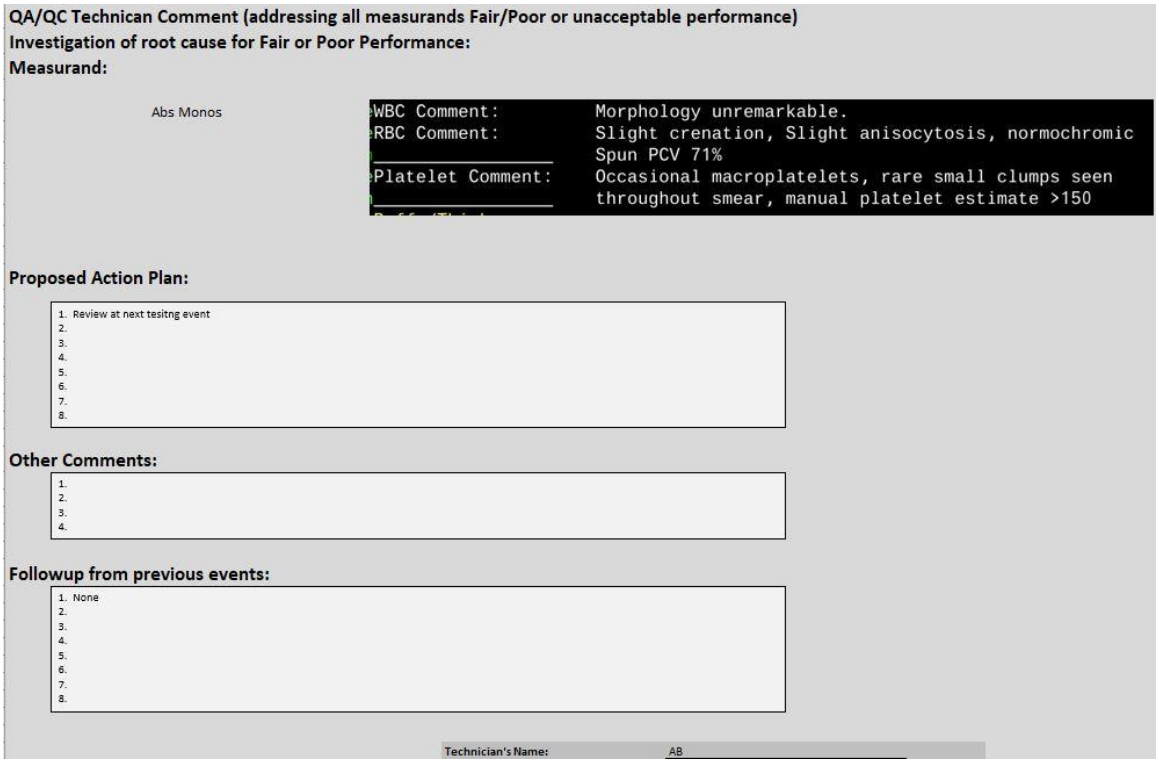
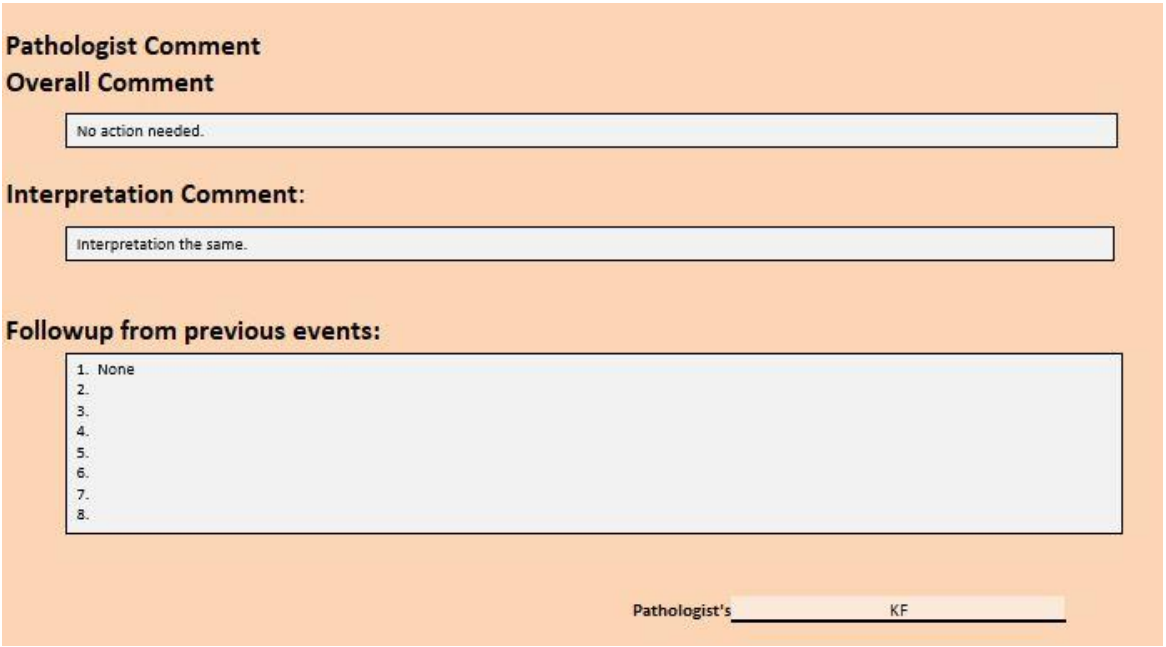


Figure 6. 4: Automated spreadsheet showing area for Pathologist comment.



6.2.4 Reference analyser EQA performance

The reference analyser remained subscribed to an EQA program (RIQAS) (Randox) where a sample was received once a month by the laboratory and processed according to the program instructions. Upon receipt of EQA reports, the EQA target results were entered into an automated EQA spreadsheet along with the results from the reference analyser. The overall performance was assessed using the performance goals described and the performance determination was automatically graded on the spreadsheet as per the performance key, see Table 6.2. Where possible, root cause was established for any performance related issues (“fair” or “poor” performance identified). Technicians were required to make any additional notes on the performance in the comments section of the spreadsheet prior to sending to the designated clinical pathologist for review and comment.

6.3 Results

6.3.1 Multi-level performance evaluation

Mean, SDs and CVs were utilised from the multi-level, 10-replicate dataset to derive the necessary observed performance information underlying the evaluations in the single analysis interlaboratory comparison program. No

analyser displayed excessive imprecision for the low, normal, or high bins using the desirable biologic variation-based goal for CV (desirable CV). The data established expectations for precision for each analyser for each bin that corresponded to a clinical decision level for haematocrit. However, reticulocyte counts yielded high CVs for many analysers and analysers frequently demonstrated imprecision for all the differential leukocyte measurands (supplementary material available, Appendix E, Tables E6.1 - E6.3).

6.3.2 Single level evaluations

The ILC events were carried out every 6-8 weeks until the 5 bins were complete with each event focussed on a different bin. The erythroid measurand performance using the performance key was consistently “excellent” to “good” across the testing events (see supplementary material, Figure F6.1 A-E) except for RDW-CV for the low, high, and very high ILC event and RETIC number for very low, low, and very high ILC events which had “fair” to “poor” performance on multiple occasions. This poor performance was due to excessive bias, as determined using the QGI, (Parry DM, 2009) with sigma metrics consistently < 5.5. The very low ILC event indicated a performance issue for the reference analyser, which was likely due to random error. Overall, for the 5 ILC events for the 5 analysers compared to the reference analyser, RBC, MCH, and MCHC only flagged

once, and this was most frequently due to bias, as these measurands failed to meet the desirable biologic variation-based goals for bias (Table 6.3).

Table 6. 3: Measurands with poor performance for the erythroid measurands for the network analysers using quality goals of TE_a (ASVCP/expert opinion), sigma metric, Desirable CV (biologic variation), Desirable Bias (biologic variation), clinical interpretation and quality goal index (QGI).

Measurand	Level/Bin	Analyser	Result	Target	Achieved ASVCP/Expert opinion TE_a (Yes/No)	Sigma metric	Achieved BV CV goal	Achieved BV bias goal	Achieved 0.33 CV_I goal	Clinical interpretation	QGI
RBC ($10^{\wedge 12}$/l)	Very high	5	7.72	9.56	**No**	< 3	**Yes**	**No**	**Yes**	Same	Bias
MCH (pg)	Very low	4a	24.4	23.4	**Yes**	4.04	**No**	**No**	**Yes**	Same	Bias
	Low	4a	28.1	26.6	**Yes**	3.08	**No**	**No**	**Yes**	Same	Bias
MCHC (g/dl)	Very low	4a	32.6	30.1	**Yes**	< 3	**No**	**No**	**Yes**	Same	Bias
	Low	4a	4.98	5.19	**Yes**	4.06	**No**	**No**	**Yes**	Same	Bias
RDW-CV (%)	Low	4a	0.021	0.085	**No**	< 3				Same	Bias
		4b	0.027	0.085	**No**	< 3				Same	Bias
		3	12.1	17.4	**No**	< 3				Same	Bias
	High	4b	18.7	17.4	**Yes**	5.2				Same	Bias
		5	12.1	17.4	**No**	< 3				Same	Bias
RET# (10^9/l)		2	15.9	9.5	**Yes**	< 3	***No biologic variation data available**				
	Very low	3	14	9.5	**Yes**	< 3				Likely random error for reference analyser	
		4a	15.1	9.5	**Yes**	< 3					
		5	15.1	9.5	**Yes**	< 3					
	Low	3	80	85.5	**Yes**	4.64				Same	Bias
		4a	21	85.5	**No**	< 3				Same	Bias
		4b	27	85.5	**Yes**	< 3				Same	Bias

***Biologic variation data not applied** to the very low, low and very high bins due to extreme SDs and CVs for these analyses

However, when the clinical interpretation of the result was the same as the target result (reference analyser), the “fair” to “poor” performance of the measurands was considered unlikely to be of any clinical significance and action plans were rarely necessary. “Excellent” to “good” performance criteria were met in more than > 95% of events for RBC, HGB, HCT, MCH and MCHC across the ILC program whereas more variable performance was seen for RDW-CV and RETIC which had “excellent” to “good” performance, including the latter measurands, > 90% of the erythroid measurands had “excellent” to “good” performance across the analytical network.

Leukocytes performed “good” to “fair” with occasional “poor” performance compared to the reference analyser. Notably, “poor”, or “fair” performance was seen more frequently for MONO, EOSO and BASO, due to bias and sigma metric < 5.5. On closer evaluation of these events, the clinical interpretation was the same and, therefore no action plan was needed (Table 6.4).

Table 6. 4: Fair or poor performance for the leukocyte measurands for the network analysers, using quality goals of TE_a (ASVCP/expert opinion), sigma metric, Desirable CV (biologic variation), Desirable Bias (biologic variation), clinical interpretation and quality goal index (QGI).

Level	Analyser	Result	Target	Achieved ASVCP/Expert opinion TE_a	Sigma metric	Achieved BV CV goal	Achieved BV bias goal	Achieved 0.33 CV_i goal	Clinical Interpretation	QGI
WBC (10^3/µl)										
Very low	2	1.8	1.64	Yes	5.21	**Biologic variation data not applied**			Same	Bias
Low	4a	3.04	3.52	Yes	3.5				Same	Bias
	4b	3.04	3.52	Yes	4.98				Same	Bias
NEUT# (10^9/l)										
Very low	4a	0.91	1	Yes	3.7	**Biologic variation data not applied**			Same	Bias
Low	3	3.1	3.52	Yes	2.86				Same	Bias
	4a	3.04	3.52	Yes	4				Same	Bias
	4b	3.04	3.52	Yes	5.08				Same	Bias
LYMPH# (10^9/l)										
Normal	3	1.5	1.59	Yes	4.7	**Yes**	**Yes**	**Yes**	Same	Bias
	5	1.94	1.59	Yes	4.78	**Yes**	**No**	**Yes**	Same	Bias
MONO# (10^9/l)										
Normal	3	1	0.9	Yes	< 3	**No**	**Yes**	**Yes**	Same	Bias
	4a	1.1	0.9	Yes	3.07	**No**	**No**	**Yes**	Same	Bias
	4b	1.17	0.9	Yes	< 3	**No**	**No**	**Yes**	Same	Bias
High	4a	0.47	0.36	Yes	< 3	**No**	**No**	**Yes**	Same	Bias
	4b	0.53	0.36	Yes	< 3	**No**	**No**	**Yes**	Same	Bias
	5	0.4	0.36	Yes	< 3	**No**	**Yes**	**Yes**	Same	Bias
Very high	4a	0.89	0.74	Yes	4.25	**Biologic variation data not applied**			Same	Bias
	4b	0.83	0.74	Yes	5.4				Same	Bias
	5	0.64	0.74	Yes	< 3				Same	Bias

*__Biologic variation data not applied__ to the very low, low and very high bins due to extreme SDs and CVs for these analyses

Platelet and plateletcrit performance were “fair” to “poor” on several occasions due to high observed bias and sigma metric < 5.5 (Table 6.5). The QGI also indicated bias was the most likely cause. There were two occasions where interpretation between the reference analyser and the network analyser may not have been the same for the very low and low single analysis ILC event. Harmonisation was not maintained for platelets for the very low and low ILC event for analyser 2 and for the low ILC event for analyser 3, however, harmonisation was maintained for all other measurands during this ILC program (see supplementary figures available in Appendix F, Figures F6.1 A-E, F6.2 A-E, F6.3 A-E).

Table 6. 5: Fair and poor performance summary for platelets for the network analysers, using quality goals of TE_a (ASVCP/expert opinion), sigma metric, clinical interpretation and quality goal index (QGI)

Level/Bin	Analyser	Result(10^9/l)	Target	Achieved ASVCP/Expert opinion TE_a?	Sigma metric	Clinical Interpretation	QGI interpretation
Very low	2	119	60	No	< 3	Not the same	Bias
	3	74	60	Yes	3.54	Same	Bias
	4a	71	60	Yes	4.72	Same	Bias
	4b	71	60	Yes	3.22	Same	Bias
	5	70	60	Yes	5.25	Same	Bias
Low	2	147	113	Yes	5.29	Not the same	Bias
	3	149	113	Yes	2.98	Not the same	Bias
	4a	116	113	Yes	3.9	Same	Bias
Normal	2	149	173	Yes	3.55	Same	Bias
High	2	120	142	Yes	3.31	Same	Bias
Very high	2	165	142	Yes	3.22	Same	Bias
	3	173	142	Yes	3.91	Same	Bias

6.3.3 Reference analyser EQA performance

The performance of the reference analyser compared to the RIQAS EQA (Randox) showed that the erythroid measurands performance was excellent according to the performance criteria (e.g., consistently meeting the recommended TE_a goals as recommending by ASVCP or by expert opinion, > 5.5 sigma and achieving desirable biologic variation-based goals for bias and CV). Only one event showed poor performance of HCT, and this was directly before analyser servicing was due.

Due to the limitations of the RIQAS EQA program to compare the WBC differential, only the WBC count was compared, and performance was “excellent” to “good” except for two occasions where performance was poor highlighted by sigma metrics < 3 and inability to achieve the desirable bias goal. Interestingly, both events occurred within a month before the bi-annual service/preventative maintenance date for the analyser (supplementary figures available in Appendix F, Figures F6.4 and F6.5).

Platelets and plateletcrit remained “good” to “fair” over the 12-month period according to the spreadsheet performance criteria as sigma metrics were below < 5.5. Four events were due to higher CVs and observed bias (%); however, this did not prompt further investigation as interpretation of the EQA target and the reference analyser results was the same.

6.3.4 Performance goals and sigma metrics

TE_a as recommended by ASVCP was not always achieved particularly for the extreme levels of clinical interest (very low, low, and very high) and, when this was observed, alternative TE_a that would not change the clinical interpretation were proposed and implemented. An example provided by the very low WBC bin convey the approach taken. At 2,000 cells/µL level +/- 15% TE_a as recommended by ASVCP was not achievable but adjusting to 20% would mean only a difference of 400 cells/µL (20% of 2000) which in clinical terms does not change the interpretation. Table 6.6 shows the proposed internal expert opinion goals based on the observed TE (TE_{obs}).

Table 6. 6: TE quality goals for internal laboratory comparison testing for very low, low, normal, high and very high haematocrit bins (%).

	Very low	Low	Normal		High	Very high
Measurand	**TE_{rec}**	**TE_{rec}**	**ASVCP**	**Revised TE_{rec}**	**TE_{rec}**	**TE_{rec}**
RBC (10^{12}/l)	25	20	**10**	10	10	10
HGB (g/dl)	25	20	**10**	10	10	10
HCT (%)	25	20	**10**	10	8	7
MCV (fl)	7	7	**7**	7	7	7
MCH (pg)	10	10	**n/a**	10	10	10
MCHC (g/dl)	10	10	**10**	10	10	10
RDW-CV	25	20	**10**	10	10	10
RETIC#	70	70	**20**	40	70	70
WBC (10^3/µl)	25	25	**15**	20	20	20
NEUT# (10^9/l)	25	25	**15**	20	20	20
LYMPH# (10^9/l)	80	80	**15**	40	40	30
MONO# (10^9/l)	150	150	**60**	50	50	50
EO# (10^9/l)	100	100	**50**	50	50	50
BASO# (10^9/l)	n/a	n/a	**n/a**	n/a	n/a	n/a
PLT (10^9/l)	40	40	**20**	40	40	40
PCT (%)	40	40	**n/a**	40	40	40

TE_{rec} = recommended allowable TE

Desirable biologic variation-based goals for bias were not achieved for the erythroid measurements with poor performance (RBC, HCT, MCHC and MCHC); however, as the clinical interpretation was the same for these measurands (Table 6.3), action plans were not necessary. For leukocytes, the biologic variation-based bias goal where applicable (normal and high events) was not achieved for LYMPH and MONO for the normal and high ILC events for analysers 4a, 4b and 5. This was due to observed bias which was 22% for LYMPH and ranged from 22 - 47% for MONO (Table 6.4).

Desirable biologic variation-based goals for CV were not achieved for erythroid measurands with poor performance except for RBC and HCT for the very high ILC event which did meet this goal. For MHC and MCHC for the very low and low event for analyser 4a, poor performance, although due to bias, also did not achieve the biologic variation-based CV goal, CVs were 1.41% and 1.64%, respectively. For leukocytes, this goal (applicable to normal and high) was not achieved for MONO for the normal and high event for analysers 3, 4a, 4b and 5 (Table 6.4). CVs ranged from 9 - 20% for the normal event and 7 - 23% for the high event, as shown in The biologic variation-based goal for laboratory systems ($0.33CV_I$) where applicable was achieved for all “poor” or “fair” performances.

Sigma metrics were consistently < 5.5 for measurands in the “poor” or “fair” performance rank and were monitored at the next ILC event and for any particularly low sigma (< 4) it was recommended to that site to consult with technical support for that analyser.

6.4 Discussion

We successfully implemented an interlaboratory comparison testing program for a network of haematology analysers that assessed and monitored bias, maintained a harmonised status and continued to use common reference intervals. We reduced costs by using modified common canine samples over a range of clinical interest levels, which, to our knowledge, has not previously been implemented for a network of veterinary haematology analysers. The laboratory staff at each site in the network provided the necessary data using commutable samples to ensure each analyser within the network was comparable to the reference analyser and that results could be used interchangeably.

For harmonisation to be clinically useful it was essential that harmonisation occurred across all clinically relevant levels of results and not just for a single level; therefore, the generation of the bins in the multi-level performance evaluation for the levels of clinical interest was an important first step in this testing program. The easiest and most efficient way to achieve this was to aim for haematocrit percentage with the corresponding dilution or concentration of RBCs, HGB, and WBCs, to generate CVs which would allow progression to the single analysis ILC testing program. Harmonisation may be more easily achieved for a single normal level, but this was not sufficient to allow for a harmonised status for our network which produces haematology results in patient samples across a range of clinical important results. CVs were necessary to assess the TE_a that could be achieved and to calculate sigma metrics. The CVs generated from each analyser were similar,

particularly for the erythroid measurands, but excluding reticulocytes, which showed more extreme variation. In a previous study using commercially available quality control material, observed CVs were also notably higher for reticulocytes (Daly *et al.*, 2022). A previous study also recommended improvements in instrument performance and better-defined decision limits particularly for reticulocytes (Nabity *et al.*, 2018). Higher CVs were noted for the extreme samples for leukocytes, platelets and plateletcrit but when the absolute numbers were assessed particularly for monocytes, eosinophils, and basophils, the analyser counts thousands of cells (many more than a 100-cell manual differential count); therefore this is more likely a better indicator of performance for these measurands. We found that the ASVCP-recommended TE_a was not always achieved for the very low, low, and very high levels due to higher CVs. Derivation of the ASVCP TE_a goals was based primarily on QCM which have different CVs (Nabity *et al.*, 2018). Previously reported CVs on Sysmex analysers can also be found for normal QCM (Daly *et al.*, 2022). Extreme samples should be assessed independently for their achievable error and this supports our previous recommendations for individual analyser evaluation within a network (Daly *et al.*, 2022) even if the model of analysers are the same. The results of this study indicate that, it could be useful for the ASVCP Quality Assurance and Laboratory Standards Committee to revise the recommended ASVCP TE_a goals considering this new information regarding canine samples at various levels of clinical interest.

The single analysis ILC events demonstrated varying performance for the analyser network compared to the reference analyser. However,

harmonisation was successfully maintained except for only 2 platelet results that would have generated discordant interpretation. Substandard performance (“poor” to “fair”) compared to the reference analyser was seen more frequently for the extreme samples and less frequently for the normal ILC event. Performance for the erythroid measurands was the most consistent and had the highest rate of “excellent” to “good” performance overall. The leukocytes and platelets were identified with “fair” or “poor” performance more frequently on the spreadsheet but on closer inspection rarely required action as clinical interpretation would not have changed for the individual network analyser compared to the reference analyser. Eosinophils had the most occurrence of “fair” to “poor” results, particularly for the very low and very high ILC, followed by plateletcrit and platelets, particularly for the very low ILC. Only two platelet-related results out of the entire ILC program and all measurands were interpreted as not the same; this occurred for platelets for the very low and the low single ILC event for the same analyser. On closer inspection, the observed bias for this analyser was higher than the other analysers in the network and because the difference in results would change the interpretation, the peripheral blood smear evaluation for platelets would be a requirement on future evaluations. The proposed action plan was to review this analyser at the next testing event, and it was recommended that a repeat evaluation of SD and CVs for this analyser should be performed. Bias was identified as the culprit for any “fair” to “poor” performance and supported when assessed with the QGI. As inaccuracy is real-time issue for commercial laboratories, the single analysis ILC program is a useful way to monitor bias over time and over various

clinical decision levels. The frequency of the monitoring can be controlled by the laboratory as it may be prudent to review this more frequently if a problem is identified, rather than waiting on the next commercial EQA sample to arrive and the subsequent wait for the report that follows.

The biologic variation-based goals for desirable CV and bias were a useful addition to monitoring performance and were not always achieved (Tables 6.3 – 6.5). The desirable rather than minimally accepted goal are considered possibly too stringent for veterinary laboratories (Nabity *et al.*, 2018) and therefore this should be assessed based on individual instrument performance. However, for our network of analysers, we found the desirable limit discriminatory and a useful indicator of performance when used for the normal and high levels. The laboratory systems goal ($0.33xCV_I$) was included as a useful goal for a network of laboratories, however, proved non-discriminatory (never identifying abnormality) in the ILC testing program. Therefore, its usefulness could be debated, but we chose to include this due to its relevance for a laboratory system and whether it will continue to be used on future ILC events has yet to be decided. As we did not use biologic variation-based goals for the extreme samples, sigma metric was a constant placeholder of performance. This allowed modifications to the performance key so that those measurands without biologic variation assessment would be ranked according to sigma metrics in addition to TE_a or internal expert opinion. A proposed sigma metric cut off for veterinary haematology analysers of 5.5 (Daly *et al.*, 2022) was included on the spreadsheet as measurands below this cut off were monitored closely following identification and prior to the next ILC event.

The use of the automated spreadsheet for the return of results was valuable. The fact the individual analysers SDs and CVs were hardcoded and automatically populated based on the HCT meant the evaluation of five separate spreadsheets for each of the five ILC events over the duration of the study was very efficient, the user of the spreadsheet was only required to input the results manually (which is what occurs with commercial EQA programs) simplifying the process. The performance grading for each measurand was made easily identifiable by automated colour corresponding to each classification.

The challenges to implementing this program were mainly surrounding logistics. The use of a third-party courier to each site on the same day was for the most part reliable, but there were two delayed occasions for one site. This was noted in the results and on review did not have an impact on performance. One delayed event occurred during warmer weather, and this was reflected in the RBC result for that ILC event. Future ILC programs will now avoid extremes of temperatures in Summer and Winter months to minimise any possible sample compromise due to temperature fluctuations. The benefits of the ILC testing program supersedes any logistical issues as these could potentially occur when using multiple subscriptions to a commercial EQA program and the latter is a far more expensive option. For example, an EQA option for once per month costs approximately 800 GBP. Within the UK, the cost of overnight postal delivery is less than 10 GBP. For six events per year this is less than 60 GBP per laboratory. Extra cost is incurred at the reference laboratory in technical time for preparing the samples and arranging for overnight delivery (estimated 1.5 - 2 hours per

event). The cost of technical and pathologist analysis of EQA and ILC events is similar, although dependent on the frequency of sample distribution; the EQA scheme used in this study released twice as many events per year as the ILC. The ILC option represents considerable cost savings since only the reference laboratory continues traditional EQA, rather than the total five separate EQA subscriptions which costs approximately £3200 annually for the five laboratories within the network.

6.5 Conclusion

An ILC testing program was successfully implemented for a network of veterinary haematology analysers using a species-specific sample relevant to the throughput of the laboratory. Samples could be manipulated to achieve various levels/bins across the range of clinical interest and the generation of individual analyser CVs was integral to progressing to the single ILC testing program. The single ILC events recorded performance based on defined grading criteria and any poor performance was easily identified using pre-determined combined quality goals and sigma metrics. Most cases of “fair” or “poor” performance were due to bias but were not clinically relevant based on the clinical interpretation. Therefore, patient results were not misclassified across the network. The ILC testing program demonstrated harmonisation could be maintained within a network of veterinary analysers using a combination of TE_a, biologic variation-based goals for desirable bias and CV, and a > 5.5 sigma strategy for the ILC samples used across the proposed

HCT ‘bins’. This combination of quality goals lends itself to a network state-of-the art structure for quality management for veterinary laboratories and a total quality management system. Significant advantages of the ILC program include ability to assess performance across a wide range of clinically important results (bins) and cost savings compared to subscription to traditional EQA programs. Future avenues for investigation include refinement of evaluation criteria used on the ILC spreadsheet.

CHAPTER 7: COMPARATIVE TESTING FOR IN-CLINIC POINT OF CARE HAEMATOLOGY ANALYSERS WITH A COMMERCIAL LABORATORY REFERENCE ANALYSER

This final stage of the research and continuation of chapter 6 outlines a pilot program for extending interlaboratory comparison and harmonisation to veterinary practices utilising point of care haematology analysers by comparing them to the network reference analyser. This chapter aims to evaluate the performance of the in-clinic POC analysers using the quality goals used throughout this research and to determine if the data can be used to harmonise the results of the in-clinic POC analysers with those of the reference laboratory and whether clinic-based clinicians and technicians found the information provided valuable for ongoing in-clinic haematology quality management. This research was led by S Daly and in collaboration with Dr Peter Graham and Dr Kathy Freeman as coauthors on the manuscript produced from this research and represents original work.

Adaptations: This chapter has been changed from its previous version for publishing to include English spelling, removal of some tables and/or figures to appendices. Citations style has also changed to fit with thesis structure and references are now consolidated at the end of this thesis.

This chapter in currently in review with journal of *Veterinary Clinical Pathology.*

7.1 Introduction

In-clinic laboratory equipment is often justified in rural practices where courier pickup is not available and in other practices as part of a comprehensive package including reference laboratory services. In-clinic, point of care (POC) testing is available for many haematologic measurands and a complete blood count (CBC) with a white blood cell (WBC) differential (diff) is estimated to be the most commonly ordered test for veterinary medicine (Whitbread, 2019). The need for rapid haematology results is not to be disputed, particularly for life threatening situations. For practices in rural areas without rapid access to a reference laboratory, POCT is powerful tool. However it's reliability is dependent on using sufficient quality assurance protocols (Flatland *et al.*, 2010; Lester *et al.*, 2013) to ensure that the patient results are accurate, reliable and reproducible. Inaccuracy (Bias) and imprecision for POC analysers have been previously reported (Rishniw *et al.*, 2012; Baral *et al.*, 2015; Becker *et al.*, 2009) and if analytical/analyser performance is not managed they may contribute to misclassification of patient samples.

Quality management for laboratory testing has traditionally been conducted using commercial quality control materials (QCM) and by subscribing to an external quality assurance (EQA) program to monitor individual analyser

performance in comparison to a peer group using the same analyser and/or method. However, QCM and EQA samples are typically of human origin and, therefore, do not share the same components or matrix as veterinary specimens (Unsal *et al.*, 2010). Comparative testing of patient samples by in-clinic laboratories and their reference laboratory is acceptable for EQA by several accreditation programs and expert organisations (ASVCP) (AAHA) (RCVS). Additional quality performance monitoring can include the use of quality goals such as total allowable error (TE_a), as recommended by ASVCP, (Nabity *et al.*, 2018) or by expert opinion which has been previously reported for a network of veterinary laboratories (Daly *et al.*, 2022). Sigma metrics are also a useful indicator of analytical performance whereby a sigma (σ) metric result ≥ 6 for an analytical method indicates 'world class' performance, reflecting ≤ 3.4 defects (erroneous results) per million results (Westgard, 2006). Sigma metrics have been utilised as a performance tool for veterinary laboratories and as a risk measure for unstable system performance (Daly *et al.*, 2021) (Daly *et al.*, 2022). Other aspects of quality management for haematology testing, whether in the in-clinic or reference laboratory, include: monitoring of refrigerator temperatures for storage of QCM and reagents; documentation of routine maintenance of the analysers according to the manufacturer's recommendations and repairs that may result in a need for increased scrutiny for potential changes in instrument performance; documented training and competency testing of operators of the instruments; and correlation of the microscopic evaluation of blood smears with automated results (Nabity *et al.*, 2018; Flatland *et al.*, 2013).

Furthermore, practices utilising a reference laboratory as well as in-clinic POC analysers for haematology testing should ensure that results between the systems, which may have different methods, produce results that are interchangeable for making medical decisions regarding patient classification, monitoring and/or response to treatment. This allows the clinician to use either the POC instrument or the reference laboratory instrument for monitoring of patients without concern for excessive bias that may result in misclassification of changes in results produced by different analysers/methods. This is defined as 'harmonisation' and this has only been reported for veterinary laboratories recently (Daly *et al.*, 2021). Harmonisation relies on standardisation of results across laboratories and/or methods when a certified standard material is not available (Miller *et al.*, 2011a). Harmonisation of POC results to the reference laboratory used by the clinic provides confidence in the interpretation of results produced by either the POC analyser or the reference laboratory analyser. By determining POC analyser bias through a system of external quality assessment, adjustments can be made to POC analyser results that allow for in-clinic/reference laboratory harmonisation.

To our knowledge, there are no published studies in which in-clinic analysers are compared to a veterinary reference analyser for external quality assessment. Therefore, this study aims to present an initial evaluation of a comparative testing program for veterinary haematology that could be used to harmonise results between the in-clinic POC haematology analyser and that of the reference laboratory used by the practice.

Aims:

1. To design a comparative testing program for use by in-clinic laboratories with a variety of POC analysers.
2. To evaluate the performance of in-clinic POC haematology analysers compared to a reference analyser.
3. To determine if the data can be used to harmonise the results of the in-clinic POC analysers for normal canine samples with those of the reference laboratory so that results are interchangeable and useful for initial diagnosis and ongoing monitoring.
4. To determine if the clinicians/technicians found the information provided valuable for ongoing in-clinic haematology quality management.

7.2 Materials and Methods

7.2.1 Study design for in-clinic and reference analyser comparative testing

7.2.1.1 In-clinic POC analysers

Seven veterinary clinics that were clients of the primary author's commercial laboratory were invited to participate in this study. All samples required processing within a single day; therefore, this study was limited to a small pool of practices within a courier collection area. Instructions were given via

email and with a follow up telephone call. They were instructed to process a 'normal' or within-reference-limit fresh EDTA whole blood canine sample on their in-clinic analyser which sufficient sample volume for analysing 5 times before submitting to the reference laboratory for analysis a further 10 times on the same day (minimum EDTA sample volume of 2000µL). The in-clinic haematology comparative testing process was designed to make it easy for the practice laboratory personnel to understand the requirements. Therefore, they were provided with an excel spreadsheet Microsoft Excel (Microsoft Excel for Mac 2011, Version 14.7.7 (170905) Last update 14.7.7) with areas to identify the in-clinic laboratory and practice, the species being evaluated to ensure canine specimens were used, the animal identification, date of processing, and the 5 sets of results produced by the in-clinic laboratory. The spreadsheet was emailed back to the reference laboratory and the sample was transported to the reference laboratory for same-day analysis.
The in-clinic POC analysers used in this program included EXIGO EosVet haematology analyser (EXIGO – Sweden(1)), IDEXX Procyte (1) (IDEXX laboratories Inc., Westbrook, ME, USA), IDEXX Lasercyte (1) (IDEXX), Woodley Insight H5 (Woodley Equipment Company Ltd, UK), Abaxis VetScan HM2 (1), (Abaxis North America, Union City, CA, USA), Abaxis VetScan HM5 (2), (Abaxis North America, Union City, CA, USA).

7.2.1.2 Reference Laboratory

The sample had to be received by the reference laboratory on the same day it was analysed by the in-clinic laboratory for inclusion in this study. Once

received, it was processed 10 times using a Sysmex XT-2000*i*V analyser (Sysmex corporation, Kobe, Japan). This analyser previously had a performance evaluation (Daly *et al.*, 2021) and quality control validation (Daly *et al.*, 2022) and was the reference analyser for the harmonisation of a network of veterinary laboratories (chapter 6). We choose to include the following measurands; Red blood cell count (RBC), Haemoglobin (HGB), Haematocrit (HCT), Mean corpuscular volume (MCV), Mean cell haemoglobin (MCH), Mean corpuscular haemoglobin concentration (MCHC), White blood cell count (WBC), Platelet count (PLT), Red cell distribution width-coefficient of variation (RDW-CV), Plateletcrit (PCT) and MPV because these were commonly available from in-clinic haematology analysers. The reference analyser was subscribed to an EQA monthly program (RIQAS) with excellent ongoing EQA performance (Randox). In the reference laboratory, samples were processed in accordance with manufacturers operating procedures and according to the laboratory SOP by a trained haematology technician. The results were entered into the excel spreadsheet submitted by the in-clinic laboratory for a quality assessment evaluation.

7.2.2 Performance evaluation

The observed mean, standard deviation (SD) and coefficient of variation (CV) (Jensen and Kjelgaard-Hansen, 2006) were calculated for all 10 measurands for the reference laboratory analyser (target value) and the in-clinic analyser. The mean of the results produced by the reference laboratory was used as the target mean to determine the bias (in units) and absolute bias (%) of the

in-clinic analyser compared to the reference analyser (Westgard, 2016a). The quality goals provided on the spreadsheet included an ASVCP or expert opinion goal for total allowable error (TE_a) and a sigma metric ≥ 4 as a starting point for evaluation.

Total error (TE) was calculated using the following formula:

$$TE\% = |Bias\ \%| + 1.96 \times CV\%$$ (Nabity *et al.*, 2018)

and compared against Total allowable error (TE_a) as recommend by ASVCP or an internal expert opinion quality goal recommended by an internal expert group of clinical pathologists and laboratory technicians from the reference laboratory, known as the Quality Education, Planning and Implementation Group.

The CVs were used to calculate sigma metrics which were used to evaluate performance for each measurand using the following formula:

$$\text{Sigma metric} = (TEa\% - |Bias\%|)/CV\%$$ (Westgard, 2016a)

Any sigma metric < 4 was assessed for whether this was due to excessive bias and/or imprecision using a quality goal index (QGI) which was determined using the following formula:

$$QGI = bias\%/1.5*CV$$

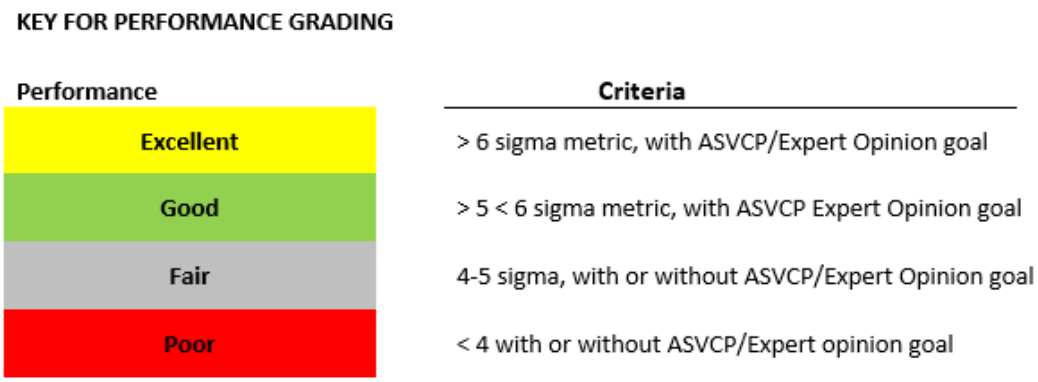
A QGI < 0.8 is indicative of imprecision, QGI > 1.2 indicates inaccuracy (bias) while a QGI of 0.8-1.2 indicate both bias and imprecision contributes to the poor performance (Parry DM, 2009).

Measurand bias (%) was deemed “high” when it exceeded 33% of the ASVCP recommended TE_a and CV’s (%) were deemed “high” they exceeded 25% of the ASVCP recommended TE_a (Westgard and Barry, 2008).

7.2.3 Performance grading key

Performance was graded according to the performance key grading system where specific criteria for “excellent”, “good”, “fair” and “poor” were applied to measurands based on sigma metric result and achievement of ASVCP recommended TE_a or expert opinion goal evaluation (Figure 7.1).

Figure 7. 1: The performance grading key and corresponding criteria to achieve performance grades; “excellent”, “good”, “fair”, “poor”.



We provided comments on each evaluation, regardless of level of performance, but only requested an action plan for “fair” and “poor” performance measurands.

To facilitate future comparison evaluations, a points system was developed to generate a standardised nominal value based on the performance grading, however this was not presented to the clinicians in this study. RBC, HGB, HCT and WBC were weighted as they are recognised as measurands that are particularly relied upon for clinical decision making in routine small animal veterinary practice.

The overall point score was calculated as follows:

Overall points = (2 X points for 4 weighted measurands) + (1 X points for 6 unweighted measurands).

The score was normalised by dividing by the number of measurands evaluated for each POC analyser.

7.2.4 In-clinic POC analyser result harmonisation

To assess the difference between the data sets, the observed mean difference between the results for the in-clinic POC analysers ($method_{poc}$) and the reference laboratory ($method_{ref}$) were evaluated and expressed in test units using the following formula:

$$Diff_{units} = Method_{poc} – Method_{ref}$$

Agreement limits were established based on the combined analytical variation for both the in-clinic POC and the reference analyser using the combined inherent imprecision and its associated 95% confidence limits (limits of agreement based on the average of the means) calculated using the following formula:

> Combined inherent imprecision = within +/- 1.96 x sqrt(CV_a-in-clinic^2 + CV_a-reference lab^2) x result of the reference laboratory (Jensen and Kjelgaard-Hansen, 2006)

Results falling within the limits for combined inherent imprecision were considered to represent 'the same' (interchangeable) result (Jensen and Kjelgaard-Hansen, 2006).

For any in-clinic POC results that were not within the combined inherent imprecision, a correction factor was applied to harmonise the result to the reference analyser by adding/subtracting the bias (units) from the in-clinic POC results to make the POC and reference analyser result the same (zero bias).

7.2.5 Technician/Clinician feedback

Each quality assessment report sent to the clinics included a section for technical comments and pathologist comments. These were aimed at educating the POCT laboratory about the significance of the results, possible

need for quality improvement, and advice regarding next steps in troubleshooting any problematic assays, including advice to consult with the technical support team of the manufacturer for their instrument.

The reference laboratory technicians were consulted through frequent meetings to give feedback on the process and time allocation for processing the comparative sample.

Follow up email and telephone calls were made to the POCT laboratory clinicians to determine the usefulness of the comparative testing event and time allocation for participating in this study.

7.3 Results

7.3.1 Study design

Six of the seven participants followed the instructions and selected and analysed the sample according to the criteria. One clinician did not process the samples 5 times but instead only 4 times due to concerning internal QC readings.

7.3.2 Performance evaluation

7.3.2.1 Bias and CV

In comparison to the Sysmex XT 2000*i*V reference analyser, the mean observed absolute bias (%) across the measurands analysed on the POC analysers ranged from 2.73 – 24.7%. The bias (%) was high for WBC for 6 of the 7 POC analysers, PLT for 6 of the 7 POC analysers, RDW-CV for 3 of the 5 analysers (2 did not return this result) and MPV 6 of the 7 POC analysers. High bias was noted for HCT and MCHC for two POC analysers and MCH for one POC analyser (Table 7.1).

Table 7. 1: Presents the observed mean, observed bias, and observed CV (%) generated by the individual in-clinic POC haematology analysers that participated in the comparative testing program using different canine samples that were chosen within the manufacturer's provided POC reference interval before sending to the reference laboratory for comparison to a Sysmex XT-2000iV

	EXIGO			Sysmex	ProCyte			Sysmex	VetScan HM2			Sysmex
Measurand (units)	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**
RBC (x 10^{12}/L)	**6.70**	-0.34	1.06	0.13	7.09	0.13	1.48	0.35	**6.68**	0.32	1.51	1.04
HGB (g/L)	16.86	-0.40	0.99	0.32	*16.62*	0.17	1.83	0.32	**16.58**	1.08	1.67	0.68
HCT (%)	48.76	-2.14	0.88	0.20	44.82	-1.12	1.22	0.46	**46.16**	1.99	1.53	1.26
MCV (fL)	72.72	0.24	0.23	0.42	63.22	-2.82	0.41	0.42	69.20	-0.29	1.21	0.74
MCH (pg)	25.12	0.60	0.59	0.34	23.46	-0.18	0.38	0.57	24.76	0.38	1.44	0.69
MCHC (g/dL)	34.56	0.64	0.60	0.32	37.10	1.29	0.66	0.48	*35.88*	0.78	1.76	0.89
WBC (x 10^9/L)	10.44	-0.80	2.40	1.57	10.60	0.02	2.12	1.30	**18.39**	5.87	5.30	1.85
Platelets (x 10^9/L)	328.60	-43.80	3.32	0.86	**197.00**	-37.60	4.43	8.02	**109.60**	-258.10	10.73	15.93
RDW-CV (%)	15.00	-2.56	0.00	0.51	n/a	n/a	n/a	1.76	**16.36**	0.45	2.31	15.95
MPV (fL)	**6.06**	-4.46	0.90	0.43	11.84	0.51	1.53	1.18	**8.06**	-2.48	3.57	2.89

	InSight V5			Sysmex	LaserCyte			Sysmex	VetScan HM5a			Sysmex
Measurand (units)	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**
RBC (x 10^{12}/L)	6.82	-0.47	0.21	0.62	**6.67**	0.36	4.19	0.62	**6.52**	0.72	0.83	0.34
HGB (g/L)	17.30	-0.47	0.88	0.46	14.66	-0.28	1.57	0.35	13.98	-0.20	0.69	0.48
HCT (%)	46.76	-4.15	0.29	0.35	**44.04**	0.11	3.26	0.56	**45.36**	4.87	0.97	0.44
MCV (fL)	68.58	-1.31	0.12	0.37	**65.98**	-3.67	1.16	0.32	*69.50*	-0.25	1.44	0.62
MCH (pg)	25.40	1.01	0.62	0.36	**22.00**	-1.70	5.00	0.66	**21.43**	-3.01	0.59	0.55
MCHC (g/dL)	37.08	2.17	0.70	0.42	**33.34**	-0.66	4.02	0.59	**30.80**	-4.19	1.33	0.87
WBC (x 10^9/L)	**8.65**	-1.05	5.19	2.12	**8.84**	0.94	6.98	1.83	5.69	0.32	2.32	2.338
Platelets (x 10^9/L)	248.80	-6.40	1.12	3.97	**240.00**	82.20	4.55	4.22	*218.00*	-41.00	3.27	1.87

RDW-CV (%)	**11.72**	-4.44	7.59	2.36	n/a	n/a	n/a	0.48	14.63	1.42	1.17	0.24
MPV (fL)	9.92	-1.36	1.31	3.26	**18.46**	5.65	9.99	1.86	8.30	-1.47	0.98	1.87

		VetScan HM2b		Sysmex
Measurand (units)	**mean**	**Bias (obs)**	**CV (%)**	**CV (%)**
RBC (x 10^{12}/L)	**7.23**	0.38	1.32	0.60
HGB (g/L)	**16.92**	-0.51	2.42	0.29
HCT (%)	*52.66*	2.04	1.36	0.63
MCV (fL)	**73.00**	-7.07	0.00	0.55
MCH (pg)	**23.42**	-2.04	3.33	0.79
MCHC (g/dL)	**32.16**	0.35	3.25	0.90
WBC (x 10^{9}/L)	**14.34**	5.55	4.77	1.28
Platelets (x 10^{9}/L)	**103.80**	-43.09	11.01	1.94
RDW-CV (%)	15.74	0.75	0.57	14.35

n/a = result not submitted
Bold = Poor performance for POC haematology analysers
Italics = Fair performance

The CV (%) values were higher for the majority of measurands when compared to the reference laboratory. The mean observed CV (%) across the measurands ranged from 0.74 – 4.0%. Notably high CV's were seen for WBC for the VetScan HM2 (5.30%), InSight V5 (5.19%), LaserCyte (6.9%), PLT for VetScan HM2 (10.7%) and VetScan HM5b (11%), RDW-CV (7.6%) for the InSight V5 and MPV (10%) for the LaserCyte (Table 7.1).

7.3.3 Quality goal performance

7.3.3.1 TE_a

Only one of the seven participating analysers (ProCyte) could achieve TE_a requirements based on total allowable error as recommended by ASVCP (RBC, HGB, HCT, MCV, MCH, MCHC, WBC) or by expert opinion (PLT, RDW-CV, MPV) for all measurands returned for comparison.
For the remaining POC analysers that failed to meet TE_a (ASVCP) or expert opinion for one or more measurand, this was seen for RBC (1), HCT (1), MCV (1), MCH (1), MCHC (1), WBC (2), PLT (1), RDW-CV (1), MPV (2) (Table 7.2).

Table 7. 2: The results for 10 haematology measurands that could achieve/not achieve total allowable error as recommended by ASVCP or by expert opinion analysed on 7 in-clinic point of care haematology analysers using individual normal canine samples

		EXIGO	**ProCyte**	**VetScan HM2**	**InSight V5**
Measurand (units)	TE_a/Expert opinion goal	Achieved/not achieved (Y/N)	Achieved/not achieved (Y/N)	Achieved/not achieved (Y/N)	Achieved/not achieved (Y/N)
RBC (x 10^{12}/L)	10	Y	Y	Y	Y
HGB (g/L)	10	Y	Y	Y	Y
HCT (%)	10	Y	Y	Y	Y
MCV (fL)	7	Y	Y	Y	Y
MCH (pg)	10	Y	Y	Y	Y
MCHC (g/dL)	10	Y	Y	Y	Y
WBC (x 10^9/L)	15	Y	Y	**N**	Y
Platelets (x 10^9/L)	30	Y	Y	**N**	Y
RDW-CV (%)	30	Y	N/A	Y	**N**
MPV (fL)	30	**N**	Y	Y	Y

		LaserCyte	**VetScan HM5a**	**VetScan HM5b**
Measurand (units)	TE_a/Expert opinion goal	Achieved/not achieved (Y/N)	Achieved/not achieved (Y/N)	Achieved/not achieved (Y/N)
RBC (x 10^{12}/L)	10	Y	**N**	Y
HGB (g/L)	10	Y	Y	Y
HCT (%)	10	Y	**N**	Y
MCV (fL)	7	Y	Y	**N**
MCH (pg)	10	Y	**N**	Y
MCHC (g/dL)	10	Y	**N**	Y
WBC (x 10^9/L)	15	Y	Y	**N**
Platelets (x 10^9/L)	30	Y	Y	Y
RDW-CV (%)	30	N/A	Y	Y
MPV (fL)	30	**N**	Y	Y

Y= Yes,; N= No; N/A= result not submitted by in-clinic participant

7.3.3.2 Sigma metrics

A total of 68 sigma metrics were calculated (2 analysers reported 9 measurands out of ten; and of these only 37 measurands (54%) achieved > 4 sigma across all the POC analysers. However, 2 analysers were able to achieve > 4 sigma for 80% of measurands (Exigo EosVet & InSight V5), one analyser (ProCyte) for 89% and one analyser for 60% of measurands (VetScan HM5a).

The remaining 3 analysers could only achieve > 4 sigma for 3 of 10 (MCV, MCH, MCHC) (VetScan HM2), 2 of 10 (RDW-CV, HCT) (VetScan HM5b) and 1 of 9 (HGB) measurands (LaserCyte).

Sigma metrics < 4 (46% of measurands) before and after harmonisation were assessed using the QGI; of these 33 of 40 (90.3%) were attributed to bias, and 3 (9.7%) were attributed to imprecision. RBC, WBC, PLT and MPV had the highest frequency of sigma metric < 4 attributed to bias (see supplementary material, Appendix E, Table E7.1).

7.3.3.3 Overall performance using the performance key grading system

Performance was graded as “poor” or “fair” for almost half (48.5%) of the comparisons to the reference laboratory and “excellent” or “good” for the remainder. (Table 7.1, see bold/italic values)

In-clinic analysers ranked in order of overall excellent performance, (indicated by the largest number of "excellent” performance grades across all

measurands; highest first) were: Woodley InSight V5, IDEXX ProCyte, EXIGO EosVet , Abaxis VetScan HM5a, Abaxis VetScan HM2, Abaxis VetScan HM5b, IDEXX LaserCyte (Table 7.3).

Using the points system to gauge the performance for each POC in-clinic analyser could simplify clinic messaging and the monitoring of future improvement by distilling quality performance across multiple measurands to a single easily conveyed and understood value. The following scores thresholds were proposed based on the data in this study:

Excellent/Good performance = > or = 48

Fair performance = 47-22

Low performance = < 22

Using the points system the analysers ranked (according to highest score) in a similar order to the performance grading key except for the IDEXX ProCyte and EXIGO EosVet which changed position (Table 7.3).

Table 7. 3: Performance rank for each POCT haematology analyser using the performance grading system compared to the scoring points system to determine best performance for 7 in-clinic POC haematology analysers when compared to a reference haematology analyser.

Grading System		**Scoring System**	
Rank	**Result**	**Rank**	**Result**
InSight V5	8 Excellent, 2 poor	InSight V5	47
ProCyte	7 Excellent, 1 Fair, 2 Poor	**Exigo**	44
Exigo	6 Excellent, 2 Good, 2 Poor	**ProCyte**	41.4
VetScan HM5a	4 Excellent, 2 Fair, 5 Poor	VetScan HM5a	35
VetScan HM2	1 Excellent, 1 Good, 1 Fair, 7 Poor	VetScan HM2	20
VetScan HM5b	1 Excellent, 1 Fair, 8 Poor	VetScan HM5b	19
LaserCyte	1 Excellent, 8 Poor	LaserCyte	17.1

Bold type = changes in rank between the two systems

7.3.4 Harmonisation

Twenty-eight bias adjustments from a total of 68 evaluations were made across all the individual in-clinic analysers following assessment using the combined inherent imprecision to harmonise results. These adjustments were required to account for significant bias as the results were not interchangeable between the in-clinic POC and reference laboratory analysers. The bias was added or subtracted from the POC result so that the 'harmonised' POC result was the same as that obtained by the reference laboratory. Table 7.4 presents the data for the pre- and post-harmonised in-clinic POC analyser results compared to the reference laboratory analyser. The post-harmonisation results for the in-clinic POC analysers are all within the RI for the reference laboratory.

Table 7. 4: Pre-and Post-harmonisation results evaluated with canine samples using point of care testing haematology analysers showing the measurands that were not within combined inherent imprecision when compared against a reference analyser and required harmonisation.

			EXIGO			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
RBC (x 10^{12}/L)	7.04	6.70	7.03	Poor	Fair	Yes
HCT (%)	50.90	48.76	50.90	Good	Excellent	Yes
RDW-CV (%)	17.56	15.00	17.56	Excellent	Excellent	Yes
MPV (fL)	10.52	6.06	10.52	Poor	Excellent	Yes
			ProCyte			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
MCV (fL)	66.04	63.22	66.04	Excellent	Excellent	Yes
MCHC (g/dL)	35.81	37.10	35.81	Excellent	Excellent	Yes
			VetScan HM2			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
WBC (x 10^{9}/L)	12.52	18.39	12.52	Poor	**Poor (< 4 sigma, CV 5.3%)**	Yes
RDW-CV (%)	15.91	16.36	15.91	Poor	Fair	Yes
MPV (fL)	10.00	8.06	9.42	Poor	**Poor (< 4 sigma, Bias still 10%)**	Yes
			InSight V5			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
RBC (x 10^{12}/L)	7.29	6.82	7.28	Excellent	Excellent	Yes
HCT (%)	50.91	46.76	50.91	Excellent	Excellent	Yes
MCV (fL)	69.89	68.58	69.89	Excellent	Excellent	Yes
MCH (pg)	24.39	25.40	24.39	Excellent	Excellent	Yes
MCHC (g/dL)	34.91	37.08	34.91	Excellent	Excellent	Yes
RDW-CV (%)	16.16	11.72	16.16	Poor	**Poor (< 4 sigma, CV 7.5%)**	Yes

			LaserCyte			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
RBC (x 10^{12}/L)	6.31	6.67	6.31	Poor	**Poor (< 4 sigma, CV 4.1%)**	Yes
MCV (fL)	69.65	65.98	69.65	Poor	Excellent	Yes
Platelets (x 10^{9}/L)	157.80	240.00	157.80	Poor	Excellent	Yes
			VetScan HM5a			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
RBC (x 10^{12}/L)	5.80	6.52	5.80	Poor	Excellent	Yes
HCT (%)	40.49	45.36	40.49	Poor	Excellent	Yes
MCH (pg)	24.43	21.43	24.44	Poor	Excellent	Yes
MCHC (g/dL)	34.99	30.80	34.99	Poor	Excellent	Yes
Platelets (x 10^{9}/L)	259.00	218.00	259.00	Fair	Excellent	Yes
RDW-CV (%)	13.21	14.63	13.21	Excellent	Excellent	Yes
MPV (fL)	9.77	8.30	9.77	Excellent	Excellent	Yes
			VetScan HM5b			
Measurand (units)	**Ref lab result**	**Initial POCT result**	**Harmonised POCT result**	**Initial performance grade**	**Post-harmonisation grade**	**Within RI (ref Lab)**
MCV (fL)	80.07	72.72	80.07	Poor	Excellent	Yes
WBC (x 10^{9}/L)	8.79	14.34	8.79	Poor	Fair	Yes
RDW-CV (%)	14.99	15.74	14.99	Excellent	Excellent	Yes

7.3.5 Participant feedback

The technical and pathologist comments on the spreadsheets were provided so the results could be easily understood by the participating clinic. Where poor performance was noted (i.e. < 4 sigma), the interpretation of the result of the in-clinic analyser and the reference analyser was the same except for one measurand at one POC laboratory: MPV where the significant bias (42.3%) meant interpretation was not the same. These comments provided context to the numerical results for the clinician, so they are more easily understood. If the interpretation of the results was the same based on the reference laboratory reference interval, no action plan was needed to address the poor performance.

Follow up telephone discussions with the participating veterinary practice were used to obtain feedback on the value of the scheme to them. Of the 7 participants, 4 had most of the poor performance (42%) and of these 4 only 1 participant was aware of having analytical error and contacted technical support for their POC analyser. The clinicians found the process simple and time allocation was acceptable for both clinicians and technicians taking approximately 40 minutes on average to select the sample, analyse results and complete the spreadsheet. The time allocation for the reference laboratory was longer at the beginning, processing, and entering the results took on average 70 minutes but this time reduced with more participants and by the final participant, the time allocation was approximately 50 minutes. Comments by technicians and pathologists required 10 min on average.

These time allocations for technician results production, technical comments and pathologist comments were considered reasonable and suitable for a comparative testing program.

7.4 Discussion

Based on the results of this study, a comparison scheme for POC haematology analysers is a useful and beneficial offering to veterinary clinicians that could enhance analytical trust for their routine CBCs using veterinary patient specimens. The clinician and/or technician can select and process a patient sample, send it to the reference laboratory the same day which keeps any preanalytical factors to a minimum and returns the results in a timely manner. Specimen repetitions were considered carefully for the program due to time, cost, and encouragement to participate in this scheme. Five replicates was selected and previously identified (Rishniw *et al.*, 2012) as a sufficient amount for performance evaluation based on a compromise between time, cost and accuracy of estimates of bias and imprecision. Many results were returned on the day of analysis. However, difficulties arose in trying to contact the practice to further discuss the individual analyser performance. For future comparisons, a summary sheet that includes a score for overall performance would be useful for each in-clinic analyser. It would allow at a glance understanding of the performance for each measurand similar to EQA reports. This was not presented to the clinicians from the POC in-clinic laboratories but added when analysing the results. Our numerical score approach was comparable to the performance grading system for

ranking the analysers according to best performance and would allow the clinician to compare results to another participant similar to a comparative testing program. This should be included in future communications with clinicians and as a summary of ongoing overall performance.

Only one clinic concluded technical support was required for their analyser (VetScan HM5a) during the conduct of the exercise and carried out 4 repetitions rather than five. This was due to noticing poor internal quality control results the same day as the scheduled comparison with our reference analyser. Our study shows that clinicians using these analysers cannot always recognise poor analyser performance in the absence of a scheme such as this. The anecdotally, infrequent use of statistical QC for POC analysers is problematic and the issue with this particular POC analyser may not have been otherwise detected. This has been previously demonstrated (Rishniw and Pion, 2016) and steps to ensure accuracy and reliability as part of a quality management system should be in place for POC analysers.

Our results demonstrated that analytical error was present for all model POC haematology analysers that took part in this comparative scheme. Using canine samples ensured evaluations were relevant to the analyser throughput and provided a good assessment for performance issues. There was a wide variation in performance which was primarily due to bias, and of significant concern for WBC, PLT, RDW-CV and MPV where each had average bias >18%. High bias has been identified in previous studies for WBC (13.32%) and PLT (17.15%) although RDW-CV and MPV were not assessed (Rishniw and Pion, 2016). Imprecision as the major contributor to poor performance (determined using the QGI) was less common (9.7%) than

bias (90%) and there was considerable overlap in results for CV (%) for haematology POC analysers found by Rishniw *et al* (2016) for the following measurands, RBC, HGB, HCT, WBC and PLT demonstrating acceptable reproducibility between current and previous studies when using 5 replicates. The use of the reference laboratory as the method of comparison has limitations due to its own inherent bias and imprecision but we were confident using the reference analyser as the target value as this analyser has been through previous quality evaluation (Daly *et al.*, 2022) and has more stringent QC methods in place (Daly *et al.*, 2023) to ensure analytical error is minimised. However, when comparing the POC analysers to the reference analyser the CVs (%) were generally lower for the reference analyser with some exceptions, PLT (ProCyte, VetScan HM2) and RDW-CV (VetScan HM2) where the CVs (%) for the reference analyser were notably higher. This is of low significance/concern for platelets which are routinely assessed using nonstatistical QC (blood smear evaluation) in the reference laboratory and should also be assessed in the POC laboratory.

TE_a or expert opinion

We found that ASVCP recommended TE_a or expert opinion was not always achieved by the in-clinic analysers (16%) and it's important to note that ASVCP TE_a goals were derived from quality control material which may have different CVs compared to those obtained with patient samples. CVs are expected to be higher with POC analysers, (Nabity *et al.*, 2018) and failure to meet ASVCP recommended TE_a or expert opinion occurred more frequently

for WBC and MPV and only one analyser achieved the TE_a goals for all measurands which suggests that ASVCP recommended TE_a goals may be an unrealistic goal for POC haematology analysers and require revision.

Sigma

Sigma metrics have been used as an excellent indicator of performance for veterinary haematology analysers previously (Daly *et al.*, 2021) (Rishniw and Pion, 2016). In this study, only 54% of measurands achieved > 4 sigma which means that 46% of measurands had poor performance that may result in the inability to compare in-clinic and reference lab results. RBC, WBC, PLT, RDW-CV as well as MPV has the highest incidences of < 4 sigma. Sigma metric < 4 equates to at least 6210 errors per million results which is a significant rate of error. Overall, sigma metrics were easily understood by the clinicians as a risk marker of performance. We found that using ASVCP recommended TE_a or expert opinion in conjunction with sigma metrics was sufficient for monitoring the performance of the POC analysers. In particular, sigma metrics showed the degree (almost 50%) of substandard results due to bias that require substantial quality improvement for POC haematology analysers.

Harmonisation

The veterinary clinics in this study routinely send patient samples whilst using POCT in parallel and for these participating clinics which are all within a

50km radius of the reference analyser, the option to apply a correction factor/adjustment for the observed bias is a possibility and would rule out any patient misclassification due to any excessive bias for the in-clinic analysers. For example, high bias (%) was on average seen for WBC (20%), PLT (25%), RDW-CV (18%) and MPV (22%) and utilising a correction factor would harmonise the results so that common reference limits could be used and initial diagnosis and follow up on patient samples could be conducted using the POC or reference analyser. We found that when comparing the results using combined inherent imprecision that 41.17% were not within combined inherent imprecision. Therefore, these results required a correction to be interpreted equivalently. Once the adjustment was made and the result harmonised (by adding or subtracting the correction factor [average bias]), sigma metrics and ASVCP recommended TE_a or expert opinion goal improved for the majority of measurands. Using the performance grading criteria, 5 measurands remained “poor” (for RBC, WBC, RDW-CV, MPV) as they still had sigma metric < 4 (mainly due to imprecision) but achieved ASVCP recommended TE_a or expert opinion goals. When assessing the individual harmonised result, the clinical interpretation was the same as the reference laboratory. To ensure a correction factor for results within the reference interval would be useful for future comparisons, the average bias could be added or subtracted from results that were at the lower and higher end of the in-clinic laboratories RIs and reassessed for compliance with ASVCP recommended TE_a goal and whether sigma metric was > 4. Additionally, for WBC and PLT, review of the blood film is always warranted but not conducted routinely in clinics so harmonising these measurands

would be a useful and prevent inaccurate interpretation of the in-clinic analyser results. For platelets, correlation with the blood smear is necessary to check for platelet clumping and this would be a useful addition to any further comparative testing program.

Future directions

For ongoing harmonisation of in-clinic analysers to a reference laboratory so that results are comparable irrespective of the method, this study would have to be repeated across a range of clinical decision levels to determine those measurands for each POC laboratory that require harmonisation and those that are currently harmonised. Periodic evaluation to ensure ongoing maintenance of harmonisation is then needed. Previous reports on a veterinary laboratory network of haematology analysers demonstrated that harmonisation was maintained for most measurands across a range of results of clinical importance using HCT as the target when manipulating samples to generate abnormal specimens (chapter 6). This program could be viewed as a low-cost alternative to EQA subscriptions whereby a reference laboratory can offer an EQA program to veterinary clinics using their services rather than them subscribing to an alternative commercial EQA program but consideration would need to be given to the amount of staffing resources available to clinics. We would suggest a periodic interlaboratory comparison test across 3 levels rather than 5 used in a previous harmonisation study and to determine whether correction for bias is needed across all levels of clinical interest. It would be feasible to generate a correction chart for results

falling within clinical decision range so the practice can manually harmonise the result or automate this via analyser specific software. We hypothesise that this is a less expensive and simpler approach to a traditional comparison of methods study with regression as it requires fewer patient results (15 versus 40 samples). The fact that many of the measurands for which excessive bias was identified as a cause of poor performance is of concern since the POCT reference intervals were often close to or the same as those of the reference laboratory, suggesting that their use may result in misclassification of some patients' results if not corrected for bias. However once harmonised, the in-clinic laboratory could use the reference laboratory reference intervals and carrying out a reference interval transference validation following harmonisation would provide the highest confidence for using the reference laboratory reference intervals.

7.5 Conclusion

To our knowledge this is a first report of a comparative testing program designed and implemented to harmonise veterinary in-clinic POC haematology analysers compared to a reference laboratory. The performance evaluation showed that in-clinic POC haematology analysers could not always reliably give accurate results or achieve quality goals and that bias was the biggest contributor to analytical error.

Further support for quality assurance and quality control for in-clinic analysers would be beneficial and veterinary reference analysers could be utilised as this support as part of a service agreement or as a subscription.

Expanding this study across the clinical decision range would be advantageous in order ensure harmonisation was maintained across a range of results of clinical importance.

CHAPTER 8: DISCUSSION AND LIMITATIONS

8.1 Reviewing the scope of the thesis

This research used an evidence-based, step-by-step approach to evaluate performance of veterinary haematology analysers operating within a commercial network and embarked on a journey towards harmonisation by (1) providing quality goals and evaluating performance of the network analysers using a commercially available quality control material, (2) performing statistical QC validation to ensure a high P_{ed} and low P_{fr} efficiently harmonising the analyser network (3) challenging the usefulness of commercial QCM in a veterinary laboratory network, (4) implementing RPT-QC as an alternative QC method, (5) establishing internal comparative testing with a designated reference analyser to monitor bias and maintain harmonisation achieved and (6) conducting a pilot program for comparative testing for POC haematology analysers with the reference analyser at one laboratory.

A disparity between the quality requirements and expectations between human and veterinary medicine laboratories was evident in the literature. For a commercial veterinary laboratory system, the deficit in specific guidelines for veterinary laboratories meant the five laboratory sites were not collaborating effectively and were operating independently despite circulating patient samples cross functionally and using the same reference intervals. This research stemmed from a desire to make necessary improvements for

QC/QA and using the few reports on laboratory harmonisation in the literature (Flatland *et al.*, 2018a; Flatland *et al.*, 2010; Theodorsson, 2016; Plebani, 2013; Plebani, 2016) as a foundation for planning and implementing the QA/QC improvements. This foundation allowed us to apply quality goals such as BV-based goals, TE_a as recommended by ASVCP/Expert opinion and sigma metrics to characterise acceptable analytical performance for harmonisation. Harmonisation was accepted and maintained when the network of analysers achieved biologic variation-based and TE_a-based goal criteria. This would facilitate the unification of the laboratory network, help it function more effectively and work as a harmonised system using common reference intervals. The aim was to raise the bar for veterinary laboratory quality, reduce the risk of patient misclassification, challenge the status quo on running minimal QCM on haematology analysers using a potentially non-commutable matrix which has not been addressed in the literature for veterinary laboratories but commutability is regarded important aspect towards harmonisation (Miller *et al.*, 2011a; Miller, 2003; Unsal *et al.*, 2010). RPT-QC was instituted as the necessary change to improve the commutability of the QC to reflect veterinary patient samples. These efforts hinged on educating the technical and pathology teams on the benefits of laboratory improvement and harmonisation.

8.2 Revisiting research findings

The process towards harmonisation started with evaluation of performance of the 6 haematology analysers within the laboratory network (chapter 2), this showed that a step-by-step process could be utilised to evaluate and optimise performance for haematology for multi-site/multi-instrument laboratories. Farr & Freeman, 2008 (Farr and Freeman, 2008) had previously highlighted the lack of veterinary specific quality standards and this chapter adds to the veterinary literature and was the first publication for a laboratory network. The second chapter hypothesised that the performance of the analysers would exhibit inherent variability across the analytical network, as previously reported (Flatland and Weiser, 2014; Rishniw *et al.*, 2012). We confirmed that for a coordinated laboratory network, i.e. using the exact same analyser models, using the same SOP, QCM and frequency of QC, that there was inherent variability across the analytical systems which may not otherwise be known. Multiple quality goals were evaluated to determine those which were most useful in detection of poor performance, and these showed that veterinary analysers fail to achieve multiple quality goals, which was also demonstrated by Rishniw (Rishniw and Pion, 2016).

This performance evaluation defined those analysers within the network that were underperforming and required attention and improvement to correct the root cause of poor performance. The strategy for this study used the quality goals as bench markers of performance to help determine the differences between the analytical performance and the root causes, including imprecision, bias, lack of standardisation and lack of consistent monitoring. In

this chapter, sigma metrics underpinned performance across the measurands and analysers as the main tool for detecting unstable performance. This concluded that < 5 sigma was an indicator of poor performance and demonstrated analysers that warranted further action and technical support. When improvements were implemented, the network of analysers were harmonised according to biologic variation-based and TE_a-based criteria. The use of sigma metrics has been reported previously (Nevalainen *et al.*, 2000; Westgard, 2001; El Sharkawy *et al.*, 2018) but this has been not previously been applied to a veterinary laboratory network. Improvements in sigma metrics were seen across the evaluation period for the network analysers which demonstrated that this degree of performance tracking was a valuable exercise. This study was a necessary improvement process that brought multiple benefits such as demonstrating analysers showing 'world class performance' and those requiring technical support. In addition, it demonstrated the varying analytical performance when using the same evaluation criteria, the quality goals and identified a combination of quality goals useful for veterinary haematology analysers and for demonstrating harmonisation of the network of analysers. Without this work laboratories in a network would continue to work independently, and poor performing analysers would remain undetected. If performance evaluations take place at an individual level, then even the analysers that perform at a decent level/perform favourably aren't benchmarked and this gives the impressions of analytical capability that could be less favourable when compared to other analysers of the same make/model.

To maintain the improved analytical performance and harmonisation achieved in chapter 2, statistical QC validation (Chapter 3) was applied to the network haematology analysers to ensure that ongoing QC had a high P_{ed} and low P_{fr}. This enabled the network to be more efficient by utilising a 1-2.5s QC rule to achieve a > 0.94 P_{ed} and < 0.05 P_{fr}. It is understood that a reduction in waste and cost may result in improved efficiency and more rapid turnaround time, which is a key performance indicator for our laboratory network, as well as for most commercial laboratories. The evaluation also showed that the previous use of manufacturer's acceptable QC limits were not suitable for achieving a high P_{ed} and that QC needed to be customised for the performance of the analysers within the group. After implementation of customised QC for the analyser network, the use of the sigma metric was further refined from a 5.0 sigma metric (identified in Chapter 2) to implement a 5.5 sigma strategy for QC. The 5.5 sigma cut-off was used for ongoing monitoring of instrument performance. This is first study of its kind for a veterinary haematology analyser network. The literature (Rishniw *et al.*, 2012; Flatland *et al.*, 2014b) states that $P_{ed} \geq 0.85$ is acceptable for veterinary haematology analysers when using a single QCM however a key finding in this research (chapter 3) showed that when manufacturers acceptable limits were used, only 3/10 measurands for the reference analyser could achieve the acceptable P_{ed} (≥ 0.85), leading to poor error detection. When this QC approach was applied to individual analyser observed performance across the network, it showed a significant improvement in P_{ed} (> 0.94) when using a 1-2.5s QC rule with the recommended TE_a therefore supporting the hypothesis that using a QC

approach does support network harmonisation. This study concluded that manufacturer's limits are not suitable to maintain an acceptable error budget and individual analyser observed performance data should be utilised for veterinary haematology analysers. In addition, it showed that sigma metrics were an excellent performance tool that can identify unstable performance which is sufficient for harmonisation maintenance and is easily adapted to veterinary laboratories.

In chapter 4, the chronic issue for veterinary laboratories regarding non-commutable matrix for QCM was tackled and whether RPT-QC could be validated across the haematology network. Material commutability has been addressed in human medicine (Miller, 2003; Vesper *et al.*, 2016) but limited research on this issue for veterinary analysers is available (Vanyo *et al.*, 2018). Westgard (Westgard, 2016b) describes a method for 'Implementing Repeat Patient Test Controls' and while this method was explored in other studies (Vanyo *et al.*, 2018; Flatland and Freeman, 2018b; Flatland and Freeman, 2020) it was not applied to a multiple instrument/multisite system. The research presented in this thesis explored using RPT-QC as an alternative QC method to monitor imprecision for the network of harmonised haematology analysers using a comprehensive QC validation approach. This further supports previous research (Manzocchi *et al.*, 2017) indicating the need for QC changes and challenges the usefulness of commercial non-animal QCM to veterinary laboratories. This study demonstrated that RPT-QC could be successfully generated and executed across the analyser network using "normal" canine samples using the quality goals as a measure of performance.

Chapter 5 delivers a template to allow other laboratories or veterinary hospitals to implement RPT-QC for their haematology analysers as a means of establishing a species-specific matrix according to their caseload. The inclusion of performance goals demonstrated that RPT-QC could achieve conditions specified in chapter 2, i.e. $P_{ed} > 0.85$ using a simple QC rule using ASVCP recommended goals for TE_a. HCT was associated with a higher TE_a than that specified by ASVCP, but a larger TE_a could be considered and applied for a calculated HCT methodology once RBC and MCV were within acceptable TE_a. These are important adaptations for veterinary laboratories that may avoid the ineffective practice of implementing quality goals that are too stringent and unachievable with the current state-of-the-art technology. The risk associated with unrealistic and unachievable targets is loss of engagement towards quality overall for an industry that requires much needed quality improvements. One important finding in this study was that the RPT-QC limits generated were rarely validated using the recommended number of data points (20 data points (Westgard, 2016b)), and 40 data points was frequently required to calculate the SD required for determining the QC limits. This chapter provides a statistical evaluation that determined that the analytical network could not use common/unified RPT-QC limits due to differences in SDs and confidence intervals whereby the QC limits rarely overlapped. Therefore this answers the hypothesis that individual analyser RPT-QC validation and limit generation is required. This supports previous conclusions that inherent variability exists between analysers even when using the same model. This answers an important question for veterinary laboratories and potentially veterinary clinics who may consider adopting

RPT-QC limits from other laboratories/practices using the same analytical instruments.

When implementing any changes to the laboratory, the technical and pathology team need to be involved in decision making; therefore, including them in the process is paramount. Their feedback allowed for valuable changes in how the RPT-QC data was managed. The automated spreadsheet was adapted to ensure their input was included and was as user-friendly as possible. The automated spreadsheet required extensive development effort and after multiple consultations with a visual basic (VB) coder outlining the design required for this study, this spreadsheet is now a useful template for other individual users and could easily be adapted to other types of testing for RPT-QC for this network. From this study, we can conclude that RPT-QC can be used as an alternative to commercial QCM. Sigma metrics can also be used to identify poor performance and, when incorporated with non-statistical QC methods, can ensure a robust monitoring system. A finding that will be of interest to the wider community is that the TE_a range for HCT for use with RPT-QC is higher than proposed by ASVCP TE_a recommendations (Nabity *et al.*, 2018) which was discovered in chapter 5.

Chapter 6 presents the design and implementation of the internal laboratory comparison (ILC) testing program which successfully maintained ongoing harmonisation and facilitated the use of common reference intervals for the haematology network. In addition to using species-specific samples for daily QC (chapter 5), this chapter addresses the same sample commutability concerns arising from EQA programs that typically use samples of human

origin. In this ILC program the species-specific sample was a cost-effective alternative to EQA and could identify bias within the network by comparison with the network reference analyser. The use of the quality goals as combined quality criteria to rank performance was useful. TE_a and biologic variation-based goals for CV and bias, were not always achieved despite clinical interpretation remaining equivalent between the analysers for the extreme levels. This indicated that adjustments to TE_a for veterinary analysers can be justified and review of the ASVCP recommendations is warranted for patient samples over the range of clinically important decision limits. Sigma metrics > 5.5 remained the benchmark of excellent performance and any poor performance flagged was easily identifiable and monitored at the next ILC event using the automated spreadsheets customised for each laboratory site. Thus, the individual periodic evaluations demonstrated a practical and systematic approach to laboratory harmonisation for a network across a range of results of clinical interest. Using canine samples over a range of clinical interest concentrations for a network of veterinary haematology analysers has not previously been reported.

The findings reported in Chapters 5 and 6, provide a protocol for veterinary laboratories to monitor imprecision and bias as part of a total quality management system using commutable samples.

The final chapter established an extension of the ILC program (chapter 6) as an offering to in-clinic POC haematology analysers (Chapter 7). Performance of point-of-care (POC) analysers was evaluated in comparison to the network reference analyser. This chapter shows how to design a comparative testing

program for external use using the previous chapters as guides to establish a performance criteria to harmonise POC analysers to a commercial reference laboratory. This comparison study shows that POC haematology analysers are not always accurate or reliable. Using the performance criteria, they failed to achieve quality goals a significant proportion of time (54% achieved sigma metric > 4, 87% of participants failed to achieve TE_a/expert opinion for one or more measurands) and this was mainly due to bias. Analytical error was identified for all model POC haematology analysers assessed in the program. A result adjustment was made which corrected the high levels of bias so that they could be interpreted equivalently. This resulted in the ability of the POC analyser to harmonise their results with the reference laboratory analyser so that either the POC or reference laboratory analysers could be used for monitoring of patients. This should lead to increased confidence for the clinicians in the usefulness of the POC analyser results and facilitate the use of reference intervals from the reference laboratory to interpret their POC results.

This is a first report of a comparative testing program that aimed to harmonise the in-clinic-POC analysers to a reference analyser.

8.3 Importance and relevance to the wider veterinary laboratory sector

Veterinary laboratory quality appears to be a small component of the publications of veterinary research. However, its application has a significant impact on producing reliable and accurate laboratory results that are important for making medical decisions about veterinary patients. This is important since inaccurate results can have serious consequences for the patient, practitioner/clinician, and the laboratory. In addition, the inability to correctly diagnose pets/animals can have a direct influence on human lives, reflecting mental and financial stress of animal owners and with increased risk of zoonotic and reverse-zoonotic diseases as well as diseases that may share common exposures in animals and people, that may occur more frequently through frequent global movement of people and animals. Therefore, veterinary research and improvements play a crucial role in mitigating and reducing disease between humans and their pets through reliable and accurate analytical diagnostics.

This research contributes to the field of veterinary laboratory quality by emphasising the definition and importance of realistic and attainable quality of analytical haematology performance.

What defines this? The combination of studies above has resulted in a shareable step-by-step guide to achieving harmonisation through use of quality goals and metrics to enhance interpretation of results produced across different laboratories within a network and the potential for

harmonisation of POC analysers to their reference analyser. As mentioned earlier, this is a template that can be adapted and implemented across the wider veterinary community and highlights the use of quality goals. We classified our quality goal criteria as TE_a as recommended by ASVCP or by internal expert opinion, biologic variation-based goals for CV and Bias and sigma metrics. However, from our first chapter we included the use of inherent imprecision and laboratory systems goal ($0.33XCV_I$) (Fraser, 1992b) which we concluded were not discriminatory for poor performance within the analytical network.

The final chapter concluded that TE_a goals as recommended by ASVCP often were not achievable by POC haematology analysers and their application may be counterproductive based on the current technological capabilities. This research provides working knowledge of applying these goals to multiple analysers and deciding/demonstrating the quality goals that were the most useful in performance monitoring. The publication of applying these goals in a veterinary laboratory makes them readily available to the veterinary community.

Quality goals that provide a relevant and useful performance management tool, are proposed as discussed above. The network analysers and the POC analysers did not always achieve TE_a as recommended by ASVCP and internal expert opinion from the laboratory quality group based on what could be controlled with QC. This is not unexpected because ASVCP TE_a goals were developed based on QCM rather than patient samples means (Nabity *et al.*, 2018). The adaptations incorporated into the laboratory network are relevant to other veterinary laboratories who wish to use species-specific

patient samples as a QC method. This study also supports using a 'reverse approach' (Korchia and Freeman, 2022) of TE_a goal setting based on observed performance characteristics for individual analysers and as a sensible way of promoting realistic targets and engaging the veterinary community. Decisions based on the TE_a that can be controlled based on the needs of veterinary reference and POC laboratories should be carefully considered so that clinicians are informed regarding the accuracy and precision associated with reference laboratory and POC testing results which are otherwise unregulated and do not have a centralised quality structure. The use of biologic variation goals for veterinary laboratories is useful but may not always be applicable or achievable. For our well maintained and harmonised analysers, we could achieve optimal or desirable BV-based goals for CV and bias for most measurands, providing confidence and reassurance that patient results would not be misclassified.

Throughout this research sigma metrics was an easily understood and adaptable metric within the laboratory. It is important metric for cost control because sigma < 4 is associated with greater complexity and frequency of QC, therefore ensuring a high sigma metric keeps cost waste to a minimum. The hierarchy as proposed by the 'the Milan Consensus' (Westgard, 2016d) will not always be easily adaptable to veterinary laboratories. This research would stimulate veterinary laboratories to assess their hierarchy of quality goals and including TE_a (recommended by ASVCP or expert opinion), BV goals for bias and CV and sigma metrics is the recommendation arising from this work. Despite extensive literature dedicated to human laboratory quality goals (Westgard and Westgard, 2016), this is still not a well understood

concept for most veterinary laboratories. This research now provides practical information about its applications in veterinary laboratories.

This study finds the commercial QCM substandard in its application for veterinary analysers due to wide manufacturer's limits that do not provide acceptable error detection. This is in addition to understanding how commutable QC material provides much needed assurances as to how patient samples will behave on the analysers, supporting the recommendation to incorporate RTP-QC and interlaboratory comparisons into the veterinary laboratory.

For the laboratory team (Technicians and Pathologists) this research provides fundamental educational material to promote a quality culture within the laboratory. It's an evidence-based account of the inherent variability within the laboratory network of haematology analysers and it identified poor performance for some analysers that was unexpected by the laboratory team. This inherent variability was evident for one site that had two analysers which demonstrated marked differences in performance (analysers 4a and 4b) and highlighted the need for ongoing performance monitoring post-harmonisation. This process showed the team that when improvements are made, and analysers are optimised and harmonised, then performance tracking shows improvement over time (chapter 2). The performance evaluations used were new introductions to the team and this led to greater understanding and appreciation of the use of quality goals, resulting in more experienced, knowledgeable, and competent staff.

This research also demonstrated insights into the stability of samples and variation expected over multiple time intervals (Chapter 4 & 6). The need to

customise QC (whether commercial QCM or RPT-QC) for the individual instrument was a significant realisation and helped personnel understand why we cannot apply one set of QC limits across multiple analysers. Overall, the changes implemented from this research allowed the team to upskill in their knowledge and techniques, which can lead to greater job satisfaction.

The reliance on and interpretation of haematology results for wellness testing, diagnosis or monitoring depends on whether these are reliable and accurate values from either the reference laboratory or POC analysers. Interpreting results from analysers that have imprecision and/or bias can lead to misclassification of the patient (Baral *et al.*, 2015; Flatland *et al.*, 2014a; Rishniw *et al.*, 2012) and this research was an opportunity to educate the clinicians about these topics. Chapter 7 demonstrated to the participants that their POC analyser had bias (90% of poor performance identified) in comparison to the reference analyser. Using the quality goals was an important way to demonstrate the degree of performance management that is required to ensure patient samples are not misclassified. In addition, the ability to harmonise to the reference analyser and potentially use the reference analysers' RIs is an opportunity welcomed by clinicians to have more confidence in their POC analysers. Chapter 7 also found the RIs were often similar to the reference analysers RIs which suggests the RIs provided by the manufacturer for POCT analysers may be ineffective in detecting important changes in the patient and may not reflect bias that is present. Conducting the comparison was an excellent way to nurture the relationship with the practice and to educate them about the application of POC results for their patients. It also draws attention to the pre-analytical factors

particularly around sample integrity and species-specific matrix for QC. It shows that individual analyser performance evaluation is required to determine if the analyser is appropriate for use for their patients.

As discussed, harmonisation for veterinary analysers is not reported in the literature and the industry relies on human laboratory regulations or recommendations that are not always helpful to a veterinary laboratory. Using this study as a template and as a quality improvement opportunity can ensure that haematology analysers are optimised, undergo sufficient quality control to ensure an appropriate sensitivity in error detection (P_{ed}) and specificity for waste reduction (P_{fr}) are present, and that individual instruments are 'fit for purpose'. This study provides guidance for achieving harmonisation that can be used by veterinary clinical pathologists, veterinary clinicians, and laboratory technicians to achieve excellence in laboratory results.

8.4 Challenges and limitations

An unexpected and challenging component to this research was personnel management, response to change and implementing changes. In the initial stages, the research project was rolled out to the QA/QC team, which comprised of a group of veterinary clinical pathologists and veterinary technicians. These individuals were responsible for implementing actions and changes at their site. Meetings were productive and engaging and early actions such as optimising QCM, SOPs and maintenance documentation

went well. When the data collection process began for RPT-QC, some reluctance became apparent at certain sites which prompted the author to visit each site and deliver presentations to ensure full engagement and address any uncertainty. This was successful and RPT-QC data was collected despite a shortfall in staffing due to the COVID pandemic. However, at the implementation stage for RPT-QC as a QC method, a lack of support at top management level meant this process was slow, frustrating, and required multiple touch points to ensure the research was on track. Poor communication at one site meant the technician did not fully understand why they were repeating patient samples and therefore did not know to note on the spreadsheet the reason(s) for QC failure(s). Numerous lunch-and-learn presentations were delivered remotely to support the team through the various changes and suggest that for any laboratory undertaking such an effort should consider appointing a champion for change at each site to support others through the process. Each site had different challenges and needed different time scales to acclimate and consider the changes thoroughly. As this was a time sensitive research project, this unexpected variable created a life lesson (Cyprus, 2024) on the psychology of change management and the importance of insuring that all stakeholders are fully on board. This would likely present itself as a similar challenge for other users following the approach laid out in this research and understanding that these initiatives can’t work without people engagement, this would allow time for educational preparation from the initial planning stages of the research. Change management (Gibbs, 2013) and resistance to change is well documented (Creasey, 2022; Phillips and Klein, 2023) and not necessarily

unique to this study but certainly having an change plan in place ahead of implementing the initiatives will save time and energy.

Electing to validate four measurands for RPT-QC can be considered a challenge, however we choose these measurands as the most useful in our network and eliminated the calculated indices because of sample ageing and because they can be assessed using non-statistical QC for haematology. This includes blood smear review to validate the platelet count and the white blood cell differential. Another challenge for some laboratories may be sample volume and availability of fresh samples; we could validate RPT-QC for up to 4 days to demonstrate a way to overcome this issue and considering our laboratory's needs and in addition to having an alternative protocol for periods of extended closure (Chapter 5). The RPT-QC design is important to ensure that fresh samples are available, and that samples and control limits are appropriate for the laboratory schedule.

8.5 Overall learning opportunities

This research required a broad background in general laboratory quality for both human and veterinary laboratories. It was necessary to understand the harmonisation approach for human laboratories (Plebani, 2013; Miller *et al.*, 2011a; Vesper *et al.*, 2016; Greenberg, 2014; Houcke *et al.*, 2013) before embarking on a total quality system applicable for harmonising and maintaining the harmonisation framework for veterinary laboratories. In keeping with the literature, we adopted the ideology of continuous

improvement which is where the concept of the research initially arose. This deep dive into total quality management (TQM) resulted in a steep learning curve on quality management, continuous improvements, defining, establishing, and maintaining a climate of quality and educating those involved. This led to an increased knowledge of quality goals that are not well developed for veterinary laboratories. These goals are at the centre of this research and understanding and upskilling on how to assess the goals that could apply to a veterinary laboratory team is an important part of quality planning. The evaluation of the goal setting to ensure our error budgets were satisfactory and could act as quantitative quality management means some goals were not discriminatory for veterinary haematology (inherent imprecision and lab systems goals) and our recommendation towards a hierarchy of quality goals is well documented in the chapters. With little corporate experience using sigma metrics within this business as a risk assessment tool, the increased knowledge, and skills in applying a sigma strategy to monitor performance was essential and strengthened the harmonisation of the network with a metric that led to a more quantitative process.

Prior to this research there was not systematic analysis of performance in the study laboratory business. Only by initiating quality goals was it possible to show how having a quality plan and evaluating our haematology analysers performance pre- and post-harmonisation led to better performance. This process resulted in improved knowledge of personnel particularly at the reference analysers site. This new skill in leading laboratory quality culture through the quality initiatives meant experience was gained in intra- and

interlaboratory leadership and education. This worked well to promote a positive attitude towards laboratory quality despite some challenges surrounding change. This was emphasised again an area of continuous professional improvement for handling change management. Having a team undergo an upskilling process i.e. quality improvements, understanding imprecision and bias and the benefits of RPT-QC was an independent educational piece that improved the technical expertise in the laboratory. The differences within the analytical network provided evidence for the need for harmonisation and the undoubtable risk of different clinical interpretations without quality management. The increased knowledge and understanding of analyser-specific performance issues and troubleshooting increased awareness of evidenced-based veterinary medicine, and offered guidelines on how to achieve harmonisation to those that were not aware of the implications for a lack of quality planning and management.

In addition to the above, this research allowed the development of certain soft skills such as improved scientific and technical writing, increased awareness of the importance of SOPs and documenting troubleshooting events and actions. New skills included the use of statistics and statistical programs and planning and designing spreadsheet design prior to a consultation process with an expert in automation of excel spreadsheets for veterinary-specific applications.

8.6 Conclusion

Chapter 2: Optimisation and harmonisation were achieved using quality goals. Performance evaluation ensured that the error budget for each analyser is within the ASVCP recommended goal for total allowable error, sigma metric strategy and desired BV-based goals. Because BV-derived bias goals were achieved, common reference intervals could be implemented across the network of analysers.

Chapter 3: Using commercial quality control material demonstrated that a higher P_{ed} (> 0.94) can be achieved using QC rules (1-2.5s) validated for each individual analyser than that achieved using manufacturers acceptable limits at the recommended TE_a.

Chapter 4: RPT-QC was evaluated as a novel QC approach and validated in each laboratory. TE_{obs} was the same or better than when using manufacturer's acceptable limits. Validation of RPT-QC limits required 40 data points rather than 20 and these limits differed across the network if analysers. The challenges that mimicked unstable performance were successful.

Chapter 5: RPT-QC was successfully implemented and integrated into a multi-site veterinary laboratory. The statistical analysis showed that RPT-QC limit generation is recommended for each individual

analyser. An automated spreadsheet was used to collect data with hard coded RPT-QC limits, Levey-Jennings chart generation and monthly pooled standard deviation, successfully highlighting OOC events.

Chapter 6: A internal laboratory comparison study and design using the same sample as a performance monitoring tool and using quality goals; TE_a/expert opinion goals, BV goals, and sigma metrics was successful in maintaining harmonisation for the selected measurands. This had cost saving and species-specific advantages and shows this program can monitor performance across a range of results of clinical importance.

Chapter 7: A first report of a comparative testing program used to harmonise POC haematology analysers to an established commercial reference analyser. This study shows that POC haematology analysers do not always achieve quality standards or produce reliable results and that this is mainly due to bias.

This sum of works shows that a step-by-step guide to achieving harmonisation can be designed and implemented for veterinary haematology analysers. This ensures that businesses can operate as a laboratory network with a low risk of patient misclassification between the analysers.

8.7 Future directions

8.7.1 Recommendations for future research

Recognition by all levels of lab management within the studied business of the importance of various interacting elements in promotion of laboratory quality culture, as outlined by Theodorsson is needed. A '*staircase of responsibility*' including technical and pathologist involvement, provision of resources and support for quality is needed as a means of improving efficiency, efficacy, and client satisfaction (Theodorsson, 2016).

A global recommendation of quality goals and some form of QC regulation would be beneficial for veterinary laboratories and for in-clinic laboratories who are lagging behind human laboratories to a degree that the human literature could be overwhelming and out of reach for veterinary laboratories. However, there is currently no appetite for regulation since the cost involved is often perceived by veterinarians as prohibitive. This research shows that a cost effective approach to quality can be achieved but that ongoing education of veterinarians is needed to demonstrate the need for and effectiveness of such approaches.

Further development of RPT-QC by using abnormal samples and including other species would extend the scope of this as a QC method. For the in-clinic laboratory, RPT-QC processes would need to be adapted as IQC should be done each day that patient samples are run. This currently is not the recommendation given by manufacturers of POC instruments to

veterinary practices with POC haematology analysers. There needs to be a cost-effective protocol specifically for implementing RPT-QC that is attractive for in-clinic laboratories. In addition, using an abnormal sample for POC comparative testing results (across a range of clinically important haematology results rather than using ‘normal’ specimens in this pilot), and ongoing development of best practices for harmonisation of POC and reference laboratory results across a range of clinically important results. There appears to be a strong need for development of supportive continuing education presentations for in-clinic laboratories (vets and vet technicians) on the important and understanding of QA/QC.

Following harmonisation, reference interval transference validation is needed for analysers wishing to use the manufacturer’s provided reference interval and theoretically should be done for the reference laboratory reference interval if the POC instrument results are harmonised to the reference laboratory. Other possible future research include nonstatistical QC for POC haematology testing, comparative testing for abnormal results (across a range of clinically important haematology results since we only did ‘normal’ specimens in this pilot), and ongoing development of best practices to harmonisation of POC and reference laboratory results across a range of clinically important results.

CHAPTER 9: REFERENCES

AACC. *The International Consortium for Harmonization of Clinical Laboratory Results* [Online]. Available: https://www.harmonization.net/ [Accessed 21st May 2022].

AAHA. Available: https://www.aaha.org/aaha-accreditation/what-is-accreditation/ [Accessed January 20 2024].

AARSAND AK, F.-C. P., WEBSTER C, COSKUN A, GONZALES-LAO E, DIAZ-GARZON J, JONKER N, SIMON M, BRAGA F, PERICH C, BONED B, MARQUES-GARCIA F, CAROBENE A, ASLAN B, SEZER E, BARTLETT WA, SANDBERG S. *EFLM Biological Variation Database* [Online]. Available: https://biologicalvariation.eu/ [Accessed March 12th 2023].

ADLM. *Association for Diagnostic & Laboratory Medicine* [Online]. Available: https://www.myadlm.org/ [Accessed 20th August 2023].

ANTONY, J., PALSUK, P., GUPTA, S., MISHRA, D. & BARACH, P. 2018. Six Sigma in healthcare: a systematic review of the literature. *International Journal of Quality & Reliability Management,* 35**,** 1075-1092.

ARMBRUSTER, D. & MILLER, R. R. 2007. The Joint Committee for Traceability in Laboratory Medicine (JCTLM): a global approach to promote the standardisation of clinical laboratory test results. *Clin Biochem Rev,* 28**,** 105-13.

ASVCP. *ASVCP POC guideline* [Online]. Available: https://www.asvcp.org/page/QALS_Guidelines [Accessed January 20 2024].

BARAL, R. M., DHAND, N. K., FREEMAN, K. P., KROCKENBERGER, M. B. & GOVENDIR, M. 2013. Biological variation and reference change values of feline plasma biochemistry analytes. *Journal of Feline Medicine and Surgery,* 16**,** 317-325.

BARAL, R. M., DHAND, N. K., MORTON, J. M., KROCKENBERGER, M. B. & GOVENDIR, M. 2015. Bias in feline plasma biochemistry results between three in-house analysers and a commercial laboratory analyser: results should not be directly compared. *Journal of Feline Medicine and Surgery,* 17**,** 653-666.

BARNETT, R. N. 1968. Medical significance of laboratory results. *Am J Clin Pathol,* 50**,** 671-6.

BARNETT, R. N. & WEISBROT, I. M. Clinical laboratory statistics. 1979.

BAUER, N., NAKAGAWA, J., DUNKER, C., FAILING, K. & MORITZ, A. 2012. Evaluation of the automated hematology analyzer Sysmex XT-2000iV compared to the ADVIA (R) 2120 for its use in dogs, cats, and horses. Part II: Accuracy of leukocyte differential and reticulocyte count, impact of anticoagulant and sample aging. *J Vet Diagn Invest,* 24**,** 74-89.

BAYAT, H. 2018. Expected long-term defect rate of analytical performance in the medical laboratory: Assured Sigma versus observed Sigma. *Biochem Med (Zagreb),* 28**,** 020101.

BECKER, M., MORITZ, A. & GIGER, U. 2009. Comparative clinical study of Canine and feline total blood cell count results with seven in-clinic and two commercial laboratory hematology analyzers. *Veterinary clinical pathology / American Society for Veterinary Clinical Pathology,* 37**,** 373-84.

BEHREND, E., HOLFORD, A., LATHAN, P., RUCINSKY, R. & SCHULMAN, R. 2018. 2018 AAHA Diabetes Management Guidelines for Dogs and Cats. *J Am Anim Hosp Assoc,* 54**,** 1-21.

BERG, J. & LANE, V. 2011. Pathology Harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory. *Ann Clin Biochem,* 48**,** 195-7.

BLAND, J. M. & ALTMAN, D. G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet,* 1**,** 307-10.

BLOIS, S. L., LANG, S. T., WOOD, R. D. & MONTEITH, G. 2015. Biologic variability and correlation of platelet function testing in healthy dogs. *Veterinary Clinical Pathology,* 44**,** 503-510.

BRAGA, F. & PANTEGHINI, M. 2013. Standardization and analytical goals for glycated hemoglobin measurement. *Clin Chem Lab Med,* 51**,** 1719-26.

BRITTIN, G. M., BRECHER, G., JOHNSON, C. A. & ELASHOFF, R. M. 1969. Stability of blood in commonly used anticoagulants. Use of refrigerated blood for quality control of the Coulter Counter Model S. *Am J Clin Pathol,* 52**,** 690-4.

CAMUS, M. S., FLATLAND, B., FREEMAN, K. P. & CRUZ CARDONA, J. A. 2015. ASVCP quality assurance guidelines: external quality assessment and comparative testing for reference and in-clinic laboratories. *Vet Clin Pathol,* 44**,** 477-92.

CAO, S. & QIN, X. 2018. Application of Sigma metrics in assessing the clinical performance of verified versus non-verified reagents for routine biochemical analytes. *Biochem Med (Zagreb),* 28**,** 020709.

CARSTAIRS, K. C., PETERS, E. & KUZIN, E. J. 1977. Development and description of the "random duplicates" method of quality control for a hematology laboratory. *Am J Clin Pathol,* 67**,** 379-85.

CEMBROWSKI, G., LUNETZKY, E., PATRICK, C. & WILSON, M. 1988. Optimized quality control procedure for hematology analyzers using retained patient specimens. *American journal of clinical pathology,* 89**,** 203-10.

CEMBROWSKI GS, S. A. 1996. *Quality Control and Statistics, Chapter 4 in Clinical Chemistry: Principles, Procedures, Correlations,* Lippincott, Philadelphia.

CERIOTTI, F., BOYD, J. C., KLEIN, G., HENNY, J., QUERALTÓ, J., KAIRISTO, V. & PANTEGHINI, M. 2008. Reference intervals for serum creatinine concentrations: assessment of available data for global application. *Clin Chem,* 54**,** 559-66.

CLSI. 2007. *GP27-A2 Using Proficiency Testing to Improve the Clincial Laboratory; Approved Guideline* [Online]. Available: https://webstore.ansi.org/preview-pages/CLSI/preview_CLSI+GP27-A2.pdf [Accessed 10th July 2023].

CLSI. 2008. *C38-A3: Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline* [Online]. Available: https://webstore.ansi.org/preview-pages/CLSI/preview_CLSI+C28-A3.pdf [Accessed 20th July 2023].

CLSI. 2013. *Clinical and Laboratory Standards Institute Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline - Third Edition* [Online]. Available: https://clsi.org/media/1435/ep09a3_sample.pdf [Accessed 19th September 2023].

COOK, A. M., MORITZ, A., FREEMAN, K. P. & BAUER, N. 2016. Quality requirements for veterinary hematology analyzers in small animals-a survey about veterinary experts' requirements and objective evaluation of analyzer performance based on a meta-analysis of method validation studies: bench top hematology analyzer. *Vet Clin Pathol,* 45**,** 466-76.

COSKUN, A. 2010. *Quality Management and Six Sigma*. India: www.Sciyo.com. ISBN 978-953-307-130-5

COTLOVE, E., HARRIS, E. K. & WILLIAMS, G. Z. 1970. Biological and Analytic Components of Variation in Long-Term Studies of Serum Constituents in Normal Subjects : III. Physiological and Medical Implications. *Clinical Chemistry,* 16**,** 1028-1032.

CREASEY, T. 2022. *5 Levels of Change Management Maturity* [Online]. Available: https://www.prosci.com/blog/change-management-maturity-model [Accessed 15th February 2024].

CYPRUS, S. 2024. *What is a Life Lesson?* [Online]. Available: https://www.languagehumanities.org/what-is-a-life-lesson.htm [Accessed 11th February 2024].

DALY, S., FREEMAN, K. P. & GRAHAM, P. A. 2021. Evaluation of Sysmex XT-2000iV analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol,* 50**,** 568-578.

DALY, S., GRAHAM, P. A. & FREEMAN, K. P. 2022. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000iV hematology analyzers. *Vet Clin Pathol,* 51**,** 565-576.

DALY, S., GRAHAM, P. A. & FREEMAN, K. P. 2023. Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000iV hematology analyzers. *Veterinary Clinical Pathology,* n/a.

DEVELOPMENT, S. E. E. A. 2012. *SEED Haematology - The Red Blood Indices* [Online]. Available: https://www.sysmex-europe.com/academy/library/educational-articles-seed/seed-the-red-blood-cell-indices-1131.html [Accessed].

DUTRA, F. R. 1966. Monitoring the Quality of Blood Cell Counts with Replicate Determinations on Routine Samples. *American Journal of Clinical Pathology,* 46**,** 286-288.

EKINS, R. 1991. Immunoassay standardization. *Scand J Clin Lab Invest Suppl,* 205**,** 33-46.

EL SHARKAWY, R., WESTGARD, S., AWAD, A. M., AHMED, A. O. I., IMAN, E. H., GABALLAH, A. & SHAHEEN, E. 2018. Comparison between Sigma metrics in four accredited Egyptian medical laboratories in some biochemical tests: an initiative towards sigma calculation harmonization. *Biochem Med (Zagreb),* 28**,** 020711.

EZRules3. Available Westgard website. Available: https://www.westgard.com. Accessed 9th June 2021]

FARR, A. J. & FREEMAN, K. P. 2008. Quality control validation, application of sigma metrics, and performance comparison between two biochemistry analyzers in a commercial veterinary laboratory. *J Vet Diagn Invest,* 20**,** 536-44.

FEI, Y., WANG, W., HE, F., ZHONG, K. & WANG, Z. 2015. Evaluating Laboratory Performance on Point-of-Care Glucose Testing with Six Sigma Metric for 151 Institutions in China. *Diabetes Technol Ther,* 17**,** 745-54.

FLATLAND, B., BARAL, R. M. & FREEMAN, K. P. 2020. Current and emerging concepts in biological and analytical variation applied in clinical practice. *Journal of Veterinary Internal Medicine,* 34**,** 2691-2700.

FLATLAND, B., BREICKNER, L. C. & FRY, M. M. 2014a. Analytical performance of a dry chemistry analyzer designed for in-clinic use. *Vet Clin Pathol,* 43**,** 206-17.

FLATLAND, B., CAMUS, M. S. & BARAL, R. M. 2018a. Analytical quality goals-a review. *Vet Clin Pathol,* 47**,** 527-538.

FLATLAND, B. & FREEMAN, K. P. 2018b. Repeat patient testing shows promise as a quality control method for veterinary hematology testing. *Vet Clin Pathol,* 47**,** 252-266.

FLATLAND, B. & FREEMAN, K. P. 2020. Repeat patient testing-based quality control shows promise for use in veterinary biochemistry testing. *Veterinary Clinical Pathology,* 49**,** 590-606.

FLATLAND, B., FREEMAN, K. P., FRIEDRICHS, K. R., VAP, L. M., GETZY, K. M., EVANS, E. W. & HARR, K. E. 2010. ASVCP quality assurance guidelines: control of general analytical factors in veterinary laboratories. *Veterinary Clinical Pathology,* 39**,** 264-277.

FLATLAND, B., FREEMAN, K. P., VAP, L. M. & HARR, K. E. 2013. ASVCP guidelines: quality assurance for point-of-care testing in veterinary medicine. *Veterinary Clinical Pathology,* 42**,** 405-423.

FLATLAND, B., FRIEDRICHS, K. R. & KLENNER, S. 2014b. Differentiating between analytical and diagnostic performance evaluation with a focus on the method comparison study and identification of bias. *Vet Clin Pathol,* 43**,** 475-86.

FLATLAND, B. & VAP, L. M. 2012. Quality management recommendations for automated and manual in-house hematology of domestic animals. *Vet Clin North Am Small Anim Pract,* 42**,** 11-22.

FLATLAND, B. & WEISER, G. 2014. The Elephant in the Room (and How to Lead It Out): In-Clinic Laboratory Quality Challenges. *Journal of the American Animal Hospital Association,* 50.

FRANZINI, C. & CERIOTTI, F. 1998. Impact of reference materials on accuracy in clinical chemistry. *Clin Biochem,* 31**,** 449-57.

FRASER, C. G. 1992a. Biological variation in clinical chemistry. An update: collated data, 1988-1991. *Arch Pathol Lab Med,* 116**,** 916-23.

FRASER, C. G. 1992b. Biological Variation: From Principles to Practice. Amer. Assoc. for Clinical Chemistry.

FRASER, C. G. 2001. Biological Variation: From Principles To Practice. *Quality Specifications***,** 29-66.

FRASER, C. G., HYLTOFT PETERSEN, P., LIBEER, J. C. & RICOS, C. 1997. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem,* 34 (Pt 1)**,** 8-12.

FRASER, C. G. & PEAKE, M. J. 1980. Problems associated with clinical chemistry quality control materials. *CRC Crit Rev Clin Lab Sci,* 12**,** 59-86.

FRASER, C. G., PETERSEN, P. H., RICOS, C. & HAECKEL, R. 1992. Proposed quality specifications for the imprecision and inaccuracy of analytical systems for clinical chemistry. *Eur J Clin Chem Clin Biochem,* 30**,** 311-7.

FREEMAN, K. & GRUENWALDT, J. 1999. Quality Control Validation in Veterinary Laboratories. *Veterinary clinical pathology / American Society for Veterinary Clinical Pathology,* 28**,** 150-155.

FREEMAN, K., KLENNER-GASTREICH, S. & KORCHIA, J. 2024. The importance of quality control validation and relationships with total error quality goals and bias in the interpretation of laboratory results. *Veterinary Clinical Pathology,* 53**,** 65-74.

FREEMAN, K. P., BARAL, R. M., DHAND, N. K., NIELSEN, S. S. & JENSEN, A. L. 2017. Recommendations for designing and conducting veterinary clinical pathology biologic variation studies. *Vet Clin Pathol,* 46**,** 211-220.

FREEMAN, K. P. & FLATLAND, B. 2017. Studies of instrument/method comparison, as part of validation for use in veterinary species. *Veterinary Clinical Pathology,* 46**,** 5-8.

FRIEDRICHS, K. R., HARR, K. E., FREEMAN, K. P., SZLADOVITS, B., WALTON, R. M., BARNHART, K. F. & BLANCO-CHAVEZ, J. 2012. ASVCP reference interval guidelines: determination of de novo reference intervals in veterinary species and other related topics. *Veterinary Clinical Pathology,* 41**,** 441-453.

FURLANELLO, T., TASCA, S., CALDIN, M., CARLI, E., PATRON, C., TRANQUILLO, M., LUBAS, G. & SOLANO-GALLEGO, L. 2006. Artifactual changes in canine blood following storage, detected using the ADVIA 120 hematology analyzer. *Veterinary Clinical Pathology,* 35**,** 42-46.

GEFFRÉ, A., FRIEDRICHS, K., HARR, K., CONCORDET, D., TRUMEL, C. & BRAUN, J. P. 2009. Reference values: a review. *Vet Clin Pathol,* 38**,** 288-98.

GIBBS, K. 2013. Managing Change. *Clinical Laboratory Management.* (eds L.S. Garcia, P. Bachner, V.S. Baselski, M.R. Lewis, A.J. Linscott, D.A. Schwab, J.C.H. Steele, A.S. Weissfeld, D.S. Wilkinson and D.M. Wolk). https://doi.org/10.1128/9781555817282.ch14

GRASBECK, R. 1969. Establishment and use of normal values. *Scand J Clin Invest,* 110**,** 1-15.

GRÄSBECK, R. & FELLMAN, J. 1968. Normal values and statistics. *Scand J Clin Lab Invest,* 21**,** 193-5.

GREENBERG, N. 2014. Update on current concepts and meanings in laboratory medicine — Standardization, traceability and harmonization. *Clinica Chimica Acta,* 432**,** 49-54.

HARR, K., FLATLAND, B., NABITY, M. & FREEMAN, K. 2013. ASVCP Guidelines: allowable total error. *Veterinary clinical pathology / American Society for Veterinary Clinical Pathology,* 42**,** 424-36.

HARRISON, H. H. & JONES, J. B. 2017. Using Sigma Quality Control to Verify and Monitor Performance in a Multi-Instrument, Multisite Integrated Health Care Network. *Clinics in Laboratory Medicine,* 37**,** 207-241.

HENS, K., BERTH, M., ARMBRUSTER, D. & WESTGARD, S. 2014. Sigma metrics used to assess analytical quality of clinical chemistry assays: importance of the allowable total error (TEa) target. *Clinical Chemistry and Laboratory Medicine (CCLM),* 52**,** 973-980.

HORN, P. S. & PESCE, A. J. 2003. Reference intervals: an update. *Clin Chim Acta,* 334**,** 5-23.

HOROWITZ, G. L. 2008. Reference Intervals: Practical Aspects. *Ejifcc,* 19**,** 95-105.

HOUCKE, S. K. V., AELST, S. V., UYTFANGHE, K. V. & THIENPONT, L. M. 2013. Harmonization of immunoassays to the all-procedure trimmed mean – proof of concept by use of data from the insulin standardization project. *Clinical Chemistry and Laboratory Medicine,* 51**,** e103-e105.

HUBERT, V., HENDRIK, E., MATTHEW, G., CHANDRA, P., GREGORY, M. W., ROBERT, R., GERHARD, S., JILL, T., LINDA, T. & JEFFREY, E. V. Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine - Approved Guideline. 2010.

ISO. 2003. *International Organization for Standardization (ISO) ISO 17511:2003* [Online]. Available: https://www.iso.org/standard/30716.html [Accessed 11th August 2023].

ISO. 2007. *International Organization for Standardization ISO/IEC Guide 99:2007* [Online]. Available: https://www.iso.org/standard/45324.html [Accessed 11th July 2023].

ISO. 2009. *International Organization for Standardization ISO 15193:2009* [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:15193:ed-2:v1:en [Accessed 9th June 2023].

ISO. 2020. *International Organization for Standardization ISO 21151:2020* [Online]. Available: https://www.iso.org/standard/69985.html [Accessed 5th September 2023].

ISO/IEC. 2023. *International Organization for Standardization ISO/IEC 17043:2023* [Online]. Available: https://www.iso.org/standard/29366.html [Accessed 2023 5th July].

JANSEN, R. T. P., COBBAERT, C. M., WEYKAMP, C. & THELEN, M. 2018. The quest for equivalence of test results: the pilgrimage of the Dutch Calibration 2.000 program for metrological traceability. *Clin Chem Lab Med,* 56**,** 1673-1684.

JASSAM, N., LINDSAY, C., HARRISON, K., THOMPSON, D., BOSOMWORTH, M. P. & BARTH, J. H. 2011. The implementation of a system for managing analytical quality in networked laboratories. *Annals of Clinical Biochemistry,* 48**,** 136-146.

JENSEN, A. L. & KJELGAARD-HANSEN, M. 2006. Method comparison in the clinical laboratory. *Veterinary Clinical Pathology,* 35**,** 276-286.

JONES, G. R., BARKER, A., TATE, J., LIM, C. F. & ROBERTSON, K. 2004. The case for common reference intervals. *Clin Biochem Rev,* 25**,** 99-104.

JONES, G. R. D. & LEGG, M. 2019. Report formatting in laboratory medicine – a call for harmony. *Clinical Chemistry and Laboratory Medicine (CCLM),* 57**,** 61-65.

KALLNER, A., KHOROVSKAYA, L. & PETTERSSON, T. 2005. A method to estimate the uncertainty of measurements in a conglomerate of instruments/laboratories. *Scand J Clin Lab Invest,* 65**,** 551-8.

KJELGAARD-HANSEN, M. & JENSEN, A. L. 2010. Subjectivity in defining quality specifications for quality control and test validation. *Veterinary Clinical Pathology,* 39**,** 134-135.

KORCHIA, J. & FREEMAN, K. P. 2022. Total observed error, total allowable error, and QC rules for canine serum and urine cortisol achievable with the Immulite 2000 Xpi cortisol immunoassay. *Journal of Veterinary Diagnostic Investigation,* 34**,** 246-257.

LESTER, S., HARR, K. E., RISHNIW, M. & PION, P. 2013. Current quality assurance concepts and considerations for quality control of in-clinic biochemistry testing. *Journal of the American Veterinary Medical Association,* 242**,** 182-192.

LIPPI, G., BANFI, G., CHURCH, S., CORNES, M., DE CARLI, G., GRANKVIST, K., KRISTENSEN, G. B., IBARZ, M., PANTEGHINI, M., PLEBANI, M., NYBO, M., SMELLIE, S., ZANINOTTO, M. & SIMUNDIC, A. M. 2015. Preanalytical quality improvement. In pursuit of harmony, on behalf of European Federation for

Clinical Chemistry and Laboratory Medicine (EFLM) Working group for Preanalytical Phase (WG-PRE). *Clin Chem Lab Med,* 53**,** 357-70.

LIPPI, G., BECAN-MCBRIDE, K., BEHÚLOVÁ, D., BOWEN, R. A., CHURCH, S., DELANGHE, J., GRANKVIST, K., KITCHEN, S., NYBO, M., NAUCK, M., NIKOLAC, N., PALICKA, V., PLEBANI, M., SANDBERG, S. & SIMUNDIC, A. M. 2013. Preanalytical quality improvement: in quality we trust. *Clin Chem Lab Med,* 51**,** 229-41.

LITTEN, J. 2017. Applying Sigma Metrics to Reduce Outliers. *Clinics in Laboratory Medicine,* 37**,** 177-186.

LUMSDEN, J. H. & MULLEN, K. 1978. On establishing reference values. *Can J Comp Med,* 42**,** 293-301.

MANZOCCHI, S., FURMAN, E. & FREEMAN, K. 2017. Comparison of 3 options for choosing control limits in biochemistry testing. *Vet Clin Pathol,* 46**,** 120-125.

MAO, X., SHAO, J., ZHANG, B. & WANG, Y. 2018. Evaluating analytical quality in clinical biochemistry laboratory using Six Sigma. *Biochem Med (Zagreb),* 28**,** 020904.

MARQUES-GARCIA, F., GARCIA-CODESAL, M. F., CARO-NARROS, M. R. & CONTRERAS-SANFELICIANO, T. 2015. Importance of implementing an analytical quality control system in a core laboratory. *Rev Calid Asist,* 30**,** 302-9.

MCLAWHON, R. W. 2011. Patient safety and clinical effectiveness as imperatives for achieving harmonization inside and outside the clinical laboratory. *Clin Chem,* 57**,** 936-8.

MEIJER, P., KYNDE, K., VAN DEN BESSELAAR, A., VAN BLERK, M. & WOODS, T. A. L. 2018. International normalized ratio (INR) testing in Europe: between-laboratory comparability of test results obtained by Quick and Owren reagents. *Clin Chem Lab Med,* 56**,** 1698-1703.

MILLER, W., MYERS, G. L., LOU GANTZER, M., KAHN, S. E., SCHONBRUNNER, E. R., THIENPONT, L. M., BUNK, D. M., CHRISTENSON, R. H., ECKFELDT, J. H., LO, S. F., NUBLING, C. M. & STURGEON, C. M. 2011a. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem,* 57**,** 1108-17.

MILLER, W. G. 2003. Specimen materials, target values and commutability for external quality assessment (proficiency testing) schemes. *Clin Chim Acta,* 327**,** 25-37.

MILLER, W. G., BRUNS, D. E., HORTIN, G. L., SANDBERG, S., AAKRE, K. M., MCQUEEN, M. J., ITOH, Y., LIESKE, J. C., SECCOMBE, D. W., JONES, G., BUNK, D. M., CURHAN, G. C. & NARVA, A. S. 2009. Current issues in measurement and reporting of urinary albumin excretion. *Clin Chem,* 55**,** 24-38.

MILLER, W. G., ECKFELDT, J. H., PASSARELLI, J., ROSNER, W. & YOUNG, I. S. 2014a. Harmonization of test results: what are the challenges; how can we make it better? *Clin Chem,* 60**,** 923-7.

MILLER, W. G., JONES, G. R., HOROWITZ, G. L. & WEYKAMP, C. 2011b. Proficiency testing/external quality assessment: current challenges and future directions. *Clin Chem,* 57**,** 1670-80.

MILLER, W. G. & MYERS, G. L. 2013. Commutability still matters. *Clin Chem,* 59**,** 1291-3.

MILLER, W. G., MYERS, G. L., GANTZER, M. L., KAHN, S. E., SCHÖNBRUNNER, E. R., THIENPONT, L. M., BUNK, D. M., CHRISTENSON, R. H., ECKFELDT, J. H., STANLEY, G. L., NUBLING, C. M. & STURGEON, C. M. 2013. [The ways of harmonization of clinical laboratory measurement techniques]. *Klin Lab Diagn***,** 54-61.

MILLER, W. G., MYERS, G. L. & REJ, R. 2006. Why commutability matters. *Clin Chem,* 52**,** 553-4.

MILLER, W. G., TATE, J. R., BARTH, J. H. & JONES, G. R. 2014b. Harmonization: the sample, the measurement, and the report. *Ann Lab Med,* 34**,** 187-97.

NABITY, M., HARR, K., CAMUS, M., FLATLAND, B. & VAP, L. 2018. ASVCP guidelines: Allowable total error hematology. *Veterinary Clinical Pathology,* 47.

NEVALAINEN, D., BERTE, L., KRAFT, C., LEIGH, E., PICASO, L. & MORGAN, T. 2000. Evaluating laboratory performance on quality indicators with the six sigma scale. *Arch Pathol Lab Med,* 124**,** 516-9.

ORSER, B. 1983. Hematology quality control and the Coulter S-Plus II. *Am J Med Technol,* 49**,** 643-7.

OSTERTAGOVÁ, E., OSTERTAG, O. & KOVÁČ, J. 2014. Methodology and Application of the Kruskal-Wallis Test. *Applied Mechanics and Materials,* 611**,** 115-120.

PANTEGHINI, M. 2009. Traceability as a unique tool to improve standardization in laboratory medicine. *Clin Biochem,* 42**,** 236-40.

PARRY DM. 2009. *Quality Goal Index* [Online]. Available: https://www.westgard.com/guest34.htm [Accessed June 2020].

PARVIN, C. A. 2017. What's New in Laboratory Statistical Quality Control Guidance? The 4th Edition of CLSI C24, Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions. *J Appl Lab Med,* 1**,** 581-584.

PETERSEN, P., FRASER, C., JØRGENSEN, L., BRANDSLUND, I., STAHL, M., GOWANS, E., LIBEER, J.-C. & RICÓS, C. 2002. Combination of Analytical Quality Specifications Based on Biological Within- and Between-Subject Variation. *Annals of clinical biochemistry,* 39**,** 543-50.

PETERSEN, P. H. & FRASER, C. G. 2010. Strategies to set global analytical quality specifications in laboratory medicine: 10 years on from the Stockholm consensus conference. *Accreditation and Quality Assurance,* 15**,** 323-330.

PETRIDES, V. & SCHNEIDER, S. 2018. Using Sigma metrics to establish analytical product performance requirements and optimize analytical performance of an in vitro diagnostic assay using a theoretical total PSA assay as an example. *Biochem Med (Zagreb),* 28**,** 020903.

PHILLIPS, J. & KLEIN, J. D. 2023. Change Management: From Theory to Practice. *TechTrends,* 67**,** 189-197.

PLEBANI, M. 2013. Harmonization in laboratory medicine: the complete picture. *Clin Chem Lab Med,* 51**,** 741-51.

PLEBANI, M. 2016. Towards a new paradigm in laboratory medicine: the five rights. *Clin Chem Lab Med,* 54**,** 1881-1891.

PLEBANI, M., SCIACOVELLI, L., AITA, A. & CHIOZZA, M. L. 2014. Harmonization of pre-analytical quality indicators. *Biochem Med (Zagreb),* 24**,** 105-13.

RADIN, N. 1967. What is a standard? *Clin Chem,* 13**,** 55-76.

RANDOX. *RIQAS Haematolgy External Quality Assessment* [Online]. Available: www.randox.com/haematology-eqa/ [Accessed 27 July 2022].

RCVS. Available: https://www.rcvs.org.uk/setting-standards/practice-standards-scheme/ [Accessed January 20 2024].

RICÓS, C., ÁLVAREZ, V., PERICH, C., FERNÁNDEZ-CALLE, P., MINCHINELA, J., CAVA, F., BIOSCA, C., BONED, B., DOMÉNECH, M., GARCÍA-LARIO, J. V., SIMÓN, M., FERNÁNDEZ, P. F., DIAZ-GARZÓN, J. & GONZÁLEZ-LAO, E. 2015. Rationale for using data on biological variation. *Clin Chem Lab Med,* 53**,** 863-70.

RIDDICK, J. H., JOHNSON, C.L. 1972. Hematology control using computerized range charts. *Laboratory Medicine,* 3**,** 32-33.

RISHNIW, M. & PION, P. D. 2016. Evaluation of performance of veterinary in-clinic hematology analyzers. *Vet Clin Pathol,* 45**,** 604-614.

RISHNIW, M., PION, P. D. & MAHER, T. 2012. The quality of veterinary in-clinic and reference laboratory biochemical testing. *Vet Clin Pathol,* 41**,** 92-109.

ROSE, M. P. 1998. Follicle stimulating hormone international standards and reference preparations for the calibration of immunoassays and bioassays. *Clin Chim Acta,* 273**,** 103-17.

SACCHINI, F. & FREEMAN, K. P. 2008. Quality Documentation Challenges for Veterinary Clinical Pathology Laboratories. *Journal of Veterinary Diagnostic Investigation,* 20**,** 266-273.

SANDBERG, S., FRASER, C. G., HORVATH, A. R., JANSEN, R., JONES, G., OOSTERHUIS, W., PETERSEN, P. H., SCHIMMEL, H., SIKARIS, K. & PANTEGHINI, M. 2015. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clinical Chemistry and Laboratory Medicine (CCLM),* 53**,** 833-835.

SCHNEIDER, A. J. 1960. Some thoughts on normal, or standard, values in clinical medicine. *Pediatrics,* 26**,** 973-84.

SCHOENMAKERS, C. H., NAUS, A. J., VERMEER, H. J., VAN LOON, D. & STEEN, G. 2011. Practical application of Sigma Metrics QC procedures in clinical chemistry. *Clin Chem Lab Med,* 49**,** 1837-43.

SEO, J. Y., LEE, S.-T. & KIM, S.-H. 2015. Performance evaluation of the new hematology analyzer Sysmex XN-series. *International Journal of Laboratory Hematology,* 37**,** 155-164.

SIMUNDIC, A. M., CORNES, M., GRANKVIST, K., LIPPI, G. & NYBO, M. 2014. Standardization of collection requirements for fasting samples: for the Working Group on Preanalytical Phase (WG-PA) of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). *Clin Chim Acta,* 432**,** 33-7.

SOBAS, F., BENATTAR, N., BELLISARIO, A., MARIN, S., NOUGIER, C., LIENHART, A. & NÉGRIER, C. 2010. Impact of quality control matrix effect: application to the calculation of uncertainty of measurement in one-stage clotting factor VIII assay. *Blood Coagul Fibrinolysis,* 21**,** 498-501.

SOLBERG, H. E. & STAMM, D. 1991. International Federation of Clinical Chemistry, Scientific Division: approved recommendation on the theory of reference values. Part 4. Control of analytical variation in the production, transfer and application of reference values. *Eur J Clin Chem Clin Biochem,* 29**,** 531-5.

SPARKES, A. H., CANNON, M., CHURCH, D., FLEEMAN, L., HARVEY, A., HOENIG, M., PETERSON, M. E., REUSCH, C. E., TAYLOR, S. & ROSENBERG, D. 2015. ISFM consensus guidelines on the practical management of diabetes mellitus in cats. *J Feline Med Surg,* 17**,** 235-50.

STEPHAN, C., KLAAS, M., MÜLLER, C., SCHNORR, D., LOENING, S. A. & JUNG, K. 2006. Interchangeability of measurements of total and free prostate-specific antigen in serum with 5 frequently used assay combinations: an update. *Clin Chem,* 52**,** 59-64.

STOCKHAM, S. L. & SCOTT, M. A. 2008. *Fundamentals of Veterinary Clinical Pathology*, Wiley.

STÖCKL, D., DEWITTE, K. & THIENPONT, L. M. 1998. 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,* 44**,** 2340-6.

STURGEON, C. M., BERGER, P., BIDART, J.-M., BIRKEN, S., BURNS, C., NORMAN, R. J., STENMAN, U.-H. K. & HCG, T. I. W. G. O. 2009. Differences in Recognition of the 1st WHO International Reference Reagents for hCG-Related Isoforms by Diagnostic Immunoassays for Human Chorionic Gonadotropin. *Clinical Chemistry,* 55**,** 1484-1491.

TECLES, F., MUÑOZ, A., CERÓN, J. J. & FREEMAN, K. 2020. The number of replicates, and pooling versus individual measurements for analytical imprecision calculations: Does it matter? *Veterinary Clinical Pathology,* 49**,** 112-118.

THEODORSSON, E. 2016. Quality Assurance in Clinical Chemistry: A Touch of Statistics and A Lot of Common Sense. *J Med Biochem,* 35**,** 103-112.

THEODORSSON, E. & MAGNUSSON, B. 2017. Full method validation in clinical chemistry. *Accreditation and Quality Assurance,* 22**,** 235-246.

THIENPONT, L. M., VAN UYTFANGHE, K., DE GRANDE, L. A. C., REYNDERS, D., DAS, B., FAIX, J. D., MACKENZIE, F., DECALLONNE, B., HISHINUMA, A., LAPAUW, B.,

TAELMAN, P., VAN CROMBRUGGE, P., VAN DEN BRUEL, A., VELKENIERS, B. & WILLIAMS, P. 2017. Harmonization of Serum Thyroid-Stimulating Hormone Measurements Paves the Way for the Adoption of a More Uniform Reference Interval. *Clin Chem,* 63**,** 1248-1260.

TONKS, D. B. 1963. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. *Clin Chem,* 9**,** 217-33.

TRIPODI, A., LIPPI, G. & PLEBANI, M. 2016. How to report results of prothrombin and activated partial thromboplastin times. *Clin Chem Lab Med,* 54**,** 215-22.

UNSAL, I., COSKUN, A., SERTESER, M., INAL, T. C. & OZPINAR, A. 2010. Toward standardization of quality assessment in laboratory medicine by using the same matrix samples for both internal and external quality assessments. *Accreditation and Quality Assurance,* 15**,** 621-627.

VANYO, L. C., FREEMAN, K. P., MELÉNDEZ-LAZO, A., TELES, M., CUENCA, R. & PASTOR, J. 2018. Comparison of traditional statistical quality control using commercially available control materials and two patient-based quality control procedures for the ADVIA 120 Hematology System. *Veterinary Clinical Pathology,* 47**,** 368-376.

VESPER, H. W., MILLER, W. G. & MYERS, G. L. 2007. Reference materials and commutability. *Clin Biochem Rev,* 28**,** 139-47.

VESPER, H. W., MYERS, G. L. & MILLER, W. G. 2016. Current practices and challenges in the standardization and harmonization of clinical laboratory tests. *Am J Clin Nutr,* 104 Suppl 3**,** 907S-12S.

VETBIOLOGICALVARIATION.ORG. Available: https://www.vetbiologicalvariation.org/ [Accessed 21st September 2022].

WAYNE, P. 2002. *CLSI document EP9-A2* [Online]. Clinical and Laboratory Standards Institute [Accessed].

WESTGARD, J. O. 2001. *Six Sigma Quality Design and Control,* Maddison, WI, Westgard QC, Inc.

WESTGARD, J. O. 2006. *Six sigma quality design and control*, Westgard QC, Incorporated.

WESTGARD, J. O. 2013a. Statistical quality control procedures. *Clin Lab Med,* 33**,** 111-24.

WESTGARD, S. 2013b. Prioritizing Risk Analysis Quality Control Plans Based on Sigma-metrics. *Clinics in laboratory medicine,* 33**,** 41-53.

WESTGARD, J.O 2016a. *Basic QC Practices: Training in Statistical Quality Control for Medical Laboratories,* Madison, WI, Westgard QC.

WESTGARD, J.O 2016b. *Implementing repeat patient test controls. In: Basic QC Practices: Training in Statistical Quality Control for Medical Laboratories,* Madison, WI, Westgard QC.

WESTGARD, J. O. 2016c. Developing a QA program to monitor the QC plan. *In:* WESTGARD, J. O. (ed.) *Basic QC Practices Training in Statistical Quality Control for Medical Laboratories.* 4th ed. Madison, WI: Westgard QC.

WESTGARD, J. O. 2016d. Useful measures and models for analytical quality management in medical laboratories. *Clinical Chemistry and Laboratory Medicine (CCLM),* 54**,** 223-233.

WESTGARD, J. O. & BARRY, P. L. 1986. Cost-effective quality control: managing the quality and productivity of analytical processes. *AACC Press*.

WESTGARD, J. O. & BARRY, P. L. 2008. *Basic method validation,* Madison, WI, Westgard QC.

WESTGARD, J. O., BARRY, P. L., HUNT, M. R. & GROTH, T. 1981. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem,* 27**,** 493-501.

WESTGARD, J. O., CAREY, R. N. & WOLD, S. 1974. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem,* 20**,** 825-33.

WESTGARD JO, K. G. 1996. *Quality Assurance. Chapter in Fundamentals of Clinical Chemistry* Philadelphia, WB Saunders Company

WESTGARD, J. O. & WESTGARD, S. A. 2015. Assessing quality on the Sigma scale from proficiency testing and external quality assessment surveys. *Clin Chem Lab Med,* 53**,** 1531-5.

WESTGARD, J. O. & WESTGARD, S. A. 2016. Quality control review: implementing a scientifically based quality control system. *Ann Clin Biochem,* 53**,** 32-50.

WESTGARD, J. O. & WESTGARD, S. A. 2017. Six Sigma Quality Management System and Design of Risk-based Statistical Quality Control. *Clinics in Laboratory Medicine,* 37**,** 85-96.

WESTGARD, S., BAYAT, H. & WESTGARD, J. O. 2018. Analytical Sigma metrics: A review of Six Sigma implementation tools for medical laboratories. *Biochem Med (Zagreb),* 28**,** 020502.

WESTGARD QC. 2015. *Quality Requirements* [Online]. Westgard QC Available: https://www.westgard.com/rilibak.htm [Accessed 19th September 2022].

WEYKAMP, C. 2013. HbA1c: a review of analytical and clinical aspects. *Ann Lab Med,* 33**,** 393-400.

WEYKAMP, C., WIELDERS, J., HELANDER, A., ANTON, R. F., BIANCHI, V., JEPPSSON, J. O., SIEBELDER, C., WHITFIELD, J. B. & SCHELLENBERG, F. 2014. Harmonization of measurement results of the alcohol biomarker carbohydrate-deficient transferrin by use of the toolbox of technical procedures of the International Consortium for Harmonization of Clinical Laboratory Results. *Clin Chem,* 60**,** 945-53.

WHEELAN, C. 2013. *Naked Statistics: Stripping the Dread from the Data*, W. W. Norton.

WHITBREAD, T. J. 2019. *Laboratory Tests Routinely Performed in Veterinary Medicine* [Online]. Merck Manal, Veterinary Manual Available: https://www.merckvetmanual.com/special-pet-topics/diagnostic-tests-and-imaging/laboratory-tests-routinely-performed-in-veterinary-medicine#top [Accessed 27/02/2024].

WOOTTON, I. D. P., KING, E. J. & SMITH, J. M. 1951. The quantitative approach to hospital biochemistry; normal values and the use of biochemical determinations for diagnosis and prognosis. *British medical bulletin,* 7**,** 307-311.

WRIGHT, M. E., CROSER, E. L., RAIDAL, S., BARAL, R. M., ROBINSON, W., LIEVAART, J. & FREEMAN, K. P. 2019. Biological variation of routine haematology and biochemistry measurands in the horse. *Equine Vet J,* 51**,** 384-390.

YAGO, M. & ALCOVER, S. 2016. Selecting Statistical Procedures for Quality Control Planning Based on Risk Management. *Clin Chem,* 62**,** 959-65.

ZANINOTTO, M., TASINATO, A., PADOAN, A., VECCHIATO, G., PINATO, A., SCIACOVELLI, L. & PLEBANI, M. 2012a. Effects of sample transportation on commonly requested laboratory tests. *Clin Chem Lab Med,* 50**,** 1755-60.

ZANINOTTO, M., TASINATO, A., PADOAN, A., VECCHIATO, G., PINATO, A., SCIACOVELLI, L. & PLEBANI, M. 2012b. An integrated system for monitoring the quality of sample transportation. *Clin Biochem,* 45**,** 688-90.

ZHANG, C., ZHAO, H., WANG, J., ZENG, J. & WANG, Z. 2015. The Application of Six Sigma Techniques in the Evaluation of Enzyme Measurement Procedures in China. *Clin Lab,* 61**,** 461-5.

Note: American spelling follows in this section.

APPENDIX A: PUBLICATION OF CHAPTER 2

Title: Evaluation of Sysmex XT-2000*i*V analyzer performance across a network of five veterinary laboratories using a commercially available quality control material

Reference: This chapter has been published as Daly SC, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000*i*V analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021; 50: 568–578. doi:10.1111/vcp.13016

Check for updates

Received: 18 January 2021 | Revised: 12 March 2021 | Accepted: 17 March 2021

DOI: 10.1111/vcp.13016

ORIGINAL ARTICLE

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

Evaluation of Sysmex XT-2000*i*V analyzer performance across a network of five veterinary laboratories using a commercially available quality control material

Susan Daly[1] | Kathleen P. Freeman[2] | Peter A. Graham[3]

[1]SYNLAB-VPG/Cork, Cork, UK

[2]SYNLAB –VPG/Exeter, Exeter, UK

[3]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, Brushville House, Dosco Business Park, South Douglas Road, Cork T12XR6V, UK.
Email: susan.daly@synlab.co.uk

Abstract

Background: Laboratory and instrument harmonization is seldom reported in the veterinary literature despite its advantages to clinical interpretation, including the use of interchangeable results and common reference intervals within a system of laboratories.

Objectives: A three-step process was employed to evaluate and optimize performance and then assess the appropriateness of common reference intervals across a network of six Sysmex XT-2000*i*V hematology analyzers at 5 commercial veterinary laboratory sites. The aims were to discover if harmonization was feasible in veterinary hematology and which quality parameters would best identify performance deviations to ensure a harmonized status could be maintained.

Methods: The performance of 10 measurands of a commercially available quality control material (Level 2—Normal e-CHECK (XE)-Hematology Control) was evaluated during three 1-month time periods. Precision and bias were assessed with Six Sigma, American Society of Veterinary Clinical Pathology (ASVCP) total error quality goals and biologic variation (BV)-based quality goal approaches to performance measurement.

Results: Instrument adjustments were made to 1 analyzer twice and 3 analyzers once between evaluations to improve performance and achieve harmonization. Sigma metrics improved from 37/50 > 6 to 58/60 > 6 and to all >5 over the course of the harmonization project. BV-based quality goals for desirable bias and for laboratory systems of 0.33 × CV_I (within-subject biologic variation) were more sensitive and useful for assessing performance than the ASVCP total error goals.

Conclusions: Optimization and harmonization were achieved, and because BV-derived bias goals were achieved, common reference intervals could be implemented across the network of analyzers.

KEYWORDS
analytical performance, biologic variation, hematology, quality system, sigma metrics, Sysmex, veterinary

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Veterinary Clinical Pathology* published by Wiley Periodicals LLC on behalf of American Society for Veterinary Clinical Pathology.

Appendix A

1 | INTRODUCTION

Interlaboratory or intralaboratory comparisons across multiple instruments in human clinical laboratories are well documented[1-4]; however, a few similar evaluations are found in the veterinary literature.[5] This could be attributed to a multitude of factors, including that human laboratory regulations are not applied to veterinary laboratories[5,6]; the results of comparative external quality assurance (EQA) performances are not routinely available and/or shared, and there is a lack of reporting of efforts or possibly an absence of efforts to harmonize results within or between veterinary laboratory businesses and networks. Harmonization of methods across laboratories means the standardization of results across methods and laboratories when no certified reference material is available. It is based on an agreed-upon standard for performance among participating laboratories to ensure that results have sufficient uniformity to be used interchangeably across a laboratory network for an individual patient.[3]

Synlab Veterinary Pathology Group (Synlab-VPG) operates six Sysmex XT-2000*i*V analyzers within their network of five laboratories located in the UK and Ireland, providing an environment in which to explore a veterinary hematology harmonization approach for the benefit of a group of veterinary clinical pathologists, client veterinarians, and ultimately, patients. Were a successful approach identified, it could provide a useful template to implement a harmonization process for the veterinary hematology community.

At the core of the harmonization process is a set of quality goals that are used to measure acceptable performance and progress toward harmonization. Such measures include error goals for veterinary hematology provided by the American Society for Veterinary Clinical Pathology (ASVCP)[7] or from internal expert opinion based on QC validation; canine biologic variation (BV) goals for bias and CV, and the canine BV goal for a laboratory system ($0.33 \times CV_I$).[8] Sigma metrics is also a useful indicator of performance where methods >6 sigma are considered to have "World Class" performance.[9]

The objectives of this study were to:

1. Evaluate the performance of the six Sysmex hematology analyzers within the five SYNLAB-VPG laboratories using a commercially available quality control material.
2. Optimize the performance of each analyzer to ensure the total error budget is contained within ASVCP guidelines or those determined by internal expert opinion.
3. Determine if BV goals for CV and bias and/or laboratory system performance would enable the use of common reference intervals.
4. Determine those quality goals that would be the most useful for identifying performance deviations that need to be addressed to maintain the interchangeability of results across the laboratory system.

We hypothesized that performance across laboratories would be similar since the same instruments were used, as well as the same reagents and quality control materials. We additionally hypothesized that the quality goals or a minimal deviation from the goals would be achievable.

2 | MATERIALS AND METHODS

2.1 | Analyzers

Six Sysmex XT-2000*i*V hematology analyzers (Sysmex Corporation) using veterinary software version 00-13 were compared among five veterinary sites within SYNLAB-VPG. These were designated as instruments one to five (site 4 has two instruments, designated as 4A and 4B). Ten measurands were selected for statistical quality performance analysis, as we considered them the most clinically relevant in current veterinary practice: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration (MCHC), white blood cell count (WBC), platelet count (PLT), red cell distribution width-coefficient of variation (RDW-CV), plateletcrit (PCT), and reticulocyte number (RETIC). Additional measurands within the capability of the analyzer and represented in the quality control material (QCM) package insert were not included in the performance evaluation. The methods used to determine these measurands are summarized in Table 1 (SEED 2012[10] and Sysmex-2000*i*V/XT -1800*i*V user manual).

The Sysmex XT-2000*i*V hematology analyzer can be operated in open or closed mode; for this study, an open mode was used, ie, the probe aspirates the required volume from an open uncapped tube.

2.2 | Quality control material

A single level of QCM (level 2—Normal e-CHECK (XE)-Hematology Control) was analyzed once per day at instrument start-up, prior to the analysis of patient samples, by a fully trained technician according to a standard operating procedure. The QCM was refrigerated when not in use, and opened vials were used within 7 days. Performance was deemed acceptable for inclusion in the study data when the daily internal quality control (IQC) value was within the manufacturer's recommended acceptable range.

The manufacturer's target mean for the QCM was chosen for comparison across the group of analyzers.

2.3 | Data collection

Over three separate evaluation periods of November (Month 1), December 2019 (Month 2), and March 2020 (Month 5), each of the laboratories collected approximately 30 consecutive days of QCM result data from six Sysmex XT-2000*i*V analyzers. Only four sites (five analyzers) could participate in the Month 1 period due to an inability to obtain the Level 2 QCM. The data were exported to Microsoft Excel (Microsoft Excel for Mac 2011, Version 14.7.7 [170905] Last update 14.7.7). Each laboratory sent individual spreadsheets for analysis, and the data were used to calculate quality evaluation metrics. Intervention recommendations such as

19391655, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [23/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

TABLE 1 Methods[10] used by the Sysmex XT-2000iV hematology analyzer for determination of the hematologic measurands

Measurand (abbreviation) (units)	Method
Red blood cell count (RBC) ($\times 10^6/\mu L$)	Impedance direct-current detection with hydrodynamic focusing (number of electrical pulses generated per volume of blood)
Hemoglobin (HGB) (g/dL)	Cyanide-free, spectrophotometric method
Hematocrit (HCT) (%)	Impedance direct-current using cumulative pulse height
Mean cell volume (MCV) (fL)	Calculation: MCV (fL) = HCT/RBC
Mean cell hemoglobin concentration (MCHC) (g/dL)	Calculation: MCHC (g/dL) = HGB/HCT
Red cell distribution width–coefficient of variation (RDW–CV) (%)	Histogram generation from impedance
Reticulocyte Count (RETIC NUMBER) ($\times 10^6/\mu L$)	Fluorescence flow cytometry
White blood cell count (WBC) ($\times 10^3/\mu L$)	WBC-BASO channel method for total WBC count using forward and side scatter light (SCFSC)
Platelet count (PLT) ($\times 10^3/\mu L$)	Impedance with hydrodynamic focusing
Plateletcrit (PCT) (%)	Impedance with hydrodynamic focusing

optimal QCM handling and instrument adjustments and/or service were applied between each evaluation period based on the results obtained and the likely root causes for the detected deviations in performance.

2.4 | Quality evaluations

To evaluate the observed analytical imprecision and bias, the SD and CV[11] bias in units and absolute bias %[12] from the manufacturer's QCM target mean were calculated as described previously[13] for all 10 measurands from each analyzer for each evaluation period.

From the initial calculations for imprecision (CV%) and bias (absolute bias %), the total error observed (TE_{obs}) was calculated and expressed as a percentage to allow comparability with analytical quality requirements[14] using the following formula[7,14];

$$TE_{obs}\% = |\text{Bias }\%| + 2 \times CV\%$$

A sigma metric approach was taken to estimate the clinical performance of the 10 measurands across all sites using the following formula,[9,15] and Total Error Allowed (TEa) from Table 2.

$$\text{Sigma metric} = (TEa\% - |\text{Bias}\%|)/CV\%$$

The sigma metric was chosen as an important reflection of analytical capability, allowing a comparison across methods. It is considered to be a reflection of risk associated with an analytical process and correlates with the "Westgard Sigma Rules."[16,17] According to Westgard, measurands with 6 sigma or greater performance are easily controlled with a simple quality control (QC) rule and low numbers of QCM data, while those with 5-6 sigma can be controlled with a short multirule ($1_{3s}/2_{2s}/R_{4s}$) and low numbers of QCM data. Those methods with 4-5 sigma performance can be controlled with the addition of a fourth rule to a multirule ($1_{3s}/2_{2s}/R_{4s}/4_{1s}$). Those methods with 3-4 sigma performance require an even more complex multirule with larger numbers of QCM data but still may not achieve desirable levels of error detection.[12] The sigma metric, by definition, will change if performance changes, reflecting fluctuations in the observed CV and bias over defined time intervals.

For measurands with a sigma metric <6, a quality goal index (QGI) was calculated using the following formula:

$$QGI = |BIAS\%|/(1.5 * CV\%).$$

This calculation allowed us to determine whether imprecision and/or inaccuracy affected the overall quality goals using the following interpretative guidance: a QGI < 0.8 was considered indicative of imprecision while a QGI > 1.2 was indicative of inaccuracy (bias). QGI results between these two boundaries indicated both bias and imprecision contributing to <6 sigma performance.[18]

BV goals were applied as three additional performance indicators: BV-based CV, BV-based bias, and a BV-based laboratory system goal.[8] Canine BV goals were chosen because canine specimens are the most common submissions across all of the five laboratories. We calculated optimal, desirable, and minimum values for TE, CV, and bias goals based on canine BV data[14] for applicable measurands and $0.33 \times CV_I$ as the goal for multiple instruments within a laboratory system.[8] The $0.33 \times CV_I$ goal allowed us to detect differences between the means of serial measurements for each measurand to ensure comparability by checking to make sure the allowable difference was $<0.33 \times CV_I$.[8] The $\pm 0.33 \times CV_I$ goal was compared with the manufacturer's target mean for the QCM.

PCT and RET did not have BV data available.

We used the following formulas[8]

BV CV goals:

$$\text{Optimal CV: } <0.25\,(CV_I)$$
$$\text{Desirable CV: } <0.5\,(CV_I)$$
$$\text{Minimally acceptable CV: } <0.75\,(CV_I)$$

BV bias goals:

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Quality goals for the evaluation of the performance of Sysmex XT-2000iV hematology analyzers for selected measurands

Measurand	ASVCP total error quality goals[a] or internal expert opinion goal (Bold) (%)	Canine biologic variation goals for coefficient of variation[b] (%) < Opt/Des/MA	Canine biologic variation goals for absolute bias[c] (%) < Opt/Des/MA	Canine biologic variation laboratory system goal (<0.33 × CV_I) (%)
Red blood cell count	10	1.38/2.75/4.13	0.88/1.76/2.64	1.815
Hemoglobin	10	1.48/, 2.95/4.43	0.94/1.89/2.83	1.947
Hematocrit	10	1.60/3.20/4.80	1.03/2.06/3.09	2.112
Mean cell volume	7	0.525/1.05/1.58	0.418/0.836/1.25	0.693
Mean cell hemoglobin concentration	10	0.650./1.30/1.95	0.353/0.706/1.059	0.858
Red cell distribution width–coefficient of variation	10	1.00/2.00/3.00	0.901/1.80/2.	1.32
Reticulocyte number	**40**	N/A	N/A	N/A
White blood cell count	15	3.03/6.05/9.08	2.16/4.31/6.47	3.99
Platelet count	20	3.50/7.00/10.50	2.58/5.17/7.49	4.62
Plateletcrit	**25**	N/A	N/A	N/A

Note: N/A, no information for canine biologic variation available.

Sigma metric = (TEa% − Abs Bias%)/CV%.

[a]Nabity et al (2017).

[b]Biologic variation goals. CV: Opt = optimal, Des = desirable, MA = minimally acceptable. Optimal CV: <0.25(CV_I); desirable CV: <0.5(CV_I); minimally acceptable CV: <0.75(CV_I).

[c]Biologic variation bias goals: optimal bias: <0.125 × $(CV_I^2 + CV_G^2)^{1/2}$; desirable bias: <0.25 × $(CV_I^2 + CV_G^2)^{1/2}$; minimally acceptable bias: <0.375 × $(CV_I^2 + CV_G^2)^{1/2}$.

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

$$\text{Optimal bias: } <0.125\left(CV_I^2+CV_G^2\right)^{1/2}$$

$$\text{Desirable bias: } <0.25\left(CV_I^2+CV_G^2\right)^{1/2}$$

$$\text{Minimally acceptable bias: } <0.375\left(CV_I^2+CV_G^2\right)^{1/2}$$

Within-subject BV is denoted as CV_I, while between subject variation is CV_G.

3 | RESULTS

3.1 | Sigma metrics and QGI

3.1.1 | Initial evaluation of five Sysmex XT-2000*i*V analyzers Month 1

All of the hematology measurands chosen for this study for Analyzer 1 and Analyzer 2 had sigma metrics >6. Of the remaining three analyzers, Analyzer 5 had only 7, and Analyzers 4A & 4B had only five measurands >6 sigma (Table 3a).

3.1.2 | QGI Month 1

The QGI interpretation for those measurands performing <6 sigma is summarized in Table 4. Bias was more often a concern (14/24) than imprecision (8/24). Both bias and imprecision accounted for two occurrences.

Adjustments with instrument service at Analyzers 4A, 4B, and 5 were conducted to help eliminate bias. Improved control material handling was instituted across all laboratories before additional evaluations were undertaken in Month 2.

3.1.3 | Follow-up Month 2 and Month 5 evaluation on six Sysmex XT-2000*i*V analyzers

As demonstrated in Table 3b,c, sigma metrics generally improved over the following two evaluation periods, resulting in all but two measurands at one site (Analyzer 4A) achieving >6 sigma. All sites and all measurands achieved sigma >5. An instrument service visit was required for Analyzer 3, and an additional visit was required for Analyzer 5 to achieve this improvement in performance. QGI review of the remaining two measurands on Analyzer 4A identified imprecision for PLT, and bias was implicated for RBC (Table 4).

3.2 | Biologic variation goals

3.2.1 | Initial evaluation of BV goals on five Sysmex XT-2000*i*V analyzers

Achievement against CV, bias, and 0.33 × CVi goals could only be assessed for those eight measurands for which canine BV data were available (Table 2). No BV data were available for PCT and RETIC.

Analytical CVs achieved optimal or desirable CV goals for Analyzers 1 and 2. Optimal or desirable CV goals were achieved for seven of the eight measurands for Analyzer 5, and six of the eight measurands for Analyzer 4A and Analyzer 4B (Table 5a).

Optimal or desirable bias goals based on BV were achieved for seven measurands for Analyzer 1. Analyzer 2 was able to achieve optimal or desirable bias goals based on BV for six measurands. Analyzers 4A and 4B were able to achieve optimal or desirable bias goals for two and four measurands, respectively. Analyzer 5 was able to achieve optimal or desirable bias goals for three measurands but failed to achieve minimally acceptable bias goals for all other measurands (Table 5b).

For Analyzer 1, the 0.33 × CV_I goal for laboratory system performance for QCM was achieved for seven measurands for which BV information is available. For Analyzer 2, the 0.33 × CV_I goal for lab system was achieved for all measurands except PLT. For Analyzer 4A, the goal was achieved for three of the eight measurands (WBC, MCV, MCHC) but not for the remaining measurands. For Analyzer 4B, the 0.33 × CV_I goal was achieved for five measurands but not for HGB, HCT, and RDW-CV. For Analyzer 5, the 0.33 × CV_I goal was achieved for two out of eight measurands (MCV, MCHC) (Table 5c).

3.2.2 | Follow-up evaluation of BV goals on six Sysmex XT-2000*i*V analyzers in months 2 and 5

Analyzer 3, included in Month 2, achieved optimal or desirable CV goals based on BV for six of the eight measurands and all eight measurands in Month 5. The others sites achieved optimal or desirable CV goals for the eight measurands at both follow-up evaluations (Table 5a).

Analyzer 1 achieved optimal or desirable bias goals for the same seven measurands for Month 2 and Month 5. Analyzer 2 achieved optimal or desirable bias goals for eight measurands for Month 2, and 6 measurands for Month 5. Analyzer 3 achieved optimal or desirable bias goals for five measurands for Month 2, and all eight measurands for Month 5.

Analyzers 4A and 4B improved during Month 2, and Analyzer 4A met optimal or desirable bias goals for seven measurands where biologic data were available for Months 2 and 5. Analyzer 4B achieved optimal or desirable bias goal for all eight measurands in Month 2, but only achieved six of eight measurands in Month 5.

Analyzer 5, in contrast to Month 1, achieved optimal or desirable bias goal for seven measurands for Month 2, and six measurands for Month 5 (Table 5b).

The 0.33 × CV_I goal was achieved by Analyzer 1 for six measurands; it did not achieve this goal for two measurands in Month 2, but this resolved by Month 5 when the goal was achieved for eight measurands. Analyzer 2 could not achieve the goal for 2 measurands in Month 2, but in Month 5 achieved this goal for six of eight measurands (but not for HCT and MCHC).

Analyzer 3 achieved the goal for 5 of 8 measurands in Month 2. This increased to 6 of 8 measurands in Month 5.

19391658, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

TABLE 3 Summary of sigma metrics for selected measurands for six Sysmex XT-2000iV hematology analyzers within a system of veterinary laboratories over three periods of evaluation

	Analyzer					
Measurand	1	2	3	4a	4b	5
(a) Initial Month 1						
WBC	6.54	10.64	N/A	9.76	9.38	8.35
RBC	15.21	13.00	N/A	3.88	5.51	7.10
HGB	7.24	23.00	N/A	6.00	2.82	5.57
HCT	9.38	12.70	N/A	3.48	4.92	4.08
MCV	11.50	12.20	N/A	12.34	12.40	8.15
MCHC	9.28	10.90	N/A	5.65	9.37	8.12
PLT	6.15	9.05	N/A	7.35	9.7	16.97
RDW-CV	18.00	25.36	N/A	0.47	0.596	2.03
PCT	6.06	8.56	N/A	6.50	7.75	7.00
Retic. No.	6.20	7.63	N/A	3.20	4.57	7.03
N σ > 6	10	10	N/A	5	5	7
N σ 5.0-5.9	0	0	N/A	1	1	1
N σ 4.0-4.9	0	0	N/A	0	2	1
N σ 3.0-3.9	0	0	N/A	3	0	0
σ N < 3	0	0	N/A	1	2	1
Intervention	None	None	N/A	ISA	ISA	ISA
	Improved QCM handling					
(b) Follow-up Month 2						
WBC	14.12	6.81	8.87	6.80	7.13	9.53
RBC	15.11	14.77	11	13.96	12.10	17.75
HGB	19.94	12.68	12.82	14.83	13.54	20.74
HCT	2.92	13.14	2.20	9.19	12.77	10.94
MCV	16.68	13.44	5.20	12.74	12.83	8.02
MCHC	14.13	12.05	5.71	10.15	9.53	11.32
PLT	8.07	7.62	4.94	9.73	6.29	6.94
RDW-CV	20.40	20.40	13.64	18.68	17.71	24.71
PCT	7.61	9.92	4.55	9.01	6.43	7.06
Retic. No.	5.59	6.48	7.83	5.59	7.10	5.04
N σ > 6	8	10	5	9	10	9
N σ 5.0-5.9	1	0	2	1	0	1
N σ 4.0-4.9	0	0	2	0	0	0
N σ 3.0-3.9	0	0	0	0	0	0
σ N < 3	1	0	1	0	0	0
Intervention	None	None	ISA	None	None	ISA
(c) Follow-up Month 5						
WBC	7.62	7.78	6.41	9.22	7.97	10.34
RBC	14.22	18.09	14.2	5.24	12.01	9.34
HGB	15.97	24.58	30.04	17.50	18.45	10.08
HCT	13.82	13.92	9.05	15.05	16.10	7.85
MCV	13.60	14.91	10.07	11.12	10.67	8.43
MCHC	13.50	13.12	8.51	10.57	12.55	10.70
PLT	7.60	9.05	6.80	5.12	7.30	7.65

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 (Continued)

	Analyzer					
Measurand	1	2	3	4a	4b	5
RDW-CV	19.17	13.10	19.53	14.40	16.52	14.55
PCT	7.42	7.57	7.35	7.9	7.9	8.04
Retic. no.	6.90	8.00	8.50	7.60	6.00	8.35
N σ > 6	10	10	10	8	10	10
N σ 5.0-5.9	0	0	0	2	0	0
N σ 4.0-4.9	0	0	0	0	0	0
N σ 3.0-3.9	0	0	0	0	0	0
σ N < 3	0	0	0	0	0	0

Abbreviations: 1, November; 2, December; 3, March; HCT, hematocrit; HGB, hemoglobin; ISA, Instrument service and Adjustment; MCHC, mean corpuscular hemoglobin concentration; MCV, mean cell volume; N, number; N/A, not available due to inability to source QCM; PCT, plateletcrit; PLT, platelet; QCM, quality control material; RBC, red blood cell; RDW–CV, red cell distribution wide–coefficient of variation; WBC, white blood cell.

Analyzer 4A achieved the goal for 7 of 8 measurands in Month 2 but by Month 5 achieved goals for all 8 measurands. Analyzer 4B achieved the goal for 6 measurands in Month 2 and achieved the goal for 7 of 8 measurands in Month 5, except for HCT.

Analyzer 5 achieved the goal for 6 of 8 measurands in Month 2. In Month 5, 7 measurands achieved this goal, apart from MCHC (Table 5c).

3.3 | Comparison of performance measures

Tables 5a, 5b, and 5c summarize the BV goals for bias, CV, and 0.33 × CV_I goal for laboratory system performance for QCM that reached optimal or desirable acceptable goals for 8 measurands where BV data are available, assessed during three evaluations. BV goals for bias and the 0.33 × CV_I goal for laboratory systems proved discriminatory when compared with the ASVCP goals, or expert opinion where sigma metrics was <6 sigma.

4 | DISCUSSION

This is the first report comparing six Sysmex analyzers in a network of five veterinary laboratories in veterinary peer-reviewed literature. We hypothesized that performance across the laboratories would be similar and would have minimal deviations from quality goals. This was true for some sites, but this study also demonstrated that 4 of the 6 analyzers initially exhibited differences in performance, despite initial harmonization efforts, ie, using the same instruments, reagents, QCM, processing frequency of QCM, and the same technical training. This study demonstrated that at certain times, Sysmex hematology analyzers might fail to meet quality requirements,[13] particularly for the erythroid measurands.

To monitor any degree of individuality for a network of Sysmex analyzers, we found sigma metrics to be a valuable indicator of performance-related issues. We decided to use this approach for a number of reasons. It is a universal benchmark that lends itself well to multiple site evaluation and continuous comparison. It incorporates both bias and imprecision, and as sigma metric values decrease, failure of the QC rule increases.[19] We also used quality goals based on BV to minimize analytical noise in serial results and ensure that bias and/or imprecision would not mask physiological changes.[8,20] This might also be useful for ongoing patient monitoring and the continued use of common RI within this network of laboratories.

We calculated 170 sigma metrics from 10 measurands (Table 3a-c), and from those, 24 (14%) did not achieve >6 sigma for "world-class" performance. For these, the QGI was useful for analyzing potential problems with those measurands by predicting whether the underperformance was due to imprecision, bias, or both, and informing the choice of intervention and adjustment remedies.

Bias was more likely to be a contributing factor than imprecision based on the QGI (58% of <6 sigma occurrences). The high degree of bias in comparison with imprecision was not unexpected, as managing bias is a regular occurrence in a clinical laboratory, but we needed to ascertain how the range of bias affected the sigma metric value. Where bias was problematic, more than half of the occurrences did not meet the minimally acceptable BV bias goals and/or the 0.33 × CV_I goal for QCM. However, the inability to achieve the minimally acceptable BV bias goal was not unique to measurands with <6 sigma. We found examples where BV bias goals were not achieved but sigma metrics were >6. This was because the bias was much lower than the total allowable error goal (based on ASVCP or expert opinion). When bias was a larger proportion of the TEa, the sigma value was <6; for these measurands, the bias ranged from 3.22% to 8.27%. Our evaluation suggests that bias >3% on the Sysmex XT-2000*i*V analyzer indicates an increased likelihood that the sigma metric value will be below "world-class" performance and have an increased likelihood to fail at least one or, in most cases, two of the quality goals (Desirable BV bias and 0.33 × CV_I goals) of the three biologic variation-based goals. For these occurrences, the imprecision varied from 0.89 to 3.61 (CVobs). Additional performance evaluations are needed to confirm the suggestion from these results that observed bias results >3.0% could be used to determine

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

TABLE 4 Summary of the investigation of measurands with sigma metrics <6 for 6 Sysmex XT-2000iV hematology analyzers in a system of veterinary laboratories using commercially available quality control material

Site	Measurand	Sigma metric	QGI	Interpretation of QGI	CVobs %	Achieved CV_A BV goal	Bias-obs %	Achieved bias BV goal	Achieved 0.33 × CV_I goal
1	RETIC	5.59	0.69	Imprecision	6.04	N/A	6.253	N/A	N/A
	HCT	2.92	1.53	Bias	3.18	YES - Des	0.72	YES - Opt	YES
3	HCT	2.20	6.56	Bias	3.08	YES - Des	3.20	NO	YES
	MCV	5.20	0.16	Imprecision	1.31	YES - MA	0.19	YES - Opt	YES
	MCHC	5.71	2.21	Bias	1.30	Yes- MA	2.542	NO	NO
	PLT	4.94	2.597	Bias	3.84	YES - Des	1.01	YES - Opt	YES
	PCT	4.55	0.04	Imprecision	5.42	N/A	0.32	N/A	N/A
4a	PLT	5.12	0.17	Imprecision	3.72	YES - Des	0.93	YES - Opt	YES
	RBC (Month 1)	3.88	1.22	Bias	1.75	YES - Des	3.22	NO	NO
	HCT	3.48	1.32	Bias	1.70	YES - Des	4.06	NO	NO
	RDW-CV	0.47	1.52	Bias	3.61	NO	8.27	NO	NO
	RET	5.59	0.77	Imprecision	5.92	N/A	6.87	N/A	N/A
	RBC (Month 5)	5.24	6.67	Bias	1.60	YES - Des	1.60	YES - Des	NO
	MCHC	5.65	0.56	Imprecision	1.43	YES - MA	1.87	NO	YES
	RET	3.20	1.13	Both imprecision and bias	8.15	N/A	13.90	N/A	N/A
4b	RBC	5.51	0.55	Imprecision	1.57	YES - Des	1.32	YES - Des	YES
	HGB	2.82	1.66	Bias	1.88	YES - Des	4.69	NO	NO
	HCT	4.92	0.99	Both Imprecision and bias	1.56	NO	2.33	YES - MA	NO
	RDW-CV	0.60	1.75	Bias	3.10	NO	8.15	NO	NO
	RET	4.75	53.71	Bias	5.61	N/A	14.37	N/A	N/A
5	HGB	5.57	3.77	Bias	0.89	YES - Opt	5.04	NO	NO
	HCT	4.08	1.72	Bias	1.49	YES - Opt	3.88	NO	NO
	RETIC	5.04	0.29	Imprecision	7.29	N/A	3.217	N/A	N/A
	RDW-CV	2.03	27.75	Bias	1.61	YES- Des	6.71	NO	NO

Note: Tools used for the investigation included Quality Goal Index and Relationships of Observed Coefficient of Variation, Observed Bias, and Canine Biologic Variation-based Quality Goals.

QGI, Quality Goal Index (%) = Abs Bias (%)/1.5 CV (%), Interpretation: The QGI reflects the extent to which both bias and precision meet their respective quality goals without considering goals for CV and bias as separate entities. The QGI is interpreted as follows: QGI <0.8 = imprecision and QGI >1.2 = inaccuracy, whereas QGI of 0.8-1.2 = both imprecision and inaccuracy.

Abbreviations: Bias-obs %, observed analytical absolute bias (%); BV, biologic variation; CV, within-subject biologic variation coefficient of variation; CV_A, analytical coefficient of variation; CVobs %, observed analytical coefficient of variation; Des, desirable; HCT, hematocrit; HGB, hemoglobin; MA, minimally acceptable; MCV, mean cell volume; N/A, not available for this measurand; Opt, optimal; RBC, red blood cell count; RDW-CV, red cell distribution width-coefficient of variation; RET, reticulocyte number.

emerging or established problems and whether instrument servicing is indicated when these occur.

The 0.33 × CV_I goal was not achieved for 11 out of 18 occurrences with the sigma metric <6; nine of these were during the initial phase, with one occurrence during the second phase and one during the third phase. Ten occurrences were attributed to bias, and one was due to both imprecision and bias. The ongoing attention to QC, QCM handling, and emphasis on QCM performance, as well as instrument servicing when excessive bias and/or low sigma metrics (<5) were identified, likely contributed to the increasing ability in achieving the BV-based laboratory systems goal over the three periods of evaluation. Continued monitoring of performance is needed to determine if this improved level of performance will be sustained. The 0.33 × CV_I goal proved to be a good discriminatory goal for poor performance based on comparison with the QCM target.

From the initial evaluation, it was clear that analyzers 4A & 4B and Analyzer 5 required attention from Sysmex technical support, as nearly half of the measurands had sigma metrics <6 sigma during this phase and for the first assessment for Analyzer 3 in Month 2 where half of their measurands were <6 sigma. This was seen particularly in the erythroid measurands and following a Sysmex service call. We

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

TABLE 5A Summary of CV biologic variation goals that reached optimally or desirably acceptable levels for eight measurands where biologic variation data were available and assessed during three evaluations

	CV BV goal		
Analyzer number	Month 1	Month 2	Month 5
1	8	8	8
2	8	8	8
3	N/A	6	8
4A	6	8	8
4B	6	8	8
5	7	8	8

Note: Not applicable (N/A) is for an analyzer that was not using correct quality control material during the first evaluation.

Abbreviation: CV BV, coefficient of variation biologic variation.

TABLE 5B Summary of bias biologic variation goals that reached optimally or desirably acceptable for eight measurands where biologic variation data are available and assessed during three evaluations

	Bias BV goal		
Analyzer number	Month 1	Month 2	Month 5
1	7	7	7
2	6	8	6
3	N/A	5	8
4A	2	7	7
4B	4	8	6
5	3	7	6

Note: N/A is for an analyzer that was not using correct quality control material during the first evaluation.

Abbreviation: BV, biologic variation.

TABLE 5C Summary of $0.33 \times CV_I$ biologic variation goal for multiple instruments within a laboratory system that reached optimal or desirable, acceptable goals for eight measurands where biologic variation data are available and assessed during three evaluations

	$0.33 \times CV_I$ BV goal		
Analyzer number	Month 1	Month 2	Month 5
1	7	6	8
2	7	6	6
3	N/A	5	6
4A	3	7	8
4B	5	6	7
5	2	6	7

Note: Not applicable (N/A) is for an analyzer that was not using the correct quality control material during the first evaluation.

Abbreviation: BV, biologic variation.

saw improvements in those measurands in terms of sigma metrics which were now >5 sigma, and many were >6 sigma on subsequent evaluations. The performance variations and need for bias corrections were discussed with the service representative prior to making adjustments.

We standardized QCM handling among the sites, ensured routine maintenance procedures were adhered to and documented in the analyzer maintenance log, and improved the technicians' understanding of QC evaluation for identifying shifts, trends, and out of control results. Following these approaches, we observed improved sigma metrics across the three phases of QC data collection.

From this study, sigma metrics <5, whether attributed to bias or imprecision, did not meet all quality goals that were set. Sigma metrics >6 had better capabilities for reaching the BV goals and continuing improvements as these measurands had fewer QCM failures. Sigma metrics <4 were also unacceptable as these measurands could only achieve 1 out of the 3 BV quality goals and from our study required technical attention.

This raises the question as to whether there is a hierarchy of quality goals that should be used to achieve harmonization. We found the choice of BV goals used in this study to be sufficient for challenging the performance of the measurands. We also found that the TE or expert opinion goals were not discriminatory since all analyzers were able to achieve these goals, even when observed performance and sigma varied widely. Collected bias and imprecision data were more usefully applied to check the achievement of the BV goals rather than TE or expert opinion goals that were not major contributors to the improvement process. BV goals were superior, in that once desirable BV-based CV goals were met we could be confident that the patient results would not be misclassified clinically, and imprecision would not mask physiological change, which is reassuring for continued patient monitoring. We could also be assured that we could continue to use common reference intervals as long as we could meet the minimally acceptable BV bias goal.[8,21]

A limitation to using BV is that it varies with measurand. Thus, the permissiveness or stringency of the BV quality goals varies according to measurand[14]; however, a single numerical quality goal is not suitable across all measurands, which have different interpretations and clinical utilities, as well as differences in analytical performance. Despite this limitation, our study demonstrated that we could achieve optimal or desirable BV quality goals for most measurands and meet minimally acceptable goals for some measurands with the accumulated data from the three evaluations. From our final evaluation (Month 5) after the recommended improvements from the first two datasets, we met these BV goals most of the time. Where we did not meet the BV goals (25 occurrences across all evaluations <6 sigma), the most serious concern was the risk of masking within-patient changes due to analytical variation. Ongoing evaluations are needed to determine if sigma metrics alone or in combination with absolute observed bias and BV-based goals (especially $0.33 \times CV_I$ goal) can be reliably used on an ongoing basis for monitoring performance.

19391676, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

This study demonstrates that these quality specifications are achievable and that harmonization is possible. However, fluctuations in the ability to achieve these goals and in the sigma metrics were seen throughout the three evaluations for some sites, while others remained relatively consistent. These differences could be due to inherent analytical variability between the analyzers, even though all were of the same model by the same manufacturer. They could also be due to the capability and training of the operators and environmental conditions at each site, and the degree of quality management engagement and varying analyzer throughput could also have contributed. A publication from Sysmex[10] also made reference to this by quoting analyzer differences from the same manufacturer will still have some disparity and also indirectly highlighted the need to further scrutinize calculated measurands. MCHC is derived from HGB and HCT values,[10] and so its validation must be compared with their performances; for example, when MCHC was <6 (Table 3b, Analyzer 3), the corresponding Sigma metrics for HCT and HGB were 2.20 and 12.82, respectively, indicating that the poor MCHC performance could likely be attributed to poor HCT performance.

From our experience to date, in conjunction with BV bias and the 0.33 × CV_I goal, we recommend using a sigma-based quality management system as the method to monitor and compare performance in a network of laboratories; however, this should be carried out on a continuous and routine basis to detect any changes in performance. The application of sigma metrics to assess and compare instruments, seen for chemistry and immunoassays methods in human laboratories, demonstrated that a sigma metric <4 would cause QC rules to fail 20 times more frequently than methods with a 6 sigma metric.[19] Additionally, for the better-performing methods, the cost of supplies and reagents was significantly lower.[4,19] More specifically, a study of sigma metrics in human laboratories using the Sysmex XN and non-XN series (not our exact model) from a network of laboratories found sigma metrics consistently >4 with the exception of HCT, which had a number of <4 sigma occurrences.[4] On review of our sigma metrics, HCT also had the most <4 sigma occurrences. This was challenging to compare with other studies as HCT is measured on Sysmex analyzers and calculated on most other hematology analyzers.

5 | CONCLUSION

The results from this study show that we could evaluate the performance of the Sysmex XT-2000*i*V across a network of laboratories and achieved harmonization and optimization of procedures. We determined that the total error budget is indeed contained within ASVCP guidelines or internal expert opinion goals. This harmonization process demonstrated that sigma metrics are a useful monitor of performance across a network of laboratories capable of highlighting measurands that require attention (and specifically sigma metrics <5 warrants investigation and attention from Sysmex technical support). Keeping sigma metrics consistently above 5 ensures an efficient and effective quality management system. We intend to use sigma metrics <5 as the threshold for action to be taken to improve the performance for hematology in our laboratories. We intended to further evaluate the use of observed bias <3% and the failure to achieve the BV-based desirables—bias and the 0.33 x CV_I goal as indicators of unacceptable performance.

We saw continuous improvements in sigma metrics and most quality specifications during the three performance evaluation phases and have demonstrated that they are a good exercise in risk assessment and quality improvement. Further investigation is needed to determine whether applying more stringent QC rules based on the observed performance could be successful in improving performance.

ACKNOWLEDGMENTS

Two of the authors are employed by SYNLAB-VPG. All expenses involved in this study were covered by SYNLAB-VPG. The authors thank the Managing Director, Dr Andrew Torrance, Director of Operations, Emma Cake and Director of VPG Cork, Lucy Gaffney for their support in this project and the following colleagues involved in the QA/QC group in the SYNLAB-VPG laboratories for their help and support: Amy Browne BSc (Hons), Claire Crompton BVSc (Hons) MRCVS, Dr Clare Doyle BVSc, FRCPath, MRCVS, Matt Garland BSc RSciTech SAC Dip Cert Nat Sci VN MRSB, Louise Jarvis BSc (Hons), Karen Jones, Sabrina Laatz, Chris Pickard, Sarah Putwain MA, VetMB DipECVCP, FRCPath, PhD MRCVS, RCVS, Dan Stentiford BSc (Hons), and Julie Vickers BVSc, DipACVP(Clin Path), FRCPath, MRCVS.

ORCID

Susan Daly https://orcid.org/0000-0003-2627-9735
Kathleen P. Freeman https://orcid.org/0000-0003-1796-0158

REFERENCES

1. Jassam N, Lindsay C, Harrison K, Thompson D, Bosomworth MP, Barth JH. The implementation of a system for managing analytical quality in networked laboratories. *Ann Clin Biochem*. 2011;48(2):136-146.
2. Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L, Morgan T. Evaluating laboratory performance on quality indicators with the six sigma scale. *Arch Pathol Lab Med*. 2000;124(4):516-519.
3. Greg Miller W, Myers GL, Lou Gantzer M, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem*. 2011;57(8):1108-1117.
4. Harrison HH, Jones JB. Using sigma quality control to verify and monitor performance in a multi-instrument, multisite integrated health care network. *Clin Lab Med*. 2017;37(1):207-241.
5. Farr AJ, Freeman KP. Quality control validation, application of sigma metrics, and performance comparison between two biochemistry analyzers in a commercial veterinary laboratory. *J Vet Diagn Invest*. 2008;20(5):536-544.
6. Sacchini F, Freeman KP. Quality documentation challenges for veterinary clinical pathology laboratories. *J Vet Diagn Invest*. 2008;20(3):266-273.
7. Nabity MB, Harr KE, Camus MS, Flatland B, Vap L. ASVCP guidelines: allowable total error hematology. *Vet Clin Pathol*. 2018;47(1):9-21.
8. Fraser CG. *Biological Variation: From Principles to Practice*. Washington, DC: AACC Press:29-66.

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

9. Westgard JO. Method validation: The decision on method performance. *Basic Method Validation*. 3rd ed. Madison, WI: Westgard QC Inc; 2008:188-196.
10. SEED Haematology - The Red Blood Indices. https://www.sysmex-europe.com/academy/library/educational-articles-seed/seed-the-red-blood-cell-indices-1131.html. Accessed June 2020.
11. Jensen AL, Kjelgaard-Hansen M. Method comparison in the clinical laboratory. *Vet Clin Pathol*. 2006;35(3):276-286.
12. Westgard JO. Selecting the right SQC procedure. *Basic QC Practices*. 4th ed. Madison, WI: Westgard QC Inc; 2016:140-151.
13. Rishniw M, Pion PD. Evaluation of performance of veterinary in-clinic hematology analyzers. *Vet Clin Pathol*. 2016;45(4):604-614.
14. Flatland B, Camus MS, Baral RM. Analytical quality goals-a review. *Vet Clin Pathol*. 2018;47(4):527-538.
15. Westgard S. Prioritizing risk analysis quality control plans based on sigma-metrics. *Clin Lab Med*. 2013;33:41-53.
16. Westgard JO, Barry PL. *Basic Method Validation*. 3rd ed. Madison, WI: Westgard QC; 2008.
17. Westgard JO. Interpreting SQC results using "Westgard Rules". *Basic QC Practices*. 4th ed. Madison, WI: Westgard QC Inc; 2016:45-58.
18. Parry DM Quality goal index. https://www.westgard.com/guest34.htm. Accessed June 2020.
19. Litten J. Applying sigma metrics to reduce outliers. *Clin Lab Med*. 2017;37(1):177-186.
20. Petersen PH, Fraser CG, Jørgensen L, et al. Combination of analytical quality specifications based on biological within- and between-subject variation. *Ann Clin Biochem*. 2002;39:543-550.
21. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem*. 1997;34(Pt 1):8-12.

How to cite this article: Daly S, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000*i*V analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021;50:568–578. doi:10.1111/vcp.13016

1939165x, 2021, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13016 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX B: PUBLICATION OF CHAPTER 3

Title: Quality control validation for a veterinary laboratory network of six Sysmex XT-2000*i*V haematology analysers

Reference: This chapter has been published as Daly SC, Freeman KP, Graham PA. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000*i*V hematology analyzers. *Vet Clin Pathol*. 2022; 51: 565-576. doi: 10.1111/vcp.13163

Received: 7 September 2021 | Revised: 15 April 2022 | Accepted: 10 May 2022

DOI: 10.1111/vcp.13163

ORIGINAL ARTICLE

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

Quality control validation for a veterinary laboratory network of six Sysmex XT-2000*i*V hematology analyzers

Susan Daly[1,2] | Peter A. Graham[2] | Kathleen P. Freeman[3]

[1]SYNLAB-VPG/Cork, South Cork Industrial Estate, Cork, Ireland

[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

[3]SYNLAB-VPG/Exeter, Unit 1b, Science Park, Exeter, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, South Cork Industrial Estate, Vicar's Road, Pouladuff, Cork T12 DNW4, Ireland.
Email: susan.daly@synlab.co.uk

Abstract

Background: Quality control (QC) validation is an important step in the laboratory harmonization process. This includes the application of statistical QC requirements, procedures, and control rules to identify and maintain ongoing stable analytical performance. This provides confidence in the production of patient results that are suitable for clinical interpretation across a network of veterinary laboratories.

Objectives: To determine that a higher probability of error detection (P_{ed}) and lower probability of false rejection (P_{fr}) using a simple control rule and one level of quality control material (QCM) could be achieved using observed analytical performance than by using the manufacturer's acceptable ranges for QCM on the Sysmex XT-2000*i*V hematology analyzers for veterinary use. We also determined whether Westgard Sigma Rules could be sufficient to monitor and maintain a sufficiently high level of analytical performance to support harmonization.

Methods: EZRules3 was used to investigate candidate QC rules and determine the P_{ed} and P_{fr} of manufacturer's acceptable limits and also analyzer-specific observed analytical performance for each of the six Sysmex analyzers within our laboratory system using the American Society of Veterinary Clinical Pathology (ASVCP)-recommended or internal expert opinion quality goals (expressed as total allowable error, TE_a) as the quality requirement. The internal expert quality goals were generated by consensus of the Quality, Education, Planning, and Implementation (QEPI) group comprised of five clinical pathologists and seven laboratory technicians and managers. Sigma metrics, which are a useful monitoring tool and can be used in conjunction with Westgard Sigma Rules, were also calculated.

Results: The QC validation using the manufacturer's acceptable limits for analyzer 1 showed only 3/10 measurands reached acceptable P_{ed} for veterinary laboratories (>0.85). For QC validation based on observed analyzer performance, the P_{ed} was >0.94 using a 1-2.5s QC rule for the majority of observations (57/60) across the group of analyzers at the recommended TE_a. We found little variation in P_{fr} between manufacturer acceptable limits and individual analyzer observed performance as this is a characteristic of the rule used, not the analyzer performance.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Veterinary Clinical Pathology* published by Wiley Periodicals LLC on behalf of American Society for Veterinary Clinical Pathology.

Conclusions: An improved probability of error detection and probability of false rejection using a 1-2.5s QC rule for individual analyzer QC was achieved compared with the use of the manufacturers' acceptable limits for hematology in veterinary laboratories. A validated QC rule (1-2.5s) in conjunction with sigma metrics (>5.5), desirable bias, and desirable CV based on biologic variation was successful to evaluate stable analytical performance supporting continued harmonization across the network of analyzers.

KEYWORDS
harmonization, hematology, quality control, sigma metrics, validation, Westgard rules

1 | INTRODUCTION

Quality goals have been established for clinical laboratory procedures based on biological variation data or expert opinion,[1–5] and analytical equipment that achieves these quality goals will generate results that are suitable for clinical decision-making.[6] Once analyzers have been evaluated and optimized to meet quality goals, statistical quality control (QC) aims to maintain performance within those goals. Statistical QC relies on one or more control rule applied to the results generated from regular analyses of quality control materials (QCM). For statistical QC to be effective in achieving its purpose, the selected rule(s) must be sensitive and specific for the identification of deteriorating analytical performance. That is, statistical control validation is a necessary step in the design and implementation of statistical QC rule(s).

In a network of veterinary laboratory hematology analyzers that have undergone a harmonization process,[7] statistical QC performs an additional role in maintaining harmonization and allowing for continued interchangeability of results and common reference intervals within the veterinary network. In the same way that QC validation supports the effectiveness of QC to ensure the stability of a single analyzer, QC validation is a necessary step in the design and implementation of QC rules for a network of harmonized analyzers.

The sensitivity and specificity of QC are reflected by the probability of error detection (P_{ed}) and probability of false rejection (P_{fr}), respectively. The P_{ed} is a measure of the frequency with which a control rule would cause analytical runs to be rejected when results contain errors beyond the inherent imprecision. Ideally, error detection should be set to 100% for medically significant errors; however, error detection at ≥90% is considered sufficient.[6] Conversely, P_{fr} is a measure of the frequency with which analytical runs are rejected when there is no apparent reason or issue. The goal is to have the highest possible P_{ed} to ensure that medically important errors are not missed and a low P_{fr} (≤5%)[6] to reduce the waste cost and efficiency impact on patient sample volume, reagent use, and result turnaround time. This optimizes the efficiency and capability of statistical QC as a tool for the demonstration of stable laboratory system performance. The analytical performance capabilities of the analyzer inform the choice of the number of control materials and statistical rules selected, which then determines the P_{ed} and P_{fr} achieved. This step-by-step process results in a quality control validation.

Laboratories using statistical QC often employ the Westgard rules, which are denoted by a shorthand notation. For example, a 1-2s rule means that control limits encompass two standard deviations (SD) on each side of the observed mean. A 1-3s refers to a rule that is set at ±3 SDs, and the rule is violated when one measurement exceeds ±3 SDs from the QCM mean. The numeral 2 placed in front of a rule, such as 2-2s, indicates the rule is violated when two consecutive control measurements or two single measurements across two QCMs are outside the 2SD control limits.[6]

Six Sigma process-improvement methodology has been applied to clinical laboratory analyses such that performance capability (bias and imprecision) relative to TE_a can be represented on a numeric scale as a "sigma metric" (σ). Sigma metrics and QC rules have been combined in Westgards Sigma Rules,[6] such that sigma metrics can, in turn, determine whether a simple single QC rule or a more complicated collection of rules is required. A sigma metric ≥6 for an analytical method indicates <3.4 errors/defects per million results,[8] defined as world-class performance, and the implied low bias and imprecision mean that these methods are easily controlled with simple QC rules and a low number of QCM measurements. It would take a large deterioration in performance to cause a clinical error, and such a large shift could be readily detected with a simple QC rule. In fact, for measurands with sigma metrics ≥6.0, acceptable P_{ed} and P_{fr} can be achieved using 1 or 2 control measurements. Conversely, measurands with lower sigma metrics require a multirule and/or larger number of control measurements and may or may not be able to achieve acceptable P_{ed} and P_{fr}.[9] Commercially available software called EZRules3 (Westgard QC, https://www.westgard.com/store/software.html) allows the user to explore and select QC rules based on the total error quality goal and observed analytical performance as well as the P_{ed} and P_{fr} that can be achieved.[8]

It has been acknowledged that a P_{ed} of ≥0.85 can be used[10] for point of care testing (POCT) in veterinary medicine or for hematologic analyses, where the use of a single QCM is preferred. The use of a single hematology control material is traditional in the UK but a P_{ed} ≥0.90 cannot be achieved with a 1-3s rule and a single QCM data point can only achieve a P_{ed} ≥0.85.[11,12]

This paper describes a network of six harmonized Sysmex XT-2000iV hematology analyzers across five locations,[7] which required a QC approach that could ensure the maintenance of individual

Appendix B

analyzer performance and supported continued harmonization. In deciding which QC approaches to evaluate, the following preferences and constraints had to be taken into consideration. First, the use of a simple control rule (e.g., 1-2.5s, 1-3s) rather than a multirule and a single level of QCM was preferred due to the simplicity of training and evaluation and based on traditional laboratory practices and economics. Second, the choice of QCM was limited because a third-party control material was not available for the Sysmex analyzer, only a manufacturer-provided QCM. It was felt that the potential disadvantages of a single level of QCM could be mitigated since all blood smears undergo additional nonstatistical quality control via microscopy by a fully trained technician to validate the automated results prior to a pathologist's review of laboratory data and correlation with clinical findings before results are reported.

The authors set out to discover whether an effective QC approach supporting network harmonization could be successfully implemented, given these constraints and preferences. This study addressed the following objectives:

1. Determine that a higher P_{ed} can be achieved using QC rules validated for each individual analyzer than that achieved using the manufacturer's acceptable limits for the control limits.
2. Determine that a P_{ed} (≥0.85) and P_{fr} (≤0.05) can be achieved using a single QCM ($n = 1$) with a simple single control rule (1-2.5s or 1-3s).
3. Assess the use of sigma metric evaluation as part of the QC approach in conjunction with the validated QC rules appropriate for each analyzer, with a goal of using it to monitor when instrument servicing is needed to maintain a high level of instrument performance.
4. Assess whether analyzer performance criteria established in a previous study (bias <3%, achievement of desirable biologic variation-based goals for CV and bias, and 0.33CV_I goals, with sigma metrics >5) would be confirmed in this study as useful contributions to an overall QC approach.

2 | MATERIAL AND METHODS

2.1 | Hematology analyzers

The data from six Sysmex XT-2000*i*V analyzers (Sysmex Corporation, Kobe, Japan) located in five laboratory locations in the UK ($n = 4$) and Ireland ($n = 1$) were evaluated. Analyzers were designated analyzers 1, 2, 3, 4a, 4b, and 5. This network of analyzers had previously undergone optimization and successful harmonization.[7] Analyzer 1 was designated the reference analyzer.

2.2 | Quality control material

The QCM used for these evaluations was Level 2—Normal e-CHECK (XE)-Hematology Control (Sysmex Corporation) provided in an 8-vial kit with one new vial used per week. Typically the lot number changed for each new kit with slight changes in the manufacturer's target means and acceptable limits. A single lot number was used for all analyzers during the study. The hematologic measurands evaluated were the white blood cell count (WBC), red blood cell count (RBC), hemoglobin concentration (HGB), hematocrit (HCT), mean cell volume (MCV), mean cell hemoglobin concentration (MCHC), reticulocyte count (RETIC), platelet count (PLT), plateletcrit (PCT), and red cell distribution width-coefficient of variation (RDW-CV). The QCM was analyzed by a fully trained technician according to standard operating procedures; when not in use, the QCM vial was refrigerated.

2.3 | Evaluation of manufacturer's acceptable limits as control rules using EZRules3

The manufacturer's acceptable limits for the hematologic measurands were evaluated using data from 1 month of QCM results from analyzer 1. The width of the manufacturer's acceptable limits for a measurand was divided by the standard deviation (SD) for the reference lab (analyzer 1) to determine the number of SDs contained within this range. That number was divided by two to determine the SDs on each side of the target mean and was then rounded to the nearest QC rule available within EZRules3 to evaluate the performance of the control rule most closely representative of the manufacturer's acceptable limit. EZRules3 allows the manual selection of simple control rules for 1-2s, 1-2.5s, 1-3s, 1-3.5s, 1-4s, 1-5s, and 1-6s. If the manufacturer's range control rule was equivalent to <±2SDs, a 1-2s rule was used. If the manufacturer's range control rule was >6 SDs, a 1-6s rule was used.

The control rule was then assessed using EZRules3 to determine the P_{ed} and P_{fr} that are possible using the startup QC design, which is a program that allows the user to follow a series of prompts, including the manufacturer's target mean as the decision-level concentration, ASVCP recommendations,[3] and/or internal expert opinion for total allowable error (TE_a) as the chosen quality requirement, and the number of QCMs set to $n = 1$. For those measurands, where no ASVCP recommendation for TE_a was available (PCT and RETIC number), expert opinion goals from an internal working group composed of pathologists and technicians were used (5 and 7 persons, respectively). The expected instability setting in the software was set to off.

2.4 | QC validation—evaluation of control rules based on observed individual analyzer performance

The QC validation for each of the six Sysmex instruments, customized for observed performance, was determined using approximately 1 month of QC data (March 2020), as reported previously.[7] We used EZRules3 to plot the observed imprecision and observed bias (calculated from the manufacturer's target mean) generated from daily repeated QC measurements for each analyzer and measurand to determine individual analyzer QC rules and associated P_{ed}, P_{fr}, and sigma metrics.

The candidate rules evaluated using the manual selection option were 1-2.5s, $n = 1$ and 1-3s, $n = 1$. A P_{ed} ≥0.90 can theoretically be achieved with a 1-2.5s rule and a single QCM if performance is sufficiently good.

2.5 | Review of criteria for measurands with poor performance

Findings from a previous study[7] indicated that Sigma metrics >5, bias <3%, and measurands achieving desirable biologic variation goals for CV and bias demonstrated confident stable analytical performance. These criteria were reviewed and compared with the findings of this study, which confirmed that harmonization was maintained during the course of the study.

3 | RESULTS

3.1 | QC validation for manufacturer's acceptable limits

The results of the QC validation for the manufacturer's acceptable limits for the QCM using the observed SD for analyzer 1 are summarized in Table 1. Acceptable criteria for P_{ed} (>0.85) were met for only three measurands (HGB, PCT, and WBC). Acceptable P_{fr} (≤0.05) was achieved for all 10 measurands using the available control rules that most closely represented the manufacturer's acceptable limits.

3.2 | QC validation based on observed analyzer performance for six Sysmex analyzers

Tables 2-4 summarize the results of the QC validation for two control rules (1-2.5s and 1-3s) based on observed analyzer performance (March 2020) for the selected measurands on six Sysmex analyzers. Only three measurands did not achieve a P_{ed} ≥0.85; RBC and PLT on analyzer 4a using a 1-3s rule ($n = 1$), and PCT using both QC rules, 1-2.5s and 1-3s ($n = 1$). All other measurands achieved a P_{ed} ≥0.85 with either control rule across the six analyzers.

P_{fr} ≤0.05 was achieved for all 10 measurands for each of the six Sysmex analyzers and both candidate control rules.

Sigma metrics were >6 for 56/60 observations. When the sigma metrics was <5.5, three measurands did not achieve a P_{ed} >0.85 (RBC, PLT, and PCT, as above for analyzer 4a); RETIC ($\sigma = 5.98$) on analyzer 4b did achieve an acceptable P_{ed} (>0.85).

3.3 | Evaluation of QC rules

The 1-2.5s rule offered the highest P_{ed} for all measurands for the group of Sysmex analyzers based on the observed analyzer performance and yielded P_{fr}s of only 1% (Tables 2-4). The manufacturer's QC limits achieved acceptable P_{ed} for only three measurands (Table 1).

3.4 | Sigma metric monitoring

A monthly review of sigma metrics showed three measurands that performed <5.5 sigma failed to achieve quality goals for P_{ed} for one (RBC and PLT) or both QC rules (PCT). On further investigation, the quality goal index (QGI)[13] demonstrated that observed imprecision was implicated for two measurands, while imprecision and bias were implicated for one measurand. The same analyzer (4a) accounted for all three measurands with poor performance, which was compared with the other analyzers, where performance was >5.5 sigma. (see Tables 2-4).

3.5 | Review and optimization of criteria for measurands with poor performance

Previous criteria established for stable analytical performance indicated that measurands with a bias <3% achieved desirable biologic variation-based goals for CV, bias, and 0.33CVi goals, with sigma metrics >5 supported analytical stability and harmonization. In this study, we noted that a sigma metric >5.5 rather than 5 demonstrated analytical stability. P_{ed} was achieved when the observed CV and Bias achieved desirable biologic variation goals, as previously reported.[7]

Imprecision limits to achieve acceptable P_{ed} (>0.85) for RBC (see Table 2, analyzer 4a) using the 1-2.5s rule; if desirable bias based on biologic variation was applied (1.761%), then the observed CV % must be ≤1.45 to achieve a P_{ed} >0.94. Similarly, for PLT on analyzer 4a (Table 3), the Sigma metric was 5.13, and for the 1-2.5s rule, we could determine that if desirable bias was achieved based on biologic variation (5.17%), then the observed CV % must be ≤2.60% (3.72%). Therefore, the 1-3s rule was not satisfactory, the 1-2.5s rule was satisfactory at 0.86 P_{ed}, and the other analyzers achieved a P_{ed} >0.94.

4 | DISCUSSION

To the authors' knowledge, this is the first reported QC validation for a network of Sysmex analyzers used in veterinary hematology and ongoing harmonization within a laboratory system. The quality goal used to determine the P_{ed} and P_{fr} was TE_a provided by the ASVCP[3] or internal expert opinion (RETIC and PCT). QC validation, using the manufacturer's acceptable limits and observed analytical performance of analyzer 1, showed that only 3/10 measurands (HGB, WBC, and PCT) achieved an acceptable P_{ed} with the recommended TE_a or expert opinion in the case of PCT (Table 1). Consequently, the manufacturer's acceptable limits could not be relied on to identify clinically important equipment malfunction for most measurands. The wide manufacturer's limits were not unexpected, as this has been highlighted previously,[14,15] and these limits are generally

TABLE 1 Summary of QC validation for Sysmex quality control Level 2, Manufacturer's acceptable QCM ranges using 1 month of QC data from the network reference laboratory Sysmex XT 2000iV hematology analyzer (analyzer 1)

Measurand (units)	RBC (×10⁶/μL)	HBG (g/dl)	HCT (%)	MCV (fL)	MCHC (g/dL)	RDW-CV (%)	RETIC (×10⁶/μL)	WBC (×103/μL)	PLT (×103/μL)	PCT (%)
Manufacturer's acceptable lower limit	4.05	11.3	32.7	76.7	31.7	13.8	63.8	6.81	188	15
Manufacturer's acceptable upper limit	4.47	11.9	36.1	84.7	35.7	18.6	132.6	7.69	254	27
Width of acceptable range	0.42	0.6	3.4	8.0	4.0	4.8	68.8	0.88	66	12
Observed SD for analyzer 1	0.03	0.15	0.35	0.44	0.32	0.09	3.71	0.16	6.50	0.79
Number of observed SD within the acceptable range	14.00	4.00	9.70	18.18	12.50	53.33	18.54	5.50	10.15	1.27
Number of observed SD on each side of the target mean	7.00	2.00	4.85	9.09	6.25	26.66	9.27	2.75	5.07	0.63
Closest available control rule	1-6s	1-2s	1-5s	1-6s	1-6s	1-6s	1-6s	1-3s	1-5s	1-2s
TE_a% (ASVCP except for PCT and RETIC #)	10	10	10	7	10	10	40	15	20	25
Observed CV % (analyzer 1)	0.66	1.32	1.00	0.54	0.96	0.54	5.81	2.12	2.87	3.65
Observed bias % (analyzer 1)	0.008	0.44	0.66	0.75	1.05	0.19	4.00	1.16	2.33	2.89
P_{ed}	>0.01	>0.98	>0.13	>0.01	>0.01	>0.01	>0.01	>0.85	>0.13	>0.98
P_{fr}	0.0	0.05	0	0	0	0	0	0	0	0.05
Sigma metric	15.14	7.24	9.34	11.57	9.32	18.17	6.20	6.53	6.16	6.06

Notes: Number of observed SDs in the manufacturer's acceptable range was divided by two to determine the number of SD on each side of the manufacturer's target mean. This was used to determine the closest control rule available for selection in the EZRULES3 program for evaluation. If the number of SD within the acceptable range/2 is much less than the closest available control rule, the width of the acceptable range will be reduced compared with the manufacturer's acceptable range, resulting in overestimation of the P_{ed} (PCT). If the SD within the acceptable range/2 is much greater than the closest available control rule, the width of the acceptable range will be greater than the manufacturer's acceptable range, resulting in underestimation of the P_{ed} (RDW-CV).

Abbreviations: ASVCP, the American Society of Veterinary Clinical Pathology; CV, coefficient of variation; HBG, hemoglobin; HCT, hematocrit; MCHC, Mean Corpuscular Hemoglobin Concentration; MCV, Mean Corpuscular Volume; P_{ed}, probability of error detection; P_{fr}, probability of false rejection; PLT, Platelet count; PCT, Plateletcrit QC, quality control; QCM, quality control material; RBC, Red Blood Cell count; RDW-CV, coefficient of variation of the Red Cell Distribution Width; RETIC, reticulocyte count; SD, standard deviation; TE_a, total allowable error; WBC = White Blood Cell count.

TABLE 2 Summary of the hematological erythroid measurand QC validation for six Sysmex XT 2000iV hematology analyzers based on observed performance (March 2020)

Hematology analyzers	Analyzer 1		Analyzer 2		Analyzer 3		Analyzer 4a		Analyzer 4b		Analyzer 5	
RBC ($\times 10^6$/μL)												
TE_a goal %	10		10		10		10		10		10	
Observed mean	4.32		4.37		4.34		4.39		4.39		4.35	
Target mean	4.32		4.32		4.32		4.32		4.32		4.32	
Observed CV %	0.69		0.49		0.68		1.60		0.70		1.00	
Observed absolute bias %	0.10		1.13		0.40		1.60		1.58		0.64	
Observed TE	1.48		2.11		1.76		4.8		2.98		2.64	
Sigma metric	14.35		18.1		14.12		5.25		12.03		9.36	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	0.88	**0.76**	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
HGB (g/dL)												
TE_a goal %	10		10		10		10		10		10	
Observed mean	12.08		11.90		11.98		12.21		12.22		12.11	
Target mean	12		12		12		12		12		12	
Observed CV %	0.59		0.37		0.32		0.47		0.44		0.90	
Observed absolute bias %	0.63		0.80		0.15		1.79		1.87		0.90	
Observed TE	1.81		1.54		0.79		2.73		2.75		2.7	
Sigma metric	15.88		24.86		30.78		17.47		18.48		10.11	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
TE_a goal %	10		10		10		10		10		10	
HCT (%)												
TE_a goal %	10		10		10		10		10		10	
Observed mean	36.1		36.5		35.9		36.3		36.6		36.3	
Target mean	35.7		35.7		35.7		35.7		35.7		35.7	
Observed CV %	0.64		0.56		1.06		0.55		0.47		1.01	
Observed absolute bias %	1.14		2.13		0.43		1.78		2.40		2.05	
Observed TE	2.42		3.25		2.55		2.88		3.34		4.07	

TABLE 2 (Continued)

Hematology analyzers	Analyzer 1		Analyzer 2		Analyzer 3		Analyzer 4a		Analyzer 4b		Analyzer 5	
Sigma metric	13.84		14.11		9.03		14.95		16.17		7.87	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
MCV (fL)												
TE_a goal %	7		7		7		7		7		7	
Observed mean	83.5		83.4		82.7		83.7		83.4		83.8	
Target mean	82.6		82.6		82.6		82.6		82.6		82.6	
Observed CV %	0.43		0.40		0.69		0.55		0.57		0.62	
Observed absolute bias %	1.09		1.04		0.08		0.92		0.92		1.45	
Observed TE	1.95		1.84		1.46		2.02		2.06		2.69	
Sigma metric	13.74		14.90		10.03		11.05		11.72		8.95	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
MCHC (g/dL)												
TE_a goal %	10		10		10		10		10		10	
Observed mean	33.4		32.6		33.4		33.5		33.3		33.3	
Target mean	33.6		33.6		33.6		33.6		33.6		33.6	
Observed CV %	0.70		0.72		1.11		0.92		0.74		0.83	
Observed absolute bias %	0.50		0.54		0.54		0.27		0.68		1.08	
Observed TE	1.9		1.98		2.76		2.11		2.16		2.74	
Sigma metric	13.57		13.14		8.52		10.58		12.59		11.89	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

(Continues)

TABLE 2 (Continued)

Hematology analyzers	Analyzer 1		Analyzer 2		Analyzer 3		Analyzer 4a		Analyzer 4b		Analyzer 5	
RDW-CV (%)												
TE_a goal %	10		10		10		10		10		10	
Observed mean	16.1		16.2		16.1		16.2		16.3		16.3	
Target mean	16.2		16.2		16.2		16.2		16.2		16.2	
Observed CV %	0.48		0.75		0.49		0.69		0.56		0.64	
Observed absolute bias %	0.69		0.09		0.33		0.02		0.59		0.57	
Observed TE	1.65		1.59		1.31		1.4		1.71		1.85	
Sigma metric	19.40		13.21		19.73		14.46		16.80		14.73	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
RETIC ($\times 10^6/\mu L$)												
TE_a goal %	40		40		40		40		40		40	
Observed mean	70.1		73.9		71.8		76.5		77.0		75.9	
Target mean	74.3		74.3		74.3		74.3		74.3		74.3	
Observed CV %	5.11		4.58		4.30		4.87		6.07		4.73	
Observed absolute bias %	4.65		3.37		3.37		2.91		3.72		0.44	
Observed TE	14.87		12.53		11.97		12.65		15.86		9.9	
Sigma metric	6.48		8		8.52		7.62		5.98		8.36	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

Notes: All instruments used the same QC material lot number. The number of QC materials is one ($n = 1$). P_{ed} <0.85 is highlighted in bold type.

Abbreviations: CV, coefficient of variation; HBG, hemoglobin; HCT, hematocrit; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; P_{ed}, probability of error detection; P_{fr}, probability of false rejection; QC, quality control; RBC, red blood cell count; RDW-CV, coefficient of variation of the red cell distribution width; RETIC, reticulocyte count; TE, total error; TE_a, total allowable error.

TABLE 3 Summary of PLT and PCT QC Validation for six Sysmex XT 2000iV hematology analyzers based on observed performance (March 2020)

Hematology analyzers													
	Analyzer 1		Analyzer 2		Analyzer 3		Analyzer 4a		Analyzer 4b		Analyzer 5		
PLT ($\times 10^3/\mu L$)													
TE_a goal %	20		20		20		20		20		20		
Observed mean	220.4		220.0		209.4		217.2		224.1		213.1		
Target mean	220		220		220		220		220		220		
Observed CV %	2.61		2.21		2.23		3.72		2.48		2.21		
Observed absolute bias %	0.16		0.00		4.81		0.93		1.87		3.13		
Observed TE	5.38		4.42		9.27		8.37		6.83		7.55		
Sigma metric	7.6		9.05		6.81		5.13		7.31		7.63		
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3 s	
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	0.86	**0.72**	>0.94	>0.85	>0.94	>0.85	
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	
PCT (%)													
TE_a goal %	25		25		25		25		25		25		
Observed mean	0.21		0.22		0.20		0.21		0.21		0.20		
Target mean	0.21		0.21		0.21		0.21		0.21		0.21		
Observed CV %	3.09		3.27		2.38		5.31		2.98		2.83		
Observed absolute bias %	2.02		0.91		6.16		0.76		1.45		2.20		
Observed TE	8.2		7.45		10.92		11.38		7.41		7.86		
Sigma metric	7.44		7.37		7.92		4.56		7.9		8.06		
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3 s	
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	**0.69**	**0.50**	>0.94	>0.85	>0.94	>0.85	
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	

Notes: All instruments used the same QC material lot number. The number of quality control materials is one ($n = 1$). P_{ed} <0.85 is highlighted by bold type.

Abbreviations: CV, coefficient of variation; P_{ed}, probability of error detection; P_{fr}, probability of false rejection; PCT, plateletcrit; PLT, platelet count; QC, quality control; TE, total error; TE_a, total allowable error.

Appendix B

TABLE 4 Summary of WBC QC validation for six Sysmex XT 2000*i*V hematology analyzers based on observed performance (March 2020)

Hematology analyzers	Analyzer 1		Analyzer 2		Analyzer 3		Analyzer 4a		Analyzer 4b		Analyzer 5	
WBC ($\times 10^3/\mu L$)												
TE_a goal %	15		15		15		15		15		1	
Observed mean	7.27		7.17		7.12		7.10		7.16		7.16	
Target mean	7.11		7.11		7.11		7.11		7.11		7.11	
Observed CV %	1.19		1.82		2.31		1.62		1.78		1.38	
Observed absolute bias %	2.33		0.82		0.18		0.04		0.77		0.70	
Observed TE	4.71		4.46		4.8		3.28		4.33		3.46	
Sigma metric	10.65		7.79		6.42		9.23		7.99		10.36	
Control rules, $n = 1$	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
P_{ed}	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85	>0.94	>0.85
P_{fr}	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

Notes: All instruments used the same QC material lot number. The number of QC materials is one ($n = 1$).

Abbreviations: CV, coefficient of variation; P_{ed}, probability of error detection; P_{fr}, probability of false rejection; QC, quality control; TE, total error; TE_a, Total allowable error; WBC, white blood cell count.

derived from groups of instruments rather than individual instruments. For the manufacturer's limits that failed to achieve an acceptable P_{ed}, error detection was as low as >0.01, using the closest available QC rule[6,8,15] of 1-5s or 1-6s. In comparison, the QC validation used the observed performance of individual analyzers for all 10 measurands, achieving a P_{ed} >0.94 with the 1-2.5s QC rule and a P_{ed} >0.85 with the 1-3s QC rule for the observed performance of analyzer 1, resulting in much higher error detection.

We determined that a higher P_{ed} could be achieved using QC rules validated for each individual analyzer rather than the use of the manufacturer's acceptable limits. We determined that we could achieve an acceptable P_{ed}, >0.94 for 95% of observations and P_{fr} ≤0.05 for all observations for our network of Sysmex analyzers. We failed to meet the P_{ed} criteria (>0.85, $n = 1$) for 3/60 (5%) of observations (analyzer 4a for RBC, PLT, and PCT) but were able to achieve the required P_{ed} >0.85 using either 1-2.5s and/or 1-3s QC rule, $n = 1$, for all other measurands. It is noteworthy that the performance of analyzer 4a improved following instrument service and was able to achieve the quality goals (data not shown). This is compared with the manufacturers' P_{ed}, where only two measurands achieved >0.94 (Table 1). We are satisfied that the QC validation from the observed analytical performance offers the most consistent probability of error detection ≥0.85 for the largest number of measurands and, therefore, greater confidence in quality control and the likelihood of network stability. Using the manufacturer's acceptable limits does not provide sufficient P_{ed} for most measurands to have peace of mind regarding detection of unstable instrument performance. The probability of false rejection was within P_{fr} criteria (≤0.05) for analyzer 1 using both manufacturer's limits and observed analytical performance. Overall, the individual analyzer validation allowed us the ability to achieve an error detection ≥0.85 and low levels of false rejection, resulting in more satisfactory QC using the 1-2.5s rule.

The 1-2.5s rule proved to be the optimal solution for one level of control material as the P_{ed} was >0.94 for the majority of observations, whereas a P_{ed} of 0.88 was the highest value achieved using the 1-3s rule at the recommended TE_a. This was achieved with little variation in P_{fr} between the rules used to assess individual analyzer validation since P_{fr} is considered a function of the rule used and the number of QCM and is not based on analyzer performance.

For the measurand (PCT) that failed to meet acceptable criteria for P_{ed} using both QC rules (1-2.5s and 1-3s), it was considered whether the TE_a based on expert opinion was too stringent; however, as the remaining analyzers performed optimally, declining performance was more likely. Moreover, on closer evaluation, the large CV (5.31%) and sigma metric <6 ($\sigma = 4.56$) suggest that for this measurand, the instrument performance was deteriorating and required attention, which was reflected in both QC rule failures.

Sigma metrics were applied as an additional performance monitoring tool and in conjunction with QC validation. The use of sigma metrics as a performance indicator is well documented.[16,17,18] Sigma metrics were >6 for all measurands for analyzer 1 when based on observed analytical performance. For measurands on other analyzers that performed with <6 sigma (bias and/or imprecision for RBC and PLT on analyzer 4a was distinct from the other measurands across the instruments that performed at >6 sigma).

On closer evaluation for these measurands, we could determine limits for observed imprecision when applying a 1-2.5s rule to achieve acceptable P_{ed}. In reviewing the literature, we could not conclusively draw relevant comparisons due to the limited research in veterinary publications regarding imprecision goals. However, we believe that the following information is of interest to others performing QC validation using hematology analyzers. Biologic variation data were not available for PCT; however, if bias was set to 0%, then an observed CV ≤4.30% would be required to achieve an acceptable P_{ed} and >6 sigma; the observed CV for analyzer 4a was 5.31%. RETIC for analyzer 4b had a sigma metric of 5.98 but could achieve an acceptable P_{ed} and P_{fr} for the 1-2.5s and 1-3s rules; however, unlike the other measurands, the observed CVs were not

distinct from the other analyzers. Biologic variation data were not available for this measurand, but from our data, if bias was 0%, an observed CV ≤7.0% is required to meet the quality goal (40%).

These observations demonstrate that, depending on the error budget, a low % CV with a high bias may still achieve the quality goal and vice versa, but they must meet desirable goals based on biologic variation. The use of sigma metrics in conjunction with the QC rules (Westgard Sigma rules) highlighted analytical instability for analyzer 4a compared with the other 5 analyzers and technical attention was required for those measurands performing at <5.5 sigma.

The measurands that were easily controlled were >5.5 sigma. These measurands had results with desirable biases and CVs based on biologic variation and could achieve high P_{ed} and low P_{fr} at the specified TE_a. By meeting these criteria, these measurands were controlled using a simple QC rule. This is more cost-effective[16,18] and less labor-intensive for the technician, offering a degree of confidence to the clinical pathologist interpreting the results in a multisite, multi-analyzer environment. For these measurands, a 1-2.5s rule was considered the best candidate rule for ongoing use. This validation study is a much-needed step in a harmonization process, ensuring that our network of analyzers is comparable. We know that analytical variability exists between instruments of the same model[7,19]; therefore, by implementing validated QC rules and monitoring analytical performance using Westgards Sigma Rules,[20] we are controlling this aspect of variability. The benefits gained from harmonization using maximally efficient high P_{ed} and low P_{fr} control approaches include a reduction of network costs, uniformity of standard operating procedures, unified quality management policy, unified training and proficiency, less rerun waste, and increased uniformity of turnaround times across laboratories.

Measurands more difficult to control were those <5.5 sigma, which had lower error detection (<0.85 P_{ed}) and higher observed CV (>3.5%) compared with better performing measurands. In most cases, these measurands failed to meet the desirable CV based on biologic variation (Tables 2-4). In review of the criteria for identifying a need for analytical and technical attention, previous findings[7] suggest that a bias >3%, failure to meet desirable biologic variation goals, and sigma metrics <5 were immediate triggers for poor performance requiring servicing. However, this study suggested some modifications to those criteria because some measurements showed suboptimal performance (PLT, RBC, and PCT for analyzer 4A) when sigma metrics were <5.5 rather than <5. For those example measurands with biological variation data available (RBC, HGB, HCT, WBC, and PLT), those that failed to meet desirable biologic variation goals with a sigma metric <5.5 warrant further investigation as they were poorly controlled with a 1-2.5s and/or 1-3s rule.

A more complicated multirule could be adopted in these cases, but in the interest of keeping QC simple and not too laborious for the technicians and network quality Managers, our recommended approach is to monitor the sigma metric, observed bias %, and observed CV %. We are confident that using (1) the 1-2.5s QC rule, (2) ensuring that the desirable biologic variation-based quality requirements for bias and CV are met, and (3) maintaining sigma metrics >5.5 sigma are excellent indicators of analytical stability sufficient to maintain harmonization. In addition, we continue to use nonstatistical measures such as a microscopic blood film evaluation by a hematologist, the consideration of clinical findings by a clinical pathologist, patient history, and a comparison/correlation of previous results. These measures are particularly important and recommended for measurands where statistical QC is not performed or does not provide high P_{ed}. When statistical QC may not be sufficient to provide necessary quality monitoring, the addition of nonstatistical QC is recommended to help ensure that accurate and reliable results are reported.[6]

Some academic work has been published on QC validation of hematology analyzers, but this has predominantly been in human medicine.[21,22] In veterinary medicine, biochemistry analyzer QC validation has encompassed more of the focus.[14,15] Early work by Freeman and Gruenwaldt,[23] as well as some comparative work,[24-26] has created a good foundation for our harmonization study and emphasized that QC validation should be a requirement for veterinary laboratories, which is further supported in this study.

5 | CONCLUSIONS

We recommend QC validation of individual hematology analyzer performance to achieve high P_{ed} and low P_{fr} rather than the use of the manufacturer's acceptable QC limits for hematology in veterinary laboratories. We validated 57/60 observations using the 1-2.5s and/or 1-3s QC rule; however, for optimal P_{ed}, we applied the 1-2.5s QC rule. This simple rule used with a single QCM is adequate for individual analyzers and for a harmonized network of analyzers if the sigma metrics remain >5.5 and the measurands meet desirable biologic variation goals for CV and bias. These standards should be applied across groups of analyzers and should aid the internal quality TE_am, pathologists, and technicians in recognizing when analytical issues arise.

ACKNOWLEDGMENTS

Two of the authors are employed by SYNLAB-VPG. All expenses involved in this study were covered by SYNLAB-VPG. The authors would like to thank the SYNLAB Managing Director, Dr Andrew Torrance, Director of Operations, Emma Cake, and Director of VPG Cork, Lucy Gaffney, for their support of this project and the following colleagues involved in the QEPI group in the SYNLAB-VPG laboratories for their help and support: Amy Browne BSc (Hons), Claire Crompton BVSc (Hons) MRCVS, Dr Clare Doyle BVSc, FRCPath, MRCVS, Matt Garland BSc RSciTech SAC Dip Cert Nat Sci VN MRSB, Louise Jarvis BSc (Hons), Karen Jones, Sabrina Laatz (BSc), Chris Pickard, Sarah Putwain MA, VetMB DipECVCP, FRCPath, PhD MRCVS, RCVS, Dan Stentiford BSc (Hons), and Julie Vickers BVSc, DipACVP(Clin Path), FRCPath, MRCVS.

DISCLOSURE

The authors have indicated that they have no affiliations or financial involvement with any organization or entity with a financial interest in, or in financial competition with, the subject matter or materials discussed in this article.

ORCID

Susan Daly https://orcid.org/0000-0003-2627-9735
Kathleen P. Freeman https://orcid.org/0000-0003-1796-0158

REFERENCES

1. Fraser CG. Biological variation in clinical chemistry. An update: collated data, 1988-1991. *Arch Pathol Lab Med*. 1992;116(9):916-923.
2. Fraser CG. *Biological Variation: From Principles to Practice*. Amer. Assoc. for Clinical Chemistry; 1992.
3. Nabity M, Harr K, Camus M, Flatland B, Vap L. ASVCP guidelines: allowable total error hematology. *Vet Clin Pathol*. 2018;47:9-21.
4. Harr K, Flatland B, Nabity M, Freeman K. ASVCP guidelines: allowable total error. *Vet Clin Pathol*. 2013;42:424-436.
5. Flatland B, Camus MS, Baral RM. Analytical quality goals-a review. *Vet Clin Pathol*. 2018;47(4):527-538.
6. Westgard JO, Basic QC. *Practices: Training in Statistical Quality Control for Medical Laboratories*. Westgard QC, Inc; 2016.
7. Daly S, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000iV analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021;50:568-578.
8. Westgard JO, Westgard SA. Quality control review: implementing a scientifically based quality control system. *Ann Clin Biochem*. 2016;53(Pt 1):32-50.
9. Westgard JO, Barry PL. *Basic Method Validation*. 3rd ed. Westgard QC; 2008.
10. Rishniw M, Pion PD. Evaluation of performance of veterinary in-clinic hematology analyzers. *Vet Clin Pathol*. 2016;45(4):604-614.
11. Westgard JO. *Six Sigma Quality Design and Control*. Westgard QC, Inc; 2001.
12. Lester S, Harr KE, Rishniw M, Pion P. Current quality assurance concepts and considerations for quality control of in-clinic biochemistry testing. *J Am Vet Med Assoc*. 2013;242(2):182-192.
13. Parry DM. Quality goal index. https://www.westgard.com/guest34.htm. Published 2009. Accessed June 2020.
14. Farr AJ, Freeman KP. Quality control validation, application of sigma metrics, and performance comparison between two biochemistry analyzers in a commercial veterinary laboratory. *J Vet Diagn Invest*. 2008;20(5):536-544.
15. Manzocchi S, Furman E, Freeman K. Comparison of 3 options for choosing control limits in biochemistry testing. *Vet Clin Pathol*. 2017;46(1):120-125.
16. Harrison HH, Jones JB. Using sigma quality control to verify and monitor performance in a multi-instrument, multisite integrated health care network. *Clin Lab Med*. 2017;37(1):207-241.
17. Westgard JO, Westgard SA. Assessing quality on the sigma scale from proficiency testing and external quality assessment surveys. *Clin Chem Lab Med*. 2015;53(10):1531-1535.
18. Litten J. Applying sigma metrics to reduce outliers. *Clin Lab Med*. 2017;37(1):177-186.
19. Development SEEa. SEED Haematology - the red blood indices. https://www.sysmex-europe.com/academy/library/educational-articles-seed/seed-the-red-blood-cell-indices-1131.html. Published 2012.
20. Westgard J, Basic QC. *Practices: Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC, Inc; 2016.
21. Jassam N, Lindsay C, Harrison K, Thompson D, Bosomworth MP, Barth JH. The implementation of a system for managing analytical quality in networked laboratories. *Ann Clin Biochem*. 2011;48(2):136-146.
22. Westgard JO. Useful measures and models for analytical quality management in medical laboratories. *Clin Chem Lab Med*. 2016;54(2):223-233.
23. Freeman K, Gruenwaldt J. Quality control validation in veterinary laboratories. *Vet Clin Pathol*. 1999;28:150-155.
24. Flatland B, Friedrichs KR, Klenner S. Differentiating between analytical and diagnostic performance evaluation with a focus on the method comparison study and identification of bias. *Vet Clin Pathol*. 2014;43(4):475-486.
25. Cook AM, Moritz A, Freeman KP, Bauer N. Quality requirements for veterinary hematology analyzers in small animals-a survey about veterinary experts' requirements and objective evaluation of analyzer performance based on a meta-analysis of method validation studies: bench top hematology analyzer. *Vet Clin Pathol*. 2016;45(3):466-476.
26. Freeman KP, Flatland B. Studies of instrument/method comparison, as part of validation for use in veterinary species. *Vet Clin Pathol*. 2017;46(1):5-8.

How to cite this article: Daly S, Graham PA, Freeman KP. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000iV hematology analyzers. *Vet Clin Pathol*. 2022;51:565-576. doi: 10.1111/vcp.13163

APPENDIX C: PUBLICATION OF CHAPTER 4

Title: Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of 4 Sysmex XT2000*i*V haematology analysers

Reference: This chapter has been published as Daly SC, Graham PA, Freeman KP. Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000*i*V hematology analyzers. *Vet Clin Pathol*. 2023; 00: 1-9. doi:10.1111/vcp.13211

Check for updates

Received: 21 May 2022 | Revised: 8 September 2022 | Accepted: 22 October 2022

DOI: 10.1111/vcp.13211

ORIGINAL ARTICLE

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000*i*V hematology analyzers

Susan Daly[1,2] | Peter A. Graham[2] | Kathleen P. Freeman[3]

[1]SYNLAB-VPG/Cork, South Cork Industrial Estate, Cork, Ireland

[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

[3]SYNLAB-VPG/Exeter, Exeter, UK

Correspondence
Susan Daly, SYNLAB-VPG/Cork, South Cork Industrial Estate, Vicar's Road, Pouladuff, Cork T12 DNW4, Ireland.
Email: susan.daly@thevpg.ie

Abstract

Background: Repeat patient testing-quality control (RPT-QC) uses retained patient samples as an alternative to commercial quality control material (QCM). We elected to calculate and validate RPT-QC limits for red blood cell count (RBC), hemoglobin (HBG), hematocrit (HCT), and white blood cell count (WBC).

Objectives: (1) To validate RPT-QC across a network of four harmonized Sysmex XT-2000*i*V hematology analyzers and determine the total error that can be controlled with RPT-QC. (2) To generate quality control (QC) limits using the standard deviation (SD) of the duplicate measurement differences and determine a suitable simple QC rule with a probability of error detection >0.85 and probability of false rejection <0.05. (3) Monitor RPT-QC using sigma metrics as a performance indicator and (4) to challenge RPT-QC to ensure acceptable sensitivity.

Methods: Fresh adult canine EDTA samples with results within reference intervals were selected and run again on days 2, 3, and 4. QC limits were generated from the SD of the duplicate measurement differences. The QC limits were challenged using interventions designed to promote unstable system performance. The total error detectable by RPT-QC was determined using EZRULES 3 software.

Results: In all, 20-40 data points were needed for RPT-QC calculations and validated using 20 additional data points. The calculated limits differed among the network of analyzers. The total error that could be controlled was the same or better than that of the manufacturer's commercially available quality control material using the same analyzer for all measurands except hematocrit, which required a higher total error goal than that proposed by ASVCP guidelines to achieve an acceptable probability of error detection. The challenges designed to mimic unstable system performance were successfully detected as out-of-control QC.

Conclusions: The challenges for RPT-QC resulted in acceptable detection of potential unstable system performance. This initial study demonstrates that RPT-QC limits differ among the network of Sysmex XT-2000*i*V analyzers, indicating a requirement

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Veterinary Clinical Pathology* published by Wiley Periodicals LLC on behalf of American Society for Veterinary Clinical Pathology.

to customize for the individual analyzer and laboratory conditions. RPT-QC could achieve ASVCP total allowable error goals for RBC, HGB, and WBC, but not for HCT. Sigma metrics were consistently >5.5 for RBC, HGB, and WBC, but not for HCT.

KEYWORDS
commercial quality control, harmonization, hematology, repeat patient testing-quality control, Sysmex

1 | INTRODUCTION

Quality control procedures for hematology analyzers are necessary to ensure that the quality of operations is maintained and that reliable results are produced.[1] The traditional method of ensuring quality controlled hematological data uses one or more commercial quality control material (QCM). Commercial QCM used in veterinary medicine is typically derived from human blood and synthetic/artificial components; therefore, may not comprise a commutable matrix for veterinary specimens. A preferable matrix would be from the animal origin, which includes components comparable to the patient throughput.[2] There is not an available supply of veterinary-specific commercial QCM for hematology; therefore, veterinary laboratories have little choice but to use a QCM with a potentially suboptimal matrix. Commercial QCM is a costly but necessary expense and has a relatively short shelf life.[3] Most importantly, it does not provide appropriate assurances as to how the veterinary specimens may behave on the analyzer and, therefore, is substandard in its application to a veterinary laboratory.

Repeat patient testing-quality control (RPT-QC) offers an alternative to commercial QCM by using retained patient samples to determine analytical stability. It relies on the principle that under specific storage conditions and time intervals, a labile EDTA specimen will deteriorate in a predictable way from the baseline measurement to the later repeated measurement of the same specimen. A repeated measurement exceeding thresholds of expected deterioration suggests instability within the analytical system. The concept of RPT-QC has been well documented in human medicine; however, its application has varied.[4–7] Recommendations for an RPT-QC protocol have been proposed,[8] in which a pilot set of data is gathered over a minimum of 20 days to generate control limits. These limits can then be used to determine whether RPT-QC results are "in-control" or "out-of-control." After this initial limit-derivation phase, a fresh RPT sample would be measured on Day 1 and again within a defined time interval and defined storage conditions. It has been recommended to use patient samples yielding within reference interval results as they represent a significant proportion of caseload and may avoid extreme variabilities between measurements.[9] The difference between original and repeated analyses can then be calculated and plotted similarly to other QC procedures.[7] There have been some positive reports in the literature[10–12] that indicate a good potential for application in veterinary laboratories.

This study addresses the following hypotheses.

1. That RPT-QC limits could be generated and validated across a network of four Sysmex XT-2000*i*V hematology analyzers.
2. That RPT-QC performance can achieve the ASVCP-recommended total allowable error (TE_a) goals using a simple QC rule, that is, 1-2.5 s or 1-3 s, a probability of error detection (P_{ed}) of ≥0.85, and a probability of false rejection (P_{fr}) of ≤0.05.
3. That Sigma metrics is a satisfactory monitoring tool for RPT-QC, and sigma metrics >6 (reflecting world class performance) can be achieved.
4. That deliberately created QC "challenges" mimic unstable performance that is detectable by RPT-QC.

2 | MATERIALS AND METHODS

2.1 | Analyzers and measurands

RPT-QC data from four Sysmex XT-2000*i*V analyzers (Sysmex Corporation, Kobe, Japan) were evaluated for the following measurands: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), and white blood cell count (WBC). These analyzers were previously optimized and harmonized[13] to ensure stable performance prior to this study. Analyzers were designated analyzer 1 (located in the first author's laboratory), 3, 4a, and 4b. The laboratory with analyzer 2 was unable to participate and is not included in this report. A single level of commercial QCM (level 2 - Normal e-CHECK (XE)-Hematology Control), which is composed of stabilized human RBCs, human WBCs, platelets, and nucleated RBC components in a preservative medium, was analyzed once per day at analyzer start-up, prior to analysis of patient samples, by a fully trained technician according to a standard operating procedure and was within accepted limits before assessing patient samples.[14] These measurands were chosen as those most useful in our network of laboratories for statistical QC.[14] Differential cell counts and platelet counts were considered to be more adequately assessed by nonstatistical QC by technician evaluations of blood smears.[14]

2.2 | RPT-QC limits generation

The process of RPT-QC data collection and calculations were conducted as recommended by Westgard.[9] Canine EDTA samples that were received by the laboratory for routine testing were selected by

1939165x, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix C

a trained hematology technician for RPT-QC based on time of sampling (fresh/same-day submissions), surplus volume, and results that were within or close to reference intervals used by the laboratory. The day of accession was denoted "Day 1." The selected sample was refrigerated at 4-8°C and reanalyzed on days 2, 3, and 4 following acclimation to room temperature, thorough mixing, and rechecking for clots at each time point prior to analysis. The Day 1 minus Day 2 (or Day 1 – Day 2) change was used to define control limits that could be used for QC on successive days during the working week (Monday through Friday or Monday through Saturday). The Day 1 minus Day 3 (or Day 1 – Day 3) change was used to define control limits that could be used for QC on Mondays for those laboratories that worked Monday-Saturday. The Day 1 minus Day 4 (or Day 1 – Day 4) control limits were for QC on Mondays for those laboratories that worked Monday-Friday. This process was carried out across the network of four Sysmex analyzers, and the data were exported to Microsoft Excel (Microsoft Excel for Mac 2011, Version 14.7.7 [170905] Last update 14.7.7). Each analyzer accumulated 50-60 RPT-QC samples comprising a dataset of up to 40 samples for Day 1 – Day 2, Day 1 – Day 3, and Day 1 – Day 4 intervals to generate control limits.

A scatter plot of the differences of the repeated measurements was produced to identify outliers by visual examination. A maximum of three outliers were excluded from each control limit dataset following a visual review of a scatter plot of repeated measurements except for one dataset, Day 1 – Day 4 HGB, from which four were identified and removed. More complicated statistical evaluation for outliers was not undertaken to keep the procedure simple. If outliers were not readily apparent by visual examination, all data points were retained and included in the calculation of the RPT-QC limits. The mean of the differences for the repeated measurements for each time interval and the standard deviation of the duplicate (SD_{dup}) measurements were calculated for each time interval (Day 1 – Day 2, Day 1 – Day 3, Day 1 – Day 4). The coefficient of variation (CV) and the difference squared (Diff Sq) were calculated from the mean of the repeated measurements, as described previously.[9–11] See Figure 1 for example of initial calculations. The control limit range was calculated using the SD_{dup} and 1-2.5 s and 1-3 s control rules.

The following formula was used to calculate the SD_{dup}:

$$SDdup = \sqrt{\frac{\sum \text{difference}^2}{2n}}$$

where n is the number of observations.[15]

The control rule chosen for use was based on evaluation of the range of QC data compared to the width of the control limit interval (Figure 2) and whether or not validation was achieved by inclusion of at least 17 or more additional data points within the chosen control limits. If the QC data were not all included within the calculated limits or the data did not appear to adequately fill the calculated control limit interval, then an additional 10-20 QC data points were combined with the first dataset and control limits were recalculated using the 30-40 total QC data points, as described earlier (Figure 3).

	D1	D2	Aver D1,D2	Diff	Diff sq
1	15.8	15.8	15.8	0	0
2	15.3	15.3	15.3	0	0
3	11.7	11.7	11.7	0	0
4	15.4	15.6	15.5	-0.2	0.04
5	16.3	16.4	16.35	-0.1	0.01
6	13.8	13.7	13.75	0.1	0.01
7	14	14	14	0	0
8	14.9	15	14.95	-0.1	0.01
9	11.6	11.7	11.65	-0.1	0.01
10	15.2	15.2	15.2	0	0
11	18.4	18.4	18.4	0	0
12	17.6	17.7	17.65	-0.1	0.01
13	15.3	15.4	15.35	-0.1	0.01
14	14.3	14	14.15	0.3	0.09
15	15.3	15.5	15.4	-0.2	0.04
16	11.7	11.8	11.75	-0.1	0.01
17	16.3	16.6	16.45	-0.3	0.09
18	11.4	11.4	11.4	0	0
19	15.4	15.5	15.45	-0.1	0.01
20	14.9	14.9	14.9	0	0
averages	14.73	14.78	14.755	-0.05	0.017

FIGURE 1 An example of data calculations using 20 data points for hemoglobin Day 1 minus Day 2 (Day 1 – Day 2), showing Day 1 (D1) measurements, Day 2 (D2) measurements, the difference of Day 1 – Day 2 (Diff), an average of Day 1 and Day 2 measurements (Aver D1, D2), and the difference of the measurements squared (Diff sq).

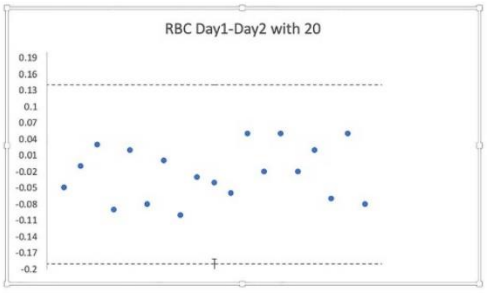


FIGURE 2 Initial scatter plot for RBC counts using 19 data points (one outlier removed) from the difference of Day 1 minus Day 2 (Day 1 – Day 2) samples used to calculate control limits for the 1-3 s rule. The dashed lines represent the upper and lower control limits (0.14 and −0.19, width 0.33). The graph shows the data does not adequately fill the control limits, which results in a poor sensitivity of detection of error.

2.3 | QC validation for calculated RPT-QC control limit interval

QC validation for the calculated RPT-QC control limits for each measurand, time period, and laboratory analyzer were undertaken using the following specifications (ASVCP recommendations for TE_a for hematology for allowable total error, number of QCM = n = 1, $P_{ed} \geq 0.85$, $P_{fr} \leq 0.05$, and simple rules used [1-2.5 s or 1-3 s]). The choice of $P_{ed} \geq 0.85$ was made because this is the minimum P_{ed} that can be achieved with a simple 1-3 s rule, and a single-level control result based on the models provided in the EZRules3 program. EZRules3 was then used to determine P_{ed} and P_{fr} that could be achieved using the recommended TE_a (ASVCP) for each of the

19391658, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

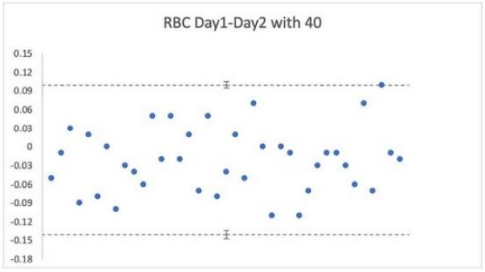


FIGURE 3 Subsequent scatter plot for RBCs using 39 data points (one outlier removed) from Day 1 minus Day 2 (Day 1 − Day 2) samples used to calculate control limits for the 1-3s rule. The dashed lines represent the upper and lower control limits (0.1 and −0.14, width 0.24). The graph shows the data adequately fills the control limits, resulting in greater sensitivity for detecting out-of-control events with smaller width control limits.

measurands. EZRules3 was also used to determine the lowest TE_a that could be achieved using the control rules chosen based on a simple QC rule and CV of the mean difference for the measurand while achieving a $P_{ed} > 0.85$ and $P_{fr} < 0.05$.

2.4 | RPT-QC and sigma metrics

The sigma metric for each measurand and RPT-QC limit were calculated according to the following formula, using the ASVCP recommendations for allowable total error, the CV from each RPT-QC interval, and absolute bias = 0%:

$$\text{Sigma} = (\text{TE}_a\% - \text{absolute bias}\%) / \text{CV}\% = \text{TE}_a\% / \text{CV}\%$$

The SD, CV, mean difference, and control limit ranges for 1-3s, and 1-2.5s rules were calculated for each measurand and dataset. The CV was calculated using the mean of the averages (see Figure 1) of the Day 1 − Day 2, Day 1 − Day 3, or Day 1 − Day 4 results as follows:

$$\text{CV} = (\text{SD}_{dup} / \text{mean of averages of the Day 1 and repeated results}) \times 100$$

2.5 | RPT challenges

RPT-QC challenges were performed at a single laboratory (analyzer 1) over 1 week and were run as a single evaluation as per procedure for analyzing RPT-QC. Challenges were set up to create intentional sample alterations to test the control limits for RPT-QC. The sample alterations included using an aged sample past days Day 1-Day 4, freezing a sample, and a low-volume sample ("short collection" not coming up to the fill line) with excess EDTA, resulting in an elevated EDTA:blood ratio sample. Samples were also manipulated by spiking the sample with saline creating a hemodiluted sample. Hemoconcentrated samples were created by removing plasma. The addition of native canine lipids obtained from a lipemic serum sample to an EDTA blood sample created a lipemic sample.

3 | RESULTS

3.1 | RPT-QC limit generation

Tables 1–4 summarize the information pertinent for the generation of the RPT-QC limits for each measurand, time period, and analyzer. Upper and lower control limits were generated from the four Sysmex analyzers for WBC, RBC, HGB, and HCT using either 20, 30, or 40 QC data for Day 1 − Day 2, Day 1 − Day 3, and Day 1 − Day 4. Most datasets had at least one outlier; only four datasets were identified with no outliers. Forty data points were required for the generation of RPT-QC limits for most measurands. Only 10 of the 48 datasets across the measurands validated the initial calculations with 20 data points for either the 1-2.5s or 1-3s QC rules. Calculated RPT-QC limits that were obtained from calculations performed with 40 data points were subsequently validated with a further 20 data points.

Figure 2 illustrates the use of a scatter plot from the differences obtained for Day 1 − Day 2 for RBC using 19 data points (analyzer 1). One outlier was removed. Figure 3 illustrates a scatter plot for 39 data points as one outlier was removed for RBC Day 1 − Day 2 with RPT-QC limits indicated by dashed lines. The scatter plots show that the 40 data points more completely fill the range specified for the RPT-QC limits.

3.2 | QC Validation for RPT-QC intervals

Tables 1–4 summarize the QC validation and total allowable error that could be achieved for each measurand. ASVCP recommended TE_a quality goals of 10% total error were achievable across all measurands and time periods across all analyzers for RBC and HGB using the specified criteria of a simple rule, single quality control assessment, and P_{ed} and P_{fr} described earlier. Analyzer 4b could not achieve the ASVCP recommended quality goal for WBC for Day 1 − Day 2 and Day 1 − Day 3, though the other three analyzers could. HCT could not achieve ASVCP recommended TE_a across all analyzers and time periods.

3.3 | RPT-QC and sigma metrics

Sigma metrics were >6 for each analyzer for all measurands except for HCT, which was <3 sigma across all analyzers and measurands using the ASVCP recommended total allowable error quality goal of ±10%.

3.4 | RPT-QC challenges

The intentional sample alterations using saline-diluted and frozen samples were detected as out-of-control by the RPT-QC for all

19391653c, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [12/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Repeat patient testing-quality control. The standard deviation (SD), coefficient of variation (CV), mean difference, control limits range, ASVCP-recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics, and QC rule obtained for white blood cell counts (WBC).

Analyzer	1			3			4a			4b		
Day (time interval)	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4
SD	0.159	0.169	0.227	0.21	0.175	0.234	0.15	0.15	0.19	0.278	0.238	0.22
CV	1.29	1.51	1.85	2.04	1.76	2.47	1.69	1.75	2.12	2.88	2.73	2.35
Mean difference	−0.111	−0.0395	−0.556	−0.06	0.098	0.022	0.009	−0.071	−0.018	0.155	−0.214	0.094
Control limits range (+/−)	0.796	0.843	1.42	1.04	1.1	1.17	0.906	0.928	1.123	1.65	1.42	1.32
ASVCP TE_a (%)	15	15	15	15	15	15	15	15	15	15	15	15
Lowest achievable TE_a (%)	8	10	12	11	10	14	11	11	13	18	17	15
Sigma metric based on ASVCP recommended TE_a	11.63	9.93	8.11	7.35	8.52	6.07	8.88	8.57	7.11	6.25	6.23	6.47
QC rule	1-2.5s	1-2.5s	1-3s	1-2.5s	1-3s	1-3s	1-3s	1-3s	1-3s	1-3s	1-2.5s	1-3s

Abbreviations: ASVCP, American Society of Veterinary Clinical Pathology; CV, coefficient of variation; D1-D2, Day 1 – Day 2; D1-D3, Day 1 – Day 3; D1-D4, Day 1 – Day 4; quality control, QC; SD, standard deviation; TE_a, total allowable error.

TABLE 2 Repeat patient testing-quality control. The standard deviation (SD), coefficient of variation (CV), mean difference, control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics, and QC rule obtained for red blood cell counts (RBC).

Analyzer	1			3			4a			4b		
Day (time interval)	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4
SD	0.039	0.042	0.062	0.061	0.104	0.084	0.033	0.049	0.039	0.082	0.076	0.065
CV	0.56	0.60	0.90	0.86	1.47	1.18	0.50	0.75	0.58	1.21	1.11	0.96
Mean difference	0.021	−0.033	−0.015	0.005	0.007	0.076	0.008	−0.013	−0.025	−0.056	−0.034	−0.017
Control limits range	0.240	0.21	0.37	0.36	0.52	0.5	0.199	0.297	0.233	0.412	0.454	0.325
ASVCP TE_a	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TE_a	4	4	6	5	8	7	3	5	4	8	7	6
Sigma metric based on ASVCP recommended TE_a	17.86	16.78	11.17	11.64	6.8	8.43	20.12	13.33	17.33	8.26	9.01	10.42
QC rule	1-3s	1-2.5s	1-3s	1-3s	1-2.5s	1-3s	1-3s	1-3s	1-3s	1-2.5s	1-3s	1-2.5s

Abbreviations: ASVCP, American Society of Veterinary Clinical Pathology; CV, coefficient of variation; D1-D2, Day 1 – Day 2; D1-D3, Day 1 – Day 3; D1-D4, Day 1 – Day 4; quality control, QC; TE_a, SD, standard deviation; total allowable error.

15391606, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13219 by Hardin Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Repeat patient testing-quality control. The standard deviation (SD); coefficient of variation (CV), mean difference, control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics, and QC rule obtained for hemoglobin concentrations (HGB).

Analyzer	1			3			4a			4b		
Day (time interval)	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4
SD	0.085	0.49	0.083	0.081	0.061	0.63	0.134	0.152	0.168	0.117	0.127	0.164
CV	0.5	0.48	0.5	0.566	0.734	0.386	0.847	0.965	1.07	0.715	0.127	0.164
Mean difference	−0.039	−0.02	−0.003	−0.034	−0.029	0.042	−0.113	−0.108	−0.154	−0.089	−0.144	−0.127
Control limits range	0.5	0.48	0.5	0.45	0.36	0.36	0.804	0.76	0.84	0.701	0.635	0.82
ASVCP TE_a	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TEa	4	4	4	3	5	3	6	6	7	5	5	7
Sigma metric based on ASVCP recommended TEa	19.88	20.83	20	17.67	13.61	7.77	11.81	10.36	9.34	13.99	12.61	10
QC rule	1-3s	1-3s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-2.5s	1-3s	1-2.5s	1-2.5s

Abbreviations: ASVCP, American Society of Veterinary Clinical Pathology; CV, coefficient of variation; D1-D2, Day 1 – Day 2; D1-D3, Day 1 – Day 3; D1-D4, Day 1 – Day 4; quality control, QC; SD, standard deviation; TE_a, total allowable error.

TABLE 4 Repeat patient testing-quality control. The standard deviation (SD), coefficient of variation (CV), mean difference, control limits range, ASVCP recommended total allowable error (TE_a) goal, lowest achievable total allowable error (TE_a), sigma metrics, and QC rule obtained for hematocrit (HCT).

Analyzer	1			3			4a			4b		
Day (time interval)	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4	D1-D2	D1-D3	D1-D4
SD	1.67	2.67	4.14	2.25	3.23	4.86	1.96	3.23	4.23	2.21	3	4.6
CV	3.47	5.41	8.24	4.47	6.21	8.34	3.94	6.37	8.23	4.34	5.84	8.78
Mean difference	−2.25	−3.98	−5.71	−3.15	−4.76	−6	−2.66	−4.42	−5.81	−2.89	−4.12	−6.22
Control limits range	8.37	13.33	20.7	11.28	16.12	15.81	9.89	16.15	21.15	11.04	14.98	22.99
ASVCP TE_a	10	10	10	10	10	10	10	10	10	10	10	10
Lowest achievable TE_a	22	34	50	25	35	45	25	39	50	27	37	54
Sigma metric based on ASVCP recommended TE_a	2.88	1.85	1.21	2.24	1.61	1.2	2.54	1.57	1.22	2.3	1.71	1.14
QC rule	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s	1-2.5s

Abbreviations: ASVCP, American Society of Veterinary Clinical Pathology; CV, coefficient of variation; D1-D2, Day 1 – Day 2; D1-D3, Day 1 – Day 3; D1-D4, Day 1 – Day 4; quality control, QC; SD, standard deviation; TE_a, total allowable error.

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [22/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

measurands. Spiking samples with native canine lipids, excessive EDTA, and incorrect storage were successful in causing errors in some but not all measurands. WBC was out-of-control with added lipid. RBC counts and HGB concentrations were out-of-control in the specimens with excessive EDTA. RBC, HGB, and HCT were out-of-control when the specimen was incorrectly stored (extended storage beyond 4 days).

4 | DISCUSSION

We found that RPT-QC limits could be generated using a single protocol across our network of laboratories using four Sysmex hematology analyzers. Additional information regarding the stability of specimens under defined storage and handling conditions is also possible from assessment of these data but beyond the scope of this study and article. The collection of RPT data across four Sysmex analyzers was the most challenging process of this study. The provision of a spreadsheet to technicians for entering data was essential for managing the RPT samples and data daily. Initial assessments across analyzers and time periods for RPT-QC demonstrated that more than 20 data points are often required to produce RPT-QC limits that will validate with a subsequent dataset of 20 results. This differed from prior recommendations to collect 20 data points[9] and prior studies that were able to validate limits determined using 20 data points.[10–12] Following the RPT-QC limit generation, the validation process required calculation of limits using 30-40 data points for more than half of measurands/time interval evaluations. We found that by only using 20 results, the data were either not included within the control limits or did not adequately fill the range, but when an additional 10-20 data points were used, then we saw a better distribution across the QC data range. Except for HCT, successful validation with 20 data points was more likely for Day 1 – Day 2 limits than Day 1 – Day 3 and Day 1 – Day 4. The calculations using 30-40 data points were more often required for Day 1 – Day 3 and Day 1 – Day 4 RPT-QC time intervals, likely due to more variability in the data due to increased aging. We found the use of scatter plots highly valuable to visualize the distribution of the data and for outlier identification. Intuitively you would think the SD of the duplicates would provide RPT-QC limits that were not too wide or too narrow. However, when assessed visually, some RPT-QC limits were too narrow for the dataset, and others appeared too wide, with the possibility of decreased sensitivity in error detection (Figure 2), which necessitated the use of additional data points to calculate the control limits.

As the sample aged, the trend was for the QC range to become wider, with higher SDs and CVs. This is not unexpected since more variation is expected with increased sample aging. However, it was interesting that the differences with increased aging were not uniformly similar within or across the laboratories. We found this particularly evident for HCT, where we saw the biggest increases in range, SD, and CV as the sample aged. Interestingly, in a previous study, we identified one analyzer performing suboptimally in comparison to the others in the network[13]; this was also reflected in the RPT-QC performance for that analyzer (analyzer 4b), demonstrating reduced analytical sensitivity.

The total error goals represented by ASVCP recommendations for hematology were achievable for the specified conditions ($P_{ed} \geq 0.85$, $P_{fr} \leq 0.05$, simple QC rule [1-2.5 s or 1-3 s], and $n = 1$) for RBC and HBG for all four analyzers and for three of the four analyzers for WBC. The ASVCP recommended TE_a goals could not be achieved for HCT on any of the analyzers with RPT-QC. This was not an unexpected finding as HCT is affected by MCV which has been previously found to increase with sample aging[10,11,14] and, therefore, was not included in this study. The inclusion of HCT was important despite the increased TE associated with differences used for RPT-QC. Although the total error that could be controlled with RPT-QC was unacceptable for interpretation of patient results, knowledge of the expected difference for the various time intervals was considered important for QC and optimizing our QC procedures for RPT-QC.

The sigma metrics using ASVCP TE_a goal for the RPT-QC for RBC and HGB are clearly reflective of world-class performance. Suboptimal sigma metrics (<3 sigma) using the ASVCP quality goal of 10% was seen for HCT for all time intervals and measurands for each analyzer; the recommended TE_a of 10% could not be achieved (see Table 4) and high CVs were seen. Analyzers 3 and 4b had lower sigma metrics compared with the other analyzers and were due to the high CVs generated. The lower sigma metrics for analyzer 4b likely reflect suboptimal performance and the need for servicing to determine if better performance can be obtained from this analyzer. It is interesting that analyzer 4b has consistently been identified with poor performance within the group based on traditional QC performance evaluation and QC validation.[13] This illustrates the fact that some analyzers may not perform as well as others. The fact that analyzers 4a and 4b were within the same laboratory, with the same environmental conditions, maintenance, reagents, control materials, and technicians, indicates that sources of variation other than those inherent to the analyzer itself were unlikely. This analyzer required more frequent servicing (up to 2 times the servicing normally expected) to maintain performance based on our internal criteria for performance monitoring (sigma metrics <5.5). The 5.5 sigma metric goal is considered useful for HCT as a basis for requiring service. For other measurands (RBC, HGB, WBC), a sigma metric <5.5 indicates a need for service.

The RPT-QC challenges were sufficient to demonstrate that specimens falling outside of our selection criteria (fresh, canine EDTA samples filled to the fill line of the collection tube) caused out-of-control events using the QC limits for the reference analyzer, demonstrating that the QC limits were sufficient to detect the induced changes in sample quality.

4.1 | Implications for our network of analyzers using RPT-QC

We can accommodate a 20% TE_a goal cut-off for WBC counts that is within reference intervals for dogs rather than the recommended TE_a

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [12/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of 15% by the ASVCP Quality Assurance and Laboratory Standards Committee,[16] based on the expert opinions of pathologists involved in our Quality Education, Performance, and Implementation Group (QEPI, comprised five clinical pathologists and seven laboratory technicians, managers, and directors with a special interest in quality). The use of 20% as a quality goal for TE_a for hematology does not result in significant changes in the interpretation of results, as indicated by the average difference of WBCs between Day 1 – Day 2, Day 1 – Day 3, and Day 1 – Day 4 of 0.155, −0.214, and 0.094, respectively.

Continued monitoring to determine if all analyzers can consistently achieve 20% total allowable error quality goals for WBC with RPT-QC is needed. The lowest achievable TE_a for HCT ranged from 22% to 27% across the analyzers for Day 1 – Day 2, from 24% to 39% for Day 1 – Day 3, and between 45% and 54% for Day 1 – Day 4 compared with the ASVCP recommendation of 10% (Tables 1–4). The TE_a for HCT that could be achieved was considered acceptable for RPT-QC because RBC and HGB were also assessed and able to achieve their respective ASVCP total error goals. In addition, a further control measure for HCT using a spun PCV (considered the gold standard for evaluation of the erythron in veterinary medicine) is generated whenever HBG×3 differs from the automated HCT by five or more units.

An alternative of Day 1 – Day 4 could be used for measurand measurements at least 6 hours apart to use on Mondays for laboratories operating on a Monday-Friday schedule. However, further research will be required to determine acceptable RPT-QC limits for this option.

Two QC rules (1-2.5 s and 1-3 s) were assessed based on their ability to achieve acceptable P_{ed} (≥85%) and P_{fr} (≤5%) when $n = 1$. The preferred QC rule is 1-3 s, where one violation exceeding ±3 SDs from the mean deems the QC run a failure, this simple approach can be easily adopted by the laboratory and pathology personnel. We found not all measurands for each time interval could be validated using the 1-3 s rule, in fact only 45% could be validated. We have tailored RPT-QC monitoring spreadsheets to include both 1-3 s and 1-2.5 s QC rules specific to each analyzer; however, educating/training the user on the differences in QC rule violations is imperative for successful implementation.

Sigma metrics were employed as a quality performance indicator, where sigma >6 is indicative of world-class performance.[8] Our previous findings[13] showed that a sigma metric <5.5 using commercially available QCM indicated instability within the analytical system with a need for further investigations and troubleshooting procedures. Additional studies are needed to determine an acceptable sigma metric cut-off for monitoring the performance of RPT-QC. For the measurands at time intervals with high CVs and unacceptable TE_a compared with the reference analyzer, we found this was reflected in the sigma metric. For example, Day 1 – Day 2 and Day 1 – Day 3 WBC intervals for analyzer 4b had a sigma metric of 6.25 and 6.23, respectively, compared with the reference analyzer, where sigma metrics were 11.63 and 9.93, respectively. The poor performance of HCT was also reflected in the sigma metrics for all analyzers in the network, demonstrating the utility of sigma metrics as a performance indicator for RPT-QC.

5 | CONCLUSIONS

We found that RPT-QC limits could be generated for a network of Sysmex XT-2000iV hematology analyzers, but that the use of 40 data points was frequently required to achieve the subsequent validation of the RPT-QC by a dataset of 20 points. The TE_a goals provided by the ASVCP for hematology can be achieved for RBC and HBG. Alteration in acceptable TE_a for WBC measurand from 15% to 20% was considered acceptable based on internal expert opinion, while larger TE_a goals for HCT (ranging from 22% to 54%, depending on the time interval) were considered acceptable based on the acceptable TE_a performance for RBC and HBG. Sigma metrics provide a unitless measurement of performance capacity and, as demonstrated in previous studies using commercially available QCM, provide a reliable indication of suboptimal performance. Continued evaluation of RPT-QC in parallel with commercial QCM is needed to determine if an RPT-QC sigma metric cut-off of <5.5 for the various measurands indicates the need for analyzer service to maintain and improve analyzer performance. The challenges produced for RPT-QC to mimic unstable system performance demonstrated that RPT-QC was sensitive to these alterations. Continued monitoring in parallel with commercial QCM is required to determine if RPT-QC is reliable for daily internal quality control for hematology.

ACKNOWLEDGMENTS

Two of the authors are employed by SYNLAB-VPG. All expenses involved in this study were covered by SYNLAB-VPG. The authors thank the SYNLAB Managing Director, Dr Fiona Gosling, Director of Operations, Emma Cake, and Director of VPG Cork, Lucy Gaffney, for their support in this project and the following colleagues involved in the QEPI group in the SYNLAB-VPG laboratories for their help and support: Amy Browne BSc (Hons), Claire Crompton BVSc (Hons) MRCVS, Dr Clare Doyle BVSc, FRCPath, MRCVS, Matt Garland BSc RSciTech SAC Dip Cert Nat Sci VN MRSB, Louise Jarvis BSc (Hons), Karen Jones, Chris Pickard, Sarah Putwain MA, VetMB DipECVCP, FRCPath, PhD MRCVS, Dan Stentiford BSc (Hons), and Julie Vickers BVSc, DipACVP (Clin Path), FRCPath, MRCVS.

CONFLICT OF INTEREST STATEMENT

The authors confirm that there is no conflict of interest.

ORCID

Susan Daly https://orcid.org/0000-0003-2627-9735
Kathleen P. Freeman https://orcid.org/0000-0003-1796-0158

REFERENCES

1. Flatland B, Weiser G. The elephant in the room (and how to lead it out): In-clinic laboratory quality challenges. *J Am Anim Hosp Assoc*. 2014;50:50-382.

19391676, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2. Unsal I, Coskun A, Serteser M, Inal TC, Ozpinar A. Toward standardization of quality assessment in laboratory medicine by using the same matrix samples for both internal and external quality assessments. *Accreditation Qual Assur*. 2010;15(11):621-627.
3. Flatland B, Vap LM. Quality management recommendations for automated and manual in-house hematology of domestic animals. *Vet Clin North Am Small Anim Pract*. 2012;42(1):11-22.
4. Dutra FR. Monitoring the quality of blood cell counts with replicate determinations on routine samples. *Am J Clin Pathol*. 1966;46(2_ts):286-288.
5. Carstairs KC, Peters E, Kuzin EJ. Development and description of the "random duplicates" method of quality control for a hematology laboratory. *Am J Clin Pathol*. 1977;67(4):379-385.
6. Orser B. Hematology quality control and the Coulter S-Plus II. *Am J Med Technol*. 1983;49(9):643-647.
7. Cembrowski GS, Lunetzky ES, Patrick CC, Wilson MK. An optimized quality control procedure for hematology analyzers with the use of retained patient specimens. *Am J Clin Pathol*. 1988;89(2):203-210.
8. Westgard J, Basic QC. *Practices: Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC; 2016.
9. Westgard J. Implementing repeat patient test controls. *Basic QC Practices: Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC; 2016.
10. Vanyo LC, Freeman KP, Meléndez-Lazo A, Teles M, Cuenca R, Pastor J. Comparison of traditional statistical quality control using commercially available control materials and two patient-based quality control procedures for the ADVIA 120 hematology system. *Vet Clin Pathol*. 2018;47(3):368-376.
11. Flatland B, Freeman KP. Repeat patient testing shows promise as a quality control method for veterinary hematology testing. *Vet Clin Pathol*. 2018;47(2):252-266.
12. Flatland B, Freeman KP. Repeat patient testing-based quality control shows promise for use in veterinary biochemistry testing. *Vet Clin Pathol*. 2020;49(4):590-606.
13. Daly S, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000iV analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021;50:568-578.
14. Daly S, Graham PA, Freeman K. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000iV hematology analyzers. *Vet Clin Pathol*. 2022;51(4):565-576.
15. Westgard JO. Developing a QA program to monitor the QC plan. In: Westgard JO, ed. *Basic QC Practices Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC; 2016:245-251.
16. Nabity M, Harr K, Camus M, Flatland B, Vap L. ASVCP guidelines: allowable total error hematology. *Vet Clin Pathol*. 2018;47:9-21.

How to cite this article: Daly S, Graham PA, Freeman KP. Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of four Sysmex XT-2000iV hematology analyzers. *Vet Clin Pathol*. 2024;53(Suppl. 1):39-47. doi:10.1111/vcp.13211

19391676, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13211 by Health Research Board, Wiley Online Library on [22/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX D: PUBLICATION OF CHAPTER 5

Title: Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT- 2000*i*V hematology analyzer in a multi- site veterinary laboratory

Reference: This chapter has been published as Daly S, Rishniw M, Graham PA, Freeman KP. Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT-2000*i*V hematology analyzer in a multi-site veterinary laboratory. *Vet Clin Pathol*. 2024; 53(Suppl. 1): 48-59. doi:10.1111/vcp.13332

Check for updates

Received: 29 March 2023 | Revised: 28 December 2023 | Accepted: 12 January 2024

DOI: 10.1111/vcp.13332

ORIGINAL ARTICLE

Veterinary Clinical Pathology
An International Journal of Laboratory Medicine
WILEY

Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT-2000*i*V hematology analyzer in a multi-site veterinary laboratory

S. Daly[1,2] | M. Rishniw[3] | P. A. Graham[2] | K. P. Freeman[3]

[1]Veterinary Pathology Group, Brushville House, Dosco Business Park, Cork, Ireland

[2]School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK

[3]Veterinary Information Network, Davis, California, USA

Correspondence
S. Daly, Veterinary Pathology Group, Brushville House, Dosco Business Park, South Douglas Road, Cork T12 XR6V, Ireland.
Email: susan.daly@thevpg.ie

Abstract

Background: Quality control material (QCM) for hematology in veterinary laboratories is limited, and repeat patient testing quality control (RPT-QC) is an alternative method using excess matrix-specific samples.

Objectives: This study aimed to determine if median differences between RPT-QC analyses for each time interval for RBC, HGB, HCT, and WBC were the same, determine if unified RPT-QC limits can be applied to a network of veterinary laboratories, compare the performance of RPT-QC to commercial QCM for the reference analyzer and evaluate the experience over a 4 month period and design, improve and implement an automated spreadsheet for RPT-QC data management.

Methods: The potential to unify individual analyzer RPT-QC limits for red blood cells (RBC), hematocrit (HCT), hemoglobin (HGB), and white blood cells (WBC) on multi-site Sysmex XT-2000-*i*V analyzers was explored by a difference of means test and confidence interval determination for the median difference for each network analyzer in comparison to the network reference analyzer. User experience of an automated RPT-QC data management Excel spreadsheet was collected by user feedback during monthly meetings. Numbers of out-of-control results and the root causes for these for RPT-QC were compared against those of a commercial QCM over a 4-month period.

Results: Differences between individual analyzer RPT-QC limits were too large to allow for unification of network limits. The automated spreadsheet successfully highlighted out-of-control events for RPT-QC. Trends or shifts were more frequent for commercial QCM based on observed performance and a 1–2.5s QC rule than for RPT-QC. Following routine troubleshooting, RPT-QC out-of-control events were resolved with an alternative RPT-QC sample indicating random error associated with excessive deterioration. Use of an automated spreadsheet for recording RPT-QC, documentation and troubleshooting of out-of-control events, and collating monthly summary calculations were considered an asset in laboratory quality management.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2024 The Authors. *Veterinary Clinical Pathology* published by Wiley Periodicals LLC on behalf of American Society for Veterinary Clinical Pathology.

Appendix D

Conclusions: RPT-QC can be successfully implemented and integrated into a multi-site veterinary laboratory. Individual analyzer RPT-QC limit generation is recommended. The deterioration of commercial QCM caused shifts or trends in QC results, which initiated more repeat analyses and investigations than did RPT-QC.

KEYWORDS
hematology, quality control, repeat patient testing, Sysmex

1 | INTRODUCTION

Quality control material (QCM) for hematology analyzers in veterinary clinical laboratories is limited and relies on commercially available QCM which is often of human origin and therefore is not a representative matrix for veterinary specimens.[1] This is problematic for veterinary laboratories that want to generate results of the highest standard.[2] Previous reports indicate that a commercial QCM using the manufacturer's acceptable ranges ('bottle values') does not reliably demonstrate out-of-control events due to wide manufacturer's acceptable ranges.[3,4] Quality control based on observed individual analyzer performance is optimal[5] for ensuring veterinary laboratories are maintaining standards according to quality benchmarks designed to highlight unstable system performance. A study of this network of veterinary hematology analyzers that were subject to QC validation using a 1–2.5 s QC rule showed a greater sensitivity for impactful error detection than using manufacturer's ranges.[5]

The only commercially available option for veterinary laboratories is to use human-derived QCM, which is costly and has a relatively short shelf life.[1] However, there is an alternative called repeat patient testing quality control (RPT-QC). RPT-QC utilizes retained excess patient samples representative of the laboratory throughput and relies on the principle that under defined time intervals and storage conditions, an EDTA hematology specimen will vary within a defined range from the baseline measurement to the repeated measurement for the same specimen. If the repeated measurement exceeds these expected limits, then instability may be present within the analytical system.[6] RPT-QC has been reported in the human literature for some time[7–10] but was not utilized widely until further research emerged following a protocol proposed by Westgard.[11] In veterinary laboratories, its application has been limited until recently.[12,13]

RPT-QC was previously validated within a network of Sysmex XT-2000*i*V veterinary hematology analyzers and then challenged to demonstrate that out-of-control events reflecting unstable system performance were detected.[14] In a previous study, a network of harmonized hematology analyzers (4 analyzers participated across 3 separate sites) generated upper and lower RPT-QC limits for each analyzer.[14] An RPT-QC system was validated and achieved total error equivalent to, or less than, total allowable error (TE_a) recommended by the American Society for Veterinary Clinical Pathology (ASVCP)[15] for red blood cells (RBC), hemoglobin (HGB), white blood cells (WBC) but not for hematocrit (HCT). However, HCT was still considered useful to include in statistical quality control; despite its poor performance,[14] understanding the expected difference for the time intervals was relevant information for quality control.

This study evaluated the implementation and suitability of RPT-QC in a harmonized network with the following objectives:

1. Determine if median differences between RPT-QC analyses for each time interval for RBC, HGB, HCT, and WBC were the same.
2. Determine whether unified RPT-QC limits could be applied across all network analyzers or if RPT-QC limits need to be developed based on the performance of individual analyzers.
3. Compare the performance of RPT-QC with that of a commercial quality control material for the reference analyzer.
4. Develop, improve, and implement an automated data management spreadsheet to simplify RPT-QC within the laboratory.
5. Evaluate and summarize the experience of implementing RPT-QC within a commercial laboratory network using RPT-QC and commercially available QCM in parallel over 4 months, including technicians' feedback, challenges, and improvements made over the trial period.

2 | MATERIALS AND METHODS

For this study, measurands evaluated using RPT-QC were RBC, HGB, HCT, and WBC, and for each of the following repeat time intervals: Day1–Day2, Day1–Day3, and Day1–Day4. Repeat intervals were selected to provide QC limits applicable to daily (Day1–Day2), weekend (Day1–Day3), and public holidays (Day1–Day4).

An alternative RPT-QC protocol was also generated for longer closure periods (>Day1–Day4 interval) or if a suitable RPT-QC sample is not available. This consisted of running a freshly collected hematology specimen without visual hemolysis or lipemia that was not more than 10% above the upper or below the lower reference limit, ie, as close to or within reference limits, and using this as the Day 1 reading. This sample was analyzed again in 3–4 h, and the result used as the "Day 2" reading. The Day1–Day2 RPT-QC limits for acceptable performance were used.

Technicians were fully trained, having undergone the required training according to the standard operating procedure (SOP) for hematology within the laboratory network, and passed competency evaluations prior to being entrusted with this task. Analyzers were maintained according to the manufacturer's instructions and a SOP. The SOP for RPT-QC was developed in the first author's laboratory and then shared across the analytical network. The analyzers used in the study were Sysmex XT-2000*i*V analyzers (Sysmex Corporation, Kobe, Japan) designated by numerical identifiers (1 [reference analyzer], 3, 4a and 4b) according to network locations. Canine EDTA samples selected by each individual laboratory site

19391655, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

were used for RPT-QC, and a commercially available control material (Level 2—Normal e-CHECK [XE]-Hematology Control, Sysmex Corporation, Kobe, Japan) was used for the daily QCM.

2.1 | Determine if differences between RPT-QC repeat analyses for each of the time intervals differed across the network

The differences between individual analyzer replicate measurements for each time interval (Day1–Day2, Day1–Day3, and Day1–Day4) for RBC, HGB, HCT, and WBC were compared across the network using the Kruskal–Wallis statistical test.

Day1–Day4) do not differ between analyzers. The disparity between individual analyzer replicate differences was explored by a difference of means test and confidence interval determination for the median difference.[16]

For the differences of the means test, individual analyzer replicate differences [the difference between the initial result (Day 1) and subsequent replicates (Day2, Day3, Day4)] were compared to those of the reference analyzer (analyzer 1). The absence of a difference between mean replicate differences was reflected by overlap in the confidence intervals for the mean replicate difference. If no overlap in the confidence intervals for the mean difference were present, these mean differences were considered to be different from one another.

$$\text{CV for RPT} - \text{QC} = (\text{SD of duplicate measurements}/\text{mean of the average of the duplicate measurements}) \times 100.$$

2.2 | Determine if unified RPT-QC limits can be applied across the network of analyzers

RPT-QC limits were calculated in a previous study,[14] and the application of a single set of RPT-QC limits across the network of analyzers requires that the differences in replicate analysis used to determine the SD_{dup} for each interval (Day1–Day2, Day1–Day3, and

2.3 | Comparison of RPT-QC with a commercial quality control material

In the reference laboratory (analyzer 1) and analyzer 3, RPT-QC and a commercially available QCM (Level 2—Normal e-CHECK [XE]-Hematology Control, Sysmex) were assayed at analyzer start-up, prior to analysis of patient samples, using an SOP for 4 months. A

TABLE 1 SD, CV, mean difference and RPT-QC limits for the reference (1) and network analyzers (3, 4a, 4b) for each time interval for RBC, HGB, HCT, WBC.

Analyzer	1			3		
Time interval	Day 1–2	Day 1–3	Day 1–4	Day 1–2	Day 1–3	Day 1–4
RBC						
SD	0.039	0.042	0.062	0.061	0.104	0.084
CV	0.560	0.600	0.900	0.859	1.470	1.180
Mean difference	0.021	−0.033	−0.051	−0.005	0.007	0.003
Control limits (LL, UL)	(−0.096, 0.138)	(−0.138, 0.072)	(−0.237, 0.135)	(−0.188, 0.178)	(−0.253, 0.267)	(−0.248, 0.255)
HGB						
SD	0.085	0.137	0.083	0.081	0.061	0.063
CV	0.500	0.477	0.500	0.566	0.734	0.386
Mean difference	−0.039	−0.013	−0.003	−0.034	−0.029	−0.042
Control limits (LL, UL)	(−0.294, 0.216)	(−0.424, 0.398)	(−0.252, 0.246)	(−0.237, 0.169)	(−0.212, 0.154)	(−0.200, 0.116)
HCT						
SD	1.670	2.670	4.140	2.250	3.230	4.860
CV	3.470	5.410	8.240	4.470	6.210	8.340
Mean difference	−2.250	−3.980	−5.710	−3.150	−4.760	−6.000
Control limits (LL, UL)	(−6.425, 1.925)	(−10.655, 2.695)	(−16.06, 4.640)	(−8.775, 2.475)	(−12.835, 3.315)	(−18.150, 6.150)
WBC						
SD	0.159	0.169	0.227	0.210	0.175	0.234
CV	1.290	1.510	1.850	2.040	1.760	2.470
Mean difference	−0.111	−0.040	−0.556	−0.060	0.098	0.022
Control limits (LL, UL)	(−0.508, 0.287)	(−0.462, 0.383)	(−1.237, 0.125)	(−0.585, 0.465)	(−0.427, 0.623)	(−0.680, 0.724)

Abbreviations: CV, coefficient of variation; HCT, hematocrit; HGB, hemoglobin; RBC, red blood cell; RPT-QC, repeat patient testing quality control; SD, standard deviation; WBC, white blood cell.

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

single RPT-QC specimen was chosen daily as previously described for RPT-QC, labeled with a sticker on the stopper, and analyzed for RPT-QC the next working day.

Commercial QCM and RPT-QC results were evaluated to determine the number of out-of-control (OOC) events, the number of repeated assays, and whether the OOC events were the result of random error (no underlying problem detected) or reflected true instability of the system. This was counted on a daily basis and comments noted on the QC spreadsheets.

As an additional comparative evaluation tool, the sigma metric for each measurand on each analyzer was calculated using the ASVCP recommendations for TE_a,[15] the CV from each RPT-QC interval and absolute bias = 0% according to the following formula.

$$\text{Sigma} = \left(TE_a\% - \text{absolutebias}\%\right) / \text{CV}$$

The CV for each RPT-QC interval was determined by dividing the SD of the duplicate measurements by the mean of the average of the duplicate measurements for each measurand ×100:

Sigma metrics for the commercial QCM were calculated daily, whereas the RPT-QC automated spreadsheet provided monthly sigma metrics for each measurand as the sigma metrics were consistently high and monthly frequency was considered sufficient.

2.4 | Automated spreadsheet development

To increase the efficiency and effectiveness of statistical QC, RPT-QC daily data management spreadsheets were developed using Microsoft® Excel for Microsoft 365 MSO (Version 2208 Build 16.0.15601.20072), 64-bit Macros were coded within Visual Basic using the Developer Tab in Excel in order to generate more automated applications/operations. The spreadsheet included each measurand data entry, hard-coded RPT-QC limits,[14] and automated out-of-control alerts for the technician operating the Sysmex hematology analyzer. A separate section, including a troubleshooting checklist, was developed for recording of out-of-control events and actions taken to bring the system back into control. A monthly Levey-Jennings (L-J) chart was automatedly generated and reviewed each month or generated more frequently, if the technician desired. Monthly refinement of the RPT-QC control limits was conducted using the pooled SD, which is a cumulative SD[11] to allow for ongoing QC limit updates. The automated spreadsheet was developed in conjunction with a veterinary

4a			4b		
Day 1–2	Day 1–3	Day 1–4	Day 1–2	Day 1–3	Day 1–4
0.033	0.049	0.039	0.082	0.076	0.065
0.500	0.750	0.577	1.210	1.110	0.960
0.008	−0.013	−0.025	−0.056	−0.034	−0.017
(−0.091, 0.107)	(−0.160, 0.134)	(−0.142, 0.092)	(−0.261, 0.149)	(−0.262, 0.194)	(−0.180, 0.146)
0.134	0.152	0.168	0.117	0.127	0.164
0.847	0.965	1.070	0.715	0.793	1.000
−0.113	−0.108	−0.154	−0.089	−0.144	−0.127
(−0.515, 0.289)	(−0.488, 0.272)	(−0.574, 0.266)	(−0.440, 0.262)	(−0.462, 0.174)	(−0.537, 0.283)
1.960	3.230	4.230	2.210	3.000	4.600
3.940	6.370	8.230	4.340	5.840	8.780
−2.660	−4.420	−5.810	−2.890	−4.120	−6.220
(−7.560, 2.240)	(−12.495, 3.655)	(−16.385, 4.765)	(−8.415, 2.635)	(−11.62, 3.380)	(−17.72, 5.280)
0.150	0.150	0.190	0.278	0.238	0.220
1.690	1.750	2.120	2.880	2.730	2.350
0.009	−0.071	0.018	0.016	−0.214	0.072
(−0.441, 0.459)	(−0.521, 0.379)	(−0.552, 0.072)	(−0.818, 0.064)	(−0.809, 0.381)	(−0.588, 0.732)

19391676, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

TABLE 2 Example calculations for RBC Day1–Day4, $n = 20$ with outlier (Y) in row 14 in bold type.

N	Day 1	Day 4	Average Day1–Day 4	Difference D1–D4	Difference squared
1	7.120	7.190	7.155	−0.070	0.005
2	7.460	7.570	7.515	−0.110	0.012
3	5.210	5.210	5.210	0.000	0.000
4	6.450	6.450	6.450	0.000	0.000
5	6.980	7.050	7.015	−0.070	0.005
6	5.750	5.920	5.835	−0.170	0.029
7	7.400	7.450	7.425	−0.050	0.003
8	5.760	5.780	5.770	−0.020	0.000
9	6.830	6.790	6.810	0.040	0.002
10	7.770	7.700	7.735	0.070	0.005
11	7.840	7.880	7.860	−0.040	0.002
12	5.100	5.210	5.155	−0.110	0.012
13	8.060	8.130	8.095	−0.070	0.005
14	**8.320**	**8.530**	**8.425**	**−0.210**	**0.044**
15	6.420	6.530	6.475	−0.110	0.012
16	6.760	6.680	6.720	0.080	0.006
17	7.230	7.380	7.305	−0.150	0.023
18	7.550	7.710	7.630	−0.160	0.026
19	8.230	8.210	8.220	0.020	0.000
20	5.650	5.560	5.605	0.090	0.008
Averages	6.895	6.947	6.921	−0.052	0.010
Without Y	6.819	6.863	6.841	−0.044	0.008
SD_{dup}	0.070	Without Y	0.052		
CV %	1.017	Without Y	0.753		
Sum Diff sq	0.198	Without Y	0.154		

Abbreviations: SD_{dup}, standard deviation of the duplicate measurements; Sum Diff sq, sum of the difference squared.

specialist in spreadsheet programming (second author), and first used in the first author's laboratory and then shared across the network.

2.5 | Technician experience and feedback

Technicians were consulted frequently during the design and trial phase through a series of regular meetings and provided ongoing feedback regarding improvements in the spreadsheet at the first author's site. The spreadsheets were checked for any inconsistencies by senior technicians.

2.6 | Statistical analysis

The Kruskal-Wallis test[17] was performed on RBC, HGB, HCT, and WBC replicates for each time interval (Day1–Day2, Day1–Day3, and Day1–Day4) for all analyzers using IBM SPSS statistical software (version 28). This test determines whether there is a significant difference between independent groups by testing the null hypothesis that the medians of the groups are the same; $P < .05$ indicated a significant difference.

TABLE 3 Kruskal–Wallis test comparing median differences across the analyzer network, significance set to $p < .05$.

Measurand	Time interval	Test statistic	p	df	N
RBC	Day1-Day2	13.019	.005	3	133
	Day1-Day3	4.111	.250	3	131
	Day1-Day4	3.862	.277	3	129
HGB	Day1-Day2	5.851	.119	3	154
	Day1-Day3	22.551	.000	3	130
	Day1-Day4	21.162	.000	3	129
HCT	Day1-Day2	14.117	.003	3	113
	Day1-Day3	9.215	.027	3	129
	Day1-Day4	3.461	.326	3	153
WBC	Day1-Day2	6.938	.074	3	134
	Day1-Day3	13.082	.004	3	112
	Day1-Day4	4.628	.201	3	151

Abbreviations: HCT, hematocrit; HGB, hemoglobin; RBC, red blood cell; WBC, white blood cell.

19391655, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

TABLE 4 Difference of means test showing standard error for the network analyzers (3, 4a, 4b) when compared to the reference analyzer (1) for each time interval and measurand (RBC, HGB, HCT, WBC).

	Time interval		
Analyzer	Day1–Day2	Day1–Day3	Day1–Day4
RBC			
3	189.890	−120.827	−187.930
4a	130.653	−111.229	−134.462
4b	356.568	5.040	−155.905
HGB			
3	−14.153	27.149	98.382
4a	115.438	84.946	159.115
4b	90.810	97.556	135.800
HCT			
3	2.139	1.010	0.266
4a	1.705	0.927	0.110
4b	2.388	0.321	0.513
WBC			
3	−17.658	−87.144	−180.096
4a	−95.223	16.103	−252.724
4b	−48.559	46.494	−238.689

Abbreviations: HCT, hematocrit; HGB, hemoglobin; RBC, red blood cell; WBC, white blood cell.

3 | RESULTS

Table 1 presents the data used for further evaluation and statistics in this study.

Table 2 presents example calculations for RBC Day1–Day4, $n = 20$.

For RBC, there is a difference between the median differences for Day1–Day2 but not for Day1–Day3 and Day1–Day4 across the analyzers. For HGB there was no difference for Day1–Day2, but Day1–Day3, and Day1–Day4 showed a difference between the median differences across the analyzers. There is a difference between the median differences for HCT for Day1–Day2 and Day1–Day3 but no difference for Day1–Day4 across the analyzers.

For WBC, there was no difference between the analyzers for Day1–Day2 and Day1–Day4, but Day1–Day3 showed differences in medians across the analyzers.

Table 3 presents data on Kruskal-Wallis test.

3.1 | Evaluation to determine if unified RPT-QC limits could be used across the network of analyzers

There was no difference between the network analyzers and the reference analyzer for the mean replicate difference for WBC, RBC, HBG, and HCT across each time interval (Day1–Day2, Day1–Day3, Day1–Day4). This suggests that common RPT-QC limits could be

TABLE 5 RBC mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analyzers (3, 4a, 4b) compared to the reference analyzer (1) for time intervals using the same QC rule.

RBC									
Time interval	Day1–Day2			Day1–Day3			Day1–Day4		
Analyzer number	1	3	4a	1	3	4a	1	3	4a
Mean difference	0.021	−0.005	0.008	−0.033	0.007		−0.051	0.003	−0.025
95% CI (LL, UL)	(0.009, 0.033)	(−0.024, 0.014)	(−0.007, 0.023)	(−0.046, −0.02)	(−0.026, 0.040)		(−0.071, −0.031)	**(−0.023, 0.030)**	(−0.044, −0.006)
Control limit, upper	0.138	0.178	0.107	0.072	0.267		0.135	0.255	0.092
95% CI (LL, UL)	(0.126, 0.150)	**(0.159, 0.197)**	**(0.092, 0.122)**	(0.059, 0.085)	**(0.234, 0.300)**		(0.115, 0.155)	**(0.229, 0.282)**	(0.073, 0.111)
Control limit, lower	−0.096	−0.188	−0.091	−0.138	−0.253		−0.237	−0.249	−0.142
95% CI (LL, UL)	(−0.108, −0.084)	**(−0.207, −0.169)**	(−0.106, −0.076)	(−0.151, −0.125)	**(−0.286, −0.220)**		(−0.257, −0.217)	(−0.275, −0.222)	**(−0.161, −0.123)**

Note: Bold type does not overlap with reference analyzer confidence intervals.
Abbreviations: LL, lower limit; RBC, red blood cell; RPT-QC, repeat patient testing quality control; UL, upper limit.

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

applied across the analyzers in the network if the SD of the duplicates were similar across the group of analyzers.

Table 4 presents the difference of means test for each analyzer compared to the reference analyzer.

The control limits across the network had overlap for the majority of the confidence intervals for the mean differences for the network analyzers compared to the reference analyzer for each time interval when the same QC rule was used as the reference analyzer. However, for RBC, HCT, and HGB, there were exceptions where no overlap occurred, indicating that common RPT-QC limits were not suitable for application across all of the analyzers.

Table 5–8 presents confidence intervals for the mean difference, and upper and lower control limits for each measurand.

The performance of RPT-QC and commercial QCM over a 4-month time period for the reference analyzer is compared in Table 9 and includes evaluations for RBC, HGB, HCT, and WBC. There were more OOC events overall with RPT-QC than with the commercial QCM. For events using a 1–2.5 s QC rule, this was within the predicted probability of false rejection (p_{fr}) (.03), but for events using a 1–3 s QC rule, this was greater than the predicted p_{fr} (.00), which was unexpected. However, control charts for RPT-QC generally showed a more stable performance than that of commercial QCM (Figure 1A,B).

Commercial QCM was within control limits (within 1–2.5 s QC rule), and sigma metrics were >5.5 except for one event where the sigma metric was <5.5 but was >5.0 (RBC) for the reference analyzer.

RPT-QC was within control limits during this same time periods, except for 4 occasions for the reference analyzer. These resolved with new patient samples subsequent to troubleshooting, comparable to using a 'new' commercial control for such incidences. Sigma metrics for RPT-QC were >5.5 for all measurands except for HCT (<3 using the ASVCP TE_a quality goal), based on monthly sigma evaluations.

3.2 | Automated spreadsheet development

Clear instructions for the use of the RPT-QC data management spreadsheet were developed (Figure 2A), and the data collection summary tab records the initial calculations for each measurand (Figure 2B). A separate tab for recording OOC data, troubleshooting, resolution, and corrective actions is included (Figure 2C).

Example images of functionality that the authors incorporated into their spreadsheet design are presented in Figure 2A–D. The development phase aimed to produce a document that was easy to use, but contained clear instructions and minimized the possibility of mistakes in data entry.

3.3 | Technician feedback

In the authors' opinions, technician feedback was positive, particularly once the automated data management spreadsheet was finalized and an RPT-QC SOP and training was instituted. Overall, the technicians (5 technicians in total across 3 laboratories) were satisfied with the spreadsheet with a notable appreciation of the separate troubleshooting and recording section, which highlighted any issues to other technicians using the analyzer.

Using RPT-QC resulted in fewer investigations of unusual results or trends when compared to 1–2.5 s rule for the commercial QCM customized for individual analyzer performance even though the commercial QCM was not out of control (OOC) i.e., results outside the specified limits for the measurand. The technicians also highlighted the advantage of having implemented RPT-QC when, on two occasions, the delivery of the commercial QCM was delayed, and on another occasion, the manufacturer's package inserts were not enclosed, precluding the determination of the assayed target value. RPT-QC was the only QC in use until the delivery issues were resolved (1 week).

TABLE 7 HCT mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analyzers (3, 4a, 4b) compared to the reference analyzer (1) for time intervals using the same QC rule.

HCT						
Time interval	**Day1–Day2**				**Day1–Day3**	
Analyzer number	**1**	**3**	**4a**	**4b**	**1**	**3**
Mean difference	−2.250	−3.150	−2.660	−2.890	−3.980	−4.760
95% CI (LL, UL)	(−2.982, −1.518)	(−4.189, −2.111)	(−3.283, −2.037)	(−3.593, −2.187)	(−4.840, −3.120)	(−6.252, −3.268)
Control limit, upper	1.925	2.475	2.240	2.635	2.695	3.315
95% CI (LL, UL)	(1.193, 2.657)	(1.436, 3.514)	(1.617, 2.863)	(1.932, 3.338)	(1.835, 3.555)	(1.823, 4.807)
Control limit, lower	−6.425	−8.775	−7.560	−8.415	−10.655	−12.835
95% CI (LL, UL)	(−7.157, −5.693)	**(−9.814, −7.736)**	(−8.183, −6.937)	**(−9.118, −7.712)**	(−11.515, −9.795)	(−14.327, −11.343)

Note: Bold type does not overlap with reference analyzer confidence intervals.

Abbreviations: HCT, hematocrit; LL, lower limit; RPT-QC, repeat patient testing quality control; UL, upper limit.

19391606, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

One challenge to implementing RPT-QC was the reliance on getting fresh patient samples following longer closure periods; however, the inclusion of the alternative RPT-QC procedure was a useful aide to cope with these extended closure periods of which there were two occurrences, and this procedure was included in the SOP.

4 | DISCUSSION

RPT-QC was previously validated for several Sysmex XT-2000*i*V veterinary hematology analyzers, and the designed challenges indicated that RPT-QC could detect unstable system performance.[14]

Overall, we were satisfied that RPT-QC could detect unstable system performance and that it was a less expensive, user-friendly, and readily available alternative to commercial QCM. The statistical evaluation was crucial in understanding the RPT-QC control limits and demonstrated that the application of unified (the same) QC limits across the network is not possible. The lack of overlap in confidence intervals for the various RPT-QC limits and detection of differences in median differences by the Kruskal-Wallis nonparametric evaluation (null hypothesis rejected) indicated that a common RPT-QC limit could not be applied across all laboratories. In addition, the different QC rules used meant the use of common QC limits would be problematic as each individual analyzer requires a statistical QC validation to determine the appropriate QC rule. It is noteworthy that for future RPT-QC validations, it may work best to use 40 data points for all calculations (rather than 20) to avoid difficulty in subsequent assessments and to use a 1–2.5 s rule to achieve $P_{ed} > 0.90$ instead of a 1–3 s rule that is capable of >0.85 P_{ed}.[14]

We did see some overlap in the confidence intervals for the upper and lower RPT-QC limits for most measurands for Day1–Day2 and Day1–Day3 time intervals using the Kruskal-Wallis test. Day1–Day4 showed no difference, and this was thought to be due to the non-proportionate deterioration of the specimen across all time intervals as indicated by differing width of the control ranges (see Table 1), and by Day1–Day4, we expect the RPT-QC specimens to be approaching the end of their useful life for RPT-QC.

For RBC, no difference in median differences was seen for Day1–Day3 indicating a varying pattern of deterioration for this

TABLE 6 HGB mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analyzers (3, 4a, 4b) compared to the reference analyzer (1) for time intervals using the same QC rule.

HGB					
Time interval	**Day1–Day2**			**Day1–Day3**	
Analyzer number	**1**	**4a**	**4b**	**1**	**3**
Mean difference	−0.039	−0.113	−0.089	−0.013	−0.029
95% CI (LL, UL)	(−0.065, −0.013)	**(−0.155, −0.071)**	(−0.127, −0.051)	(−0.057, 0.031)	(−0.048, −0.010)
Control limit, upper	0.216	0.289	0.262	0.398	0.154
95% CI (LL, UL)	(0.190, 0.242)	**(0.247, 0.331)**	(0.224, 0.300)	(0.354, 0.442)	**(0.135, 0.173)**
Control limit, lower	−0.294	−0.515	−0.440	−0.424	−0.212
95% CI (LL, UL)	(−0.320, −0.268)	**(−0.557, −0.473)**	**(−0.478, −0.402)**	(−0.468, −0.380)	**(−0.231, −0.193)**

Note: Bold type does not overlap with reference analyzer confidence intervals.
Abbreviations: HGB, hemoglobin; LL, lower limit; RPT-QC, repeat patient testing quality control; UL, upper limit.

		Day1–Day4			
4a	**4b**	**1**	**3**	**4a**	**4b**
−4.420	−4.120	−5.710	−6.000	−5.810	−6.220
(−5.461, −3.379)	(−5.087, −3.153)	(−7.026, −4.394)	(−7.566, −4.434)	(−7.138, −4.482)	(−7.664, −4.776)
3.655	3.380	4.640	6.150	4.765	5.280
(2.614, 4.696)	(2.413, 4.347)	(3.324, 5.956)	(4.584, 7.716)	(3.437, 6.093)	(3.836, 6.724)
−12.495	−11.620	−16.060	−18.150	−16.385	−17.720
(−13.536, −11.454)	(−12.587, −10.653)	(−17.376, −14.744)	(−19.716, −16.585)	(−17.713, −15.057)	(−19.164, −16.276)

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13352 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

TABLE 8 WBC mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analyzers (3, 4a, 4b) compared to the reference analyzer (1) for time intervals using the same QC rule.

WBC								
Time interval	Day1–Day2		Day1–Day3		Day1–Day4			
Analyzer number	1	3	1	4b	1	3	4a	4b
Mean difference	−0.111	−0.060	−0.040	−0.214	−0.556	0.022	0.018	0.072
95% CI (LL, UL)	(−0.162, −0.060)	(−0.152, 0.032)	(−0.093, 0.014)	**(−0.321, −0.107)**	(−0.627, −0.484)	**(−0.063, 0.107)**	**(−0.042, 0.078)**	**(0.001, 0.142)**
Control limit, upper	0.287	0.465	0.383	0.381	0.125	0.724	0.072	0.732
95% CI (LL, UL)	(0.235, 0.338)	**(0.373, 0.557)**	(0.329, 0.437)	(0.274, 0.488)	(0.054, 0.196)	**(0.639, 0.809)**	(0.012, 0.132)	**(0.661, 0.802)**
Control limit, lower	−0.509	−0.585	−0.462	−0.809	−1.237	−0.680	−0.552	−0.588
95% CI (LL, UL)	(−0.560, −0.457)	(−0.677, −0.493)	(−0.516, −0.408)	**(−0.916, −0.702)**	(−1.308, −1.166)	**(−0.765, −0.595)**	**(−0.612, −0.492)**	**(−0.659, −0.518)**

Note: Bold type does not overlap with reference analyzer confidence intervals.

Abbreviations: LL, lower limit; RPT-QC, repeat patient testing quality control; UL, upper limit; WBC, white blood cell.

TABLE 9 Comparison of RPT-QC and commercial QCM over a 4-month period for the reference analyzer (analyzer 1).

	RPT-QC (1–2.5 s or 1–3 s rules based on QC Validation)	Commercial QCM (1–2.5 s rule based on QC Validation)
OOC	4	0
False rejections	4	0
Instability/deterioration	Not studied after Day1–Day 4	By day 4 and/or 5
Ease of use	Moderate (with training)	Moderate (with training)
Shifts/trends	No trends or shifts in data noted; results fluctuate on each side of the mean difference, as expected	Trends or shifts noted by 4-5th day after opening new vial
Supply chain	Readily available but planning required for extended closure periods	Twice delayed delivery in 12 month period
Cost	Technician time and reagents	€1000 annually per analyzer plus technician time and reagents

Abbreviations: OCC, out of control; QCM, quality control material; RPT-QC, repeat patient testing quality control.

measurand that was not as uniform as the other measurands. This nonlinear variability in deterioration also indicates that individual analyzer RPT-QC validation for QC limits is needed. This study supports the need for individual analyzer RPT-QC validation and QC limit generation for each site. The apparent variability of deterioration at each site may reflect different laboratory conditions such as temperature, humidity, maintenance, and operator impact on performance.

This study found that unusual results/trends were more frequent for commercial QCM based on observed performance using a 1–2.5 s QC rule than compared to RPT-QC (see Table 9). However, no OOC results were identified in commercial QCM data over the 4-month period in the reference laboratory. Trends in commercial QCM data were seen toward the end of the open vial shelf life, usually by day 4 and/or day 5, indicating deterioration, although the degree was not considered to be of clinical significance. This resulted in increased time spent on evaluating and rerunning the QCM to determine if any trends were likely to be of significance, even if not outside of the 1–2.5 s range.

RPT-QC had 4 OOC (RBC, HGB) events over 4 months duration for the reference analyzer highlighted by the automated spreadsheet. The use of an RPT-QC troubleshooting checklist failed to demonstrate an underlying cause for these events, and a new RPT-QC specimen was then used to resolve each out-of-control event. Since a new RPT-QC specimen produced in-control results, it is likely that the original OOC result was due to random error associated with an unexpected level of specimen variation, which resulted in an increase in the width of the QC range, which was observed with increasing storage time; instability in the system was not evident when the commercial QCM was run.

Sigma metrics were flagged on two occasions for the reference analyzer using the commercial QCM despite the QC being deemed in control, highlighting poor performance as per criteria established in a previous study[5] for the network of analyzers using a 5.5 sigma quality strategy. This indicates that commercial QCM control limits, even when validated for observed performance, are not reliable as the sole indicator of system performance, and a combined quality strategy is recommended.[11] Sigma metrics for RPT-QC were generated monthly and identified poor performance for HCT using the recommended ASVCP TE_a quality goal. The 'reverse approach' to the determination[18] of achievable total error based on the use of a 1–2.5 s rule, with one level of control material, indicated that total error goals for HCT that can be controlled should be in the range of 22–54% depending on the interval between patient sample

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

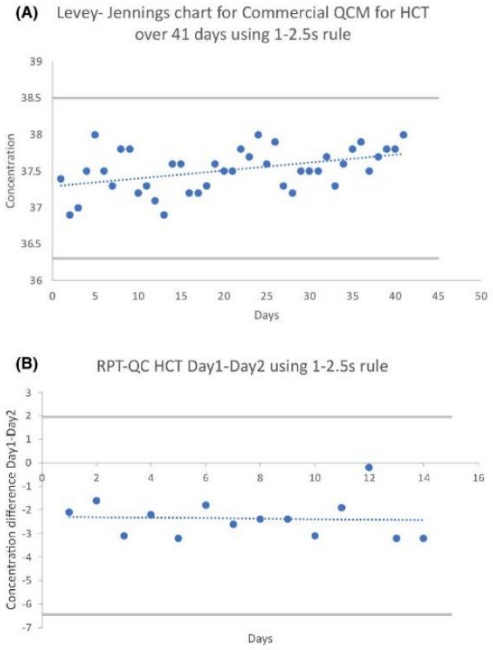


FIGURE 1 (A) Levey-Jennings (L-J) chart for commercial quality control material (QCM) for hematocrit (HCT) using a 1–2.5 s quality control (QC) rule; the x-axis represents days, and the y-axis represents concentration for HCT. The gray lines represent the upper and lower reference intervals. The trendline was added to show a rise in concentration over time. (B) L-J chart for repeat patient testing quality control (RPT-QC) Day1–Day2 for HCT using a 1–2.5 s QC rule; the x-axis represents days, and the y-axis is a difference in concentration for HCT. The gray lines represent the upper and lower RPT-QC limits. The trendline is the mean difference and shows stable concentration over time.

evaluations. Use of these as expert opinion goals resulted in sigma metrics >6 for HCT.

The use of an automated data management spreadsheet for RPT-QC meant less time spent on data entry and analysis than when using manual Excel spreadsheets that were originally used for QC recording. They also facilitated the central electronic archiving of records and sharing of data amongst the laboratories. Manual Excel spreadsheets represent an alternative for point-of-care testing or reference laboratories with a single hematology analyzer. The automated spreadsheet was developed with a laboratory network in mind, but it worked well for the individual site (reference laboratory) where it was initially trialed and refined for any coding issues. The instructions on the spreadsheet prevented erroneous use and worked well as a quick guide to technicians not using RPT-QC frequently. The automated spreadsheet was helpful in ensuring accurate identification of OOC events as these were flagged on the spreadsheet and then transferred to another section where the troubleshooting process was documented. This was useful as it meant only the in-control data were retained for calculating the monthly pooled SD and RPT-QC limits for the subsequent month. The revised RPT-QC limits were usually only slightly altered from the previous month, indicating the validity of the mean replicate difference and calculated RPT-QC limits, but the spreadsheet allowed refinement with increasing data for new RPT-QC limits for the next month. The level of automation freed up the technician from manual calculations and/or coding errors when copying/pasting formulas. The monthly L-J chart and monthly sigma metric were useful in monitoring the ongoing performance of the 5.5 sigma quality assurance strategy.[5]

Overall, the spreadsheet was user-friendly as all data collection, OOC, unusual or scheduled events, and troubleshooting were in one place.

The experience of using RPT-QC in parallel with a commercial QCM was positive. The feedback from the technical team was that the use of the automated data management spreadsheet was helpful. The involvement of all technicians in the development of the automated spreadsheet was also helpful in engaging them in the process and development and using the improvements suggested. It had all the required capabilities needed to carry out RPT-QC monitoring. A disadvantage is that the spreadsheet design involved coding that was outside the capabilities of the laboratory technicians and, therefore, any retrospective or prospective changes would need to be carried out by personnel proficient in this area. An outside consultant (second author) was required for proficiency in advanced automation of spreadsheets based on the technicians' feedback.

As with any procedure in the laboratory, a SOP was drafted. This aided the technical team with training and was used in the rollout of RPT-QC to the laboratory network.

Supply chain issues with commercial QCM were noted during COVID and following Brexit, further supporting the need for alternative QC methods. The availability and elimination of the cost of commercial QCM are major advantages of RPT-QC. RPT-QC does not incur costs for the laboratory as it uses excess patient samples. Any disadvantage for RPT-QC resulting from extended closures for holidays beyond the Day1–Day4 time interval is mitigated by having an alternative protocol that may be used during periods of long closure. The spreadsheet was also designed to allow the use of alternative time intervals where Day1–Day2 specimens are not widely available.

In addition, the canine RPT-QC specimens are a suitable commutable matrix for our veterinary laboratory network and a representative matrix for veterinary patients.[2] Canine specimens are rarely unavailable, and various levels of clinical interest could also be used. Controlling costs in the commercial laboratory is a consideration and reducing cost by approximately €1000 per analyzer (for a single level of commercial QCM) per year by implementing RPT-QC results in substantial savings for a laboratory. The cost savings are an added benefit following all of the above. In summary, the major advantages of using RPT-QC are cost savings, a QC matrix commutable with patient samples, the ability to refine control limits based on the use of a pooled SD_{dup} over time, and the use of an automated spreadsheet.

19391653, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

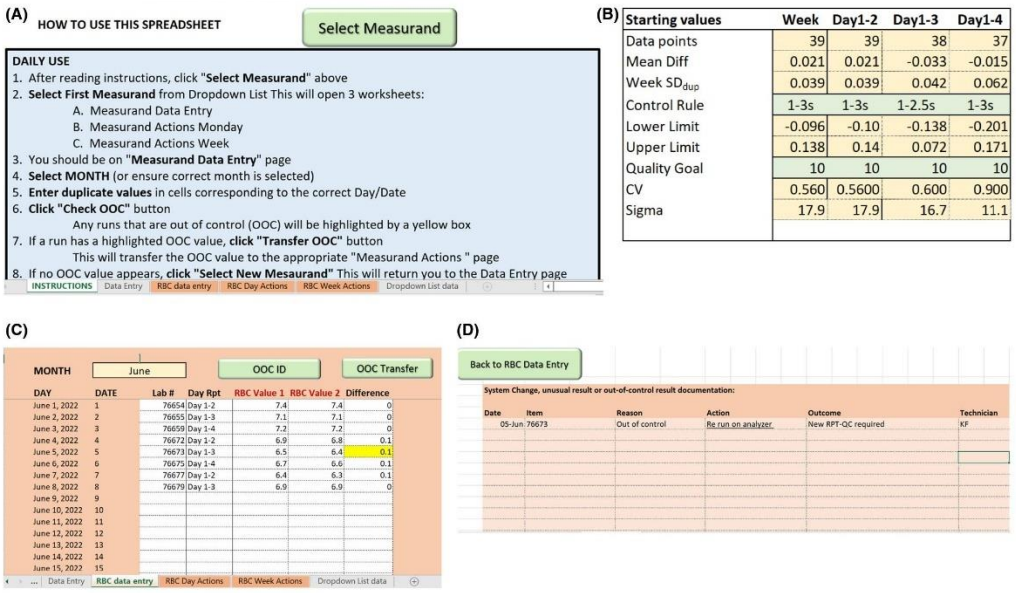


FIGURE 2 (A) Image of the automated data management spreadsheet user instructions and toolbar for RBC measurand. (B) Initial calculation summary for RBC for each time interval Day1–Day2, Day1–Day3, Day1–Day4, and the pooled monthly SD and refined control limits. (C) Data entry for RBC showing one out-of-control event (Lab#76673) that alerts the technician by highlighting the data point required for further troubleshooting. (D) Troubleshooting page where actions taken can be logged, which is a permanent record detailing the reason, action, outcome, and responsible technician.

We chose to evaluate RBC, HGB, HCT, and WBC, in particular for RPT-QC, to streamline the process of RPT-QC and considered these adequate due to the high emphasis on nonstatistical QC for hematology, including technical review of blood smears. We elected not to include the calculated RBC indices (MCV, MCH, MCHC) due to the effects of specimen aging. RPT-QC should also work for other hematology measurements, but further review of the effects of specimen aging on platelets, reticulocytes, and WBC differential counts is needed. The technical review of the blood smear was considered adequate to QC these methods, with consideration of the correlation of the microscopic findings with the automated hematology patient results.

5 | CONCLUSIONS

This study showed that RPT-QC can be used within single veterinary laboratories and across a laboratory network to function as a commutable and cost-effective method of quality control. We demonstrated statistically that RPT-QC requires individual QC limit generation for each analyzer because some median differences in the repeated results are different across the time intervals and the network analyzers. RPT-QC does not have excessive repeats, shifts, or trends in data compared with QC based on observed performance using a commercial QCM. The positive feedback from technicians using the automated data management spreadsheet meant all data entry and monitoring was in one place, and L-J charts were a great visual aid and summary of results, necessitating the need for spreadsheet automation. The refined SD_{dup} will require further investigation, but early analyses indicated the QC limit generation from pooled monthly SDs in normal RPT-QC will result in ongoing refinements of RPT-QC limits, leading to improved application of quality control.

CONFLICT OF INTEREST STATEMENT

The authors have indicated that they have no affiliations or financial involvement with any organization or entity with a financial interest in, or in financial competition with, the subject matter or materials discussed in this article.

ORCID

S. Daly https://orcid.org/0000-0003-2627-9735
M. Rishniw https://orcid.org/0000-0002-0477-1780
K. P. Freeman https://orcid.org/0000-0003-1796-0158

REFERENCES

1. Flatland B, Vap LM. Quality management recommendations for automated and manual in-house hematology of domestic animals.

19391656, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D

Vet Clin North Am Small Anim Pract. 2012;42(1):11-22. doi:10.1016/j.cvsm.2011.09.004

2. Unsal I, Coskun A, Serteser M, Inal TC, Ozpinar A. Toward standardization of quality assessment in laboratory medicine by using the same matrix samples for both internal and external quality assessments. *Accreditation and Qual Assur*. 2010;15(11):621-627. doi:10.1007/s00769-010-0694-5
3. Daly S, Freeman KP, Graham PA. Evaluation of Sysmex XT-2000iV analyzer performance across a network of five veterinary laboratories using a commercially available quality control material. *Vet Clin Pathol*. 2021;50(4):568-578. doi:10.1111/vcp.13016
4. Manzocchi S, Furman E, Freeman K. Comparison of 3 options for choosing control limits in biochemistry testing. *Vet Clin Pathol*. 2017;46(1):120-125. doi:10.1111/vcp.12453
5. Daly S, Graham PA, Freeman KP. Quality control validation for a veterinary laboratory network of six Sysmex XT-2000iV hematology analyzers. *Vet Clin Pathol*. 2022;51(4):565-576. doi:10.1111/vcp.13163
6. Westgard J, Basic QC. *Practices: Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC, Inc; 2016:233-244.
7. Dutra FR. Monitoring the quality of blood cell counts with replicate determinations on routine samples. *Am J Clin Pathol*. 1966;46(2_ts):286-288. doi:10.1093/ajcp/46.2_ts.286
8. Carstairs KC, Peters E, Kuzin EJ. Development and description of the "random duplicates" method of quality control for a hematology laboratory. *Am J Clin Pathol*. 1977;67(4):379-385. doi:10.1093/ajcp/67.4.379
9. Orser B. Hematology quality control and the coulter S-plus II. *Am J Med Technol*. 1983;49(9):643-647.
10. Riddick JH, Johnson CL. Hematology control using computerized range charts. *Lab Med*. 1972;3:32-33.
11. Westgard J. Implementing repeat patient test controls. *Basic QC Practices: Training in Statistical Quality Control for Medical Laboratories*. 4th ed. Westgard QC, Inc; 2016.
12. Flatland B, Freeman KP. Repeat patient testing shows promise as a quality control method for veterinary hematology testing. *Vet Clin Pathol*. 2018;47(2):252-266. doi:10.1111/vcp.12593
13. Vanyo LC, Freeman KP, Meléndez-Lazo A, Teles M, Cuenca R, Pastor J. Comparison of traditional statistical quality control using commercially available control materials and two patient-based quality control procedures for the ADVIA 120 hematology system. *Vet Clin Pathol*. 2018;47(3):368-376. doi:10.1111/vcp.12645
14. Daly S, Graham PA, Freeman K. Repeat patient testing-quality control with canine samples shows promise as an alternative to commercial quality control material for a network of 4 Sysmex XT2000iV hematology analyzers. *Vet Clin Pathol*. 2023;00:1-9. doi:10.1111/vcp.13211
15. Nabity M, Harr K, Camus M, Flatland B, Vap L. ASVCP guidelines: allowable total error hematology. *Vet Clin Pathol*. 2018;47:9-21. doi:10.1111/vcp.12583
16. Wheelan C. *Naked Statistics: Stripping the Dread from the Data*. W. W. Norton; 2013.
17. Ostertagová E, Ostertag O, Kováč J. Methodology and application of the Kruskal-Wallis test. *Appl Mech Mater*. 2014;611:115-120. doi:10.4028/www.scientific.net/AMM.611.115
18. Korchia J, Freeman KP. Total observed error, total allowable error, and QC rules for canine serum and urine cortisol achievable with the Immulite 2000 Xpi cortisol immunoassay. *J Vet Diagn Invest*. 2022;34(2):246-257. doi:10.1177/10406387221076129

How to cite this article: Daly S, Rishniw M, Graham PA, Freeman KP. Repeat patient testing-quality control compared to commercial quality control material for the Sysmex XT-2000iV hematology analyzer in a multi-site veterinary laboratory. *Vet Clin Pathol*. 2024;53(Suppl. 1):48-59. doi:10.1111/vcp.13332

19390165x, 2024, S1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/vcp.13332 by Health Research Board, Wiley Online Library on [21/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX E – SUPPLEMENTARY TABLES

Supplementary table list:

Chapter 3

Table E 3.1: Summary of the haematological erythroid measurand QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020).

Table E3.2: Summary of Platelet (PLT) and Plateletcrit (PCT) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020).

Table E3.3: Summary of White Blood Cell Count (WBC) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020).

Chapter 5

Table E5.1: SD, CV, mean difference and RPT-QC limits for the reference (1) and network analysers (3, 4a, 4b) for each time interval for RBC, HGB, HCT, WBC.

Table E5.2: Example calculations for RBC Day1-Day4, n=20 with outlier (Y) in row 14 in bold type.

Table E5.3: RBC mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

Table E5.4: HGB mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

Table E5.5: HCT mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

Table E5.6: WBC mean difference, upper and lower RPT-QC limits and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

Table E5.7: Comparison of RPT-QC and commercial QCM over a 4-month period for the reference analyser (analyser 1).

Chapter 6

Table E6.1: Mean and CV(%) range for the erythroid measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

Table E6.2: Mean and CV(%) range for the leukocyte measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

Table E6.3: Mean and CV(%) range for the platelet measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

Chapter 7

Table E7.1: The number (n) of each measurand evaluated on 7 point of care haematology analysers using different canine EDTA samples that could achieve > 4 sigma metric and using the QGI whether those that didn't achieve > 4 sigma was due to bias, imprecision or both bias and imprecision.

CHAPTER 3 TABLES

Table E3. 1: Summary of the haematological erythroid measurand QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All instruments used the same quality control material lot number. The number of quality control materials is one (n=1). The number of quality control materials is one (n=1). Ped < 0.85 are highlighted in bold type.

	Analyser 1		Analyser 2		Analyser 3		Analyser 4a		Analyser 4b		Analyser 5	
	RBC											
	(x10^6/uL)											
TEa goal %	10		10		10		10		10		10	
Observed mean	4.32		4.37		4.34		4.39		4.39		4.35	
Target mean	4.32		4.32		4.32		4.32		4.32		4.32	
Observed CV %	0.69		0.49		0.68		1.60		0.70		1.00	
Observed absolute bias %	0.10		1.13		0.40		1.60		1.58		0.64	
Observed TE	1.48		2.11		1.76		4.8		2.98		2.64	
Sigma metric	14.35		18.1		14.12		5.25		12.03		9.36	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s

Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	0.88	**0.76**	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
					HGB							
					(g/dL)							
TEa goal %	10		10		10		10		10		10	
Observed mean	12.08		11.90		11.98		12.21		12.22		12.11	
Target mean	12		12		12		12		12		12	
Observed CV %	0.59		0.37		0.32		0.47		0.44		0.90	
Observed absolute bias %	0.63		0.80		0.15		1.79		1.87		0.90	
Observed TE	1.81		1.54		0.79		2.73		2.75		2.7	
Sigma metric	15.88		24.86		30.78		17.47		18.48		10.11	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
TEa goal %	10		10		10		10		10		10	

						HCT (%)						
TEa goal %	10		10		10		10		10		10	
Observed mean	36.1		36.5		35.9		36.3		36.6		36.3	
Target mean	35.7		35.7		35.7		35.7		35.7		35.7	
Observed CV %	0.64		0.56		1.06		0.55		0.47		1.01	
Observed absolute bias %	1.14		2.13		0.43		1.78		2.40		2.05	
Observed TE	2.42		3.25		2.55		2.88		3.34		4.07	
Sigma metric	13.84		14.11		9.03		14.95		16.17		7.87	
Control rules, n=1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
						MCV (fL)						
TEa goal %	7		7		7		7		7		7	
Observed mean	83.5		83.4		82.7		83.7		83.4		83.8	

Target mean	82.6		82.6		82.6		82.6		82.6		82.6	
Observed CV %	0.43		0.40		0.69		0.55		0.57		0.62	
Observed absolute bias %	1.09		1.04		0.08		0.92		0.92		1.45	
Observed TE	1.95		1.84		1.46		2.02		2.06		2.69	
Sigma metric	13.74		14.90		10.03		11.05		11.72		8.95	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
MCHC (g/dL)												
TEa goal %	10		10		10		10		10		10	
Observed mean	33.4		32.6		33.4		33.5		33.3		33.3	
Target mean	33.6		33.6		33.6		33.6		33.6		33.6	
Observed CV %	0.70		0.72		1.11		0.92		0.74		0.83	
Observed absolute bias %	0.50		0.54		0.54		0.27		0.68		1.08	

Observed TE	1.9		1.98		2.76		2.11		2.16		2.74	
Sigma metric	13.57		13.14		8.52		10.58		12.59		11.89	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
RDW-CV (%)												
TEa goal %	10		10		10		10		10		10	
Observed mean	16.1		16.2		16.1		16.2		16.3		16.3	
Target mean	16.2		16.2		16.2		16.2		16.2		16.2	
Observed CV %	0.48		0.75		0.49		0.69		0.56		0.64	
Observed absolute bias %	0.69		0.09		0.33		0.02		0.59		0.57	
Observed TE	1.65		1.59		1.31		1.4		1.71		1.85	
Sigma metric	19.40		13.21		19.73		14.46		16.80		14.73	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85

Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
					RETIC							
					(x10^6/uL)							
TEa goal %	40		40		40		40		40		40	
Observed mean	70.1		73.9		71.8		76.5		77.0		75.9	
Target mean	74.3		74.3		74.3		74.3		74.3		74.3	
Observed CV %	5.11		4.58		4.30		4.87		6.07		4.73	
Observed absolute bias %	4.65		3.37		3.37		2.91		3.72		0.44	
Observed TE	14.87		12.53		11.97		12.65		15.86		9.9	
Sigma metric	6.48		8		8.52		7.62		5.98		8.36	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

RBC = Red blood cell count; HBG = Haemoglobin; HCT = Haematocrit; MCV = Mean Corpuscular Volume; MCHC = Mean Corpuscular Haemoglobin Concentration; RDW-CV = coefficient of variation of the Red Cell Distribution Width; RETIC = Reticulocyte count; TEa = Total allowable error; CV = coefficient of variation; Ped = probability of error detection; Pfr = probability of false rejection; TE = total error

Table E3. 2: Summary of Platelet (PLT) and Plateletcrit (PCT) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All instruments used the same quality control material lot number. The number of quality control materials is one (n=1). Ped < 0.85 is highlighted by bold type.

	Analyser 1		Analyser 2		Analyser 3		Analyser 4a		Analyser 4b		Analyser 5	
	PLT											
	(x10^3/uL)											
TEa goal %	20		20		20		20		20		20	
Observed mean	220.4		220.0		209.4		217.2		224.1		213.1	
Target mean	220		220		220		220		220		220	
Observed CV %	2.61		2.21		2.23		3.72		2.48		2.21	
Observed absolute bias %	0.16		0.00		4.81		0.93		1.87		3.13	
Observed TE	5.38		4.42		9.27		8.37		6.83		7.55	
Sigma metric	7.6		9.05		6.81		5.13		7.31		7.63	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	0.86	**0.72**	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

	PCT (%)											
TEa goal %	25		25		25		25		25		25	
Observed mean	0.21		0.22		0.20		0.21		0.21		0.20	
Target mean	0.21		0.21		0.21		0.21		0.21		0.21	
Observed CV %	3.09		3.27		2.38		5.31		2.98		2.83	
Observed absolute bias %	2.02		0.91		6.16		0.76		1.45		2.20	
Observed TE	8.2		7.45		10.92		11.38		7.41		7.86	
Sigma metric	7.44		7.37		7.92		4.56		7.9		8.06	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	**0.69**	**0.50**	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

PLT = Platelet count; PCT = Plateletcrit; TEa = Total allowable error; CV = coefficient of variation; Ped = probability of error detection; Pfr = probability of false rejection; TE = total error

Table E3. 3: Summary of White Blood Cell Count (WBC) QC Validation for 6 Sysmex XT 2000iV haematology analysers based on observed performance (March 2020). All instruments used the same quality control material lot number. The number of quality control materials is one (n=1).

	Analyser 1		Analyser 2		Analyser 3		Analyser 4a		Analyser 4b		Analyser 5	
					WBC							
					(x10^3/uL)							
TEa goal %	15		15		15		15		15		1	
Observed mean	7.27		7.17		7.12		7.10		7.16		7.16	
Target mean	7.11		7.11		7.11		7.11		7.11		7.11	
Observed CV %	1.19		1.82		2.31		1.62		1.78		1.38	
Observed absolute bias %	2.33		0.82		0.18		0.04		0.77		0.70	
Observed TE	4.71		4.46		4.8		3.28		4.33		3.46	
Sigma metric	10.65		7.79		6.42		9.23		7.99		10.36	
Control rules, n =1	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s	1-2.5s	1-3s
Ped	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85	> 0.94	> 0.85
Pfr	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00

WBC = White Blood Cell count; TEa = Total allowable error; CV = coefficient of variation; Ped = probability of error detection; Pfr = probability of false rejection; TE = total error

CHAPTER 5 TABLES

Table E5. 1: SD, CV, mean difference and RPT-QC limits for the reference (1) and network analysers (3, 4a, 4b) for each time interval for RBC, HGB, HCT, WBC.

Analyser	1			3			4a			4b		
Time Interval	**Day 1-2**	**Day 1-3**	**Day 1-4**	**Day 1-2**	**Day 1-3**	**Day 1-4**	**Day 1-2**	**Day 1-3**	**Day 1-4**	**Day 1-2**	**Day 1-3**	**Day 1-4**
						RBC						
SD	0.039	0.042	0.062	0.061	0.104	0.084	0.033	0.049	0.039	0.082	0.076	0.065
CV	0.560	0.600	0.900	0.859	1.470	1.180	0.500	0.750	0.577	1.210	1.110	0.960
Mean difference	0.021	-0.033	-0.051	-0.005	0.007	0.003	0.008	-0.013	-0.025	-0.056	-0.034	-0.017
Control limits (LL,UL)	(-0.096, 0.138)	(-0.138, 0.072)	(-0.237, 0.135)	(-0.188, 0.178)	(-0.253, 0.267)	(-0.248, 0.255)	(-0.091, 0.107)	(-0.160, 0.134)	(-0.142, 0.092)	(-0.261, 0.149)	(-0.262, 0.194)	(-0.180, 0.146)
						HGB						
SD	0.085	0.137	0.083	0.081	0.061	0.063	0.134	0.152	0.168	0.117	0.127	0.164
CV	0.500	0.477	0.500	0.566	0.734	0.386	0.847	0.965	1.070	0.715	0.793	1.000
Mean difference	-0.039	-0.013	-0.003	-0.034	-0.029	-0.042	-0.113	-0.108	-0.154	-0.089	-0.144	-0.127
Control limits (LL,UL)	(-0.294, 0.216)	(-0.424, 0.398)	(-0.252, 0.246)	(-0.237, 0.169)	(-0.212, 0.154)	(-0.200, 0.116)	(-0.515, 0.289)	(-0.488, 0.272)	(-0.574, 0.266)	(-0.440, 0.262)	(-0.462, 0.174)	(-0.537, 0.283)
						HCT						

SD	1.670	2.670	4.140	2.250	3.230	4.860	1.960	3.230	4.230	2.210	3.000	4.600
CV	3.470	5.410	8.240	4.470	6.210	8.340	3.940	6.370	8.230	4.340	5.840	8.780
Mean difference	-2.250	-3.980	-5.710	-3.150	-4.760	-6.000	-2.660	-4.420	-5.810	-2.890	-4.120	-6.220
Control limits (LL,UL)	(-6.425, 1.925)	(-10.655, 2.695)	(-16.06, 4.640)	(-8.775, 2.475)	(-12.835, 3.315)	(-18.150, 6.150)	(-7.560, 2.240)	(-12.495, 3.655)	(-16.385, 4.765)	(-8.415, 2.635)	(-11.62, 3.380)	(-17.72, 5.280)
						WBC						
SD	0.159	0.169	0.227	0.210	0.175	0.234	0.150	0.150	0.190	0.278	0.238	0.220
CV	1.290	1.510	1.850	2.040	1.760	2.470	1.690	1.750	2.120	2.880	2.730	2.350
Mean difference	-0.111	-0.040	-0.556	-0.060	0.098	0.022	0.009	-0.071	0.018	0.016	-0.214	0.072
Control limits (LL,UL)	(-0.508, 0.287)	(-0.462, 0.383)	(-1.237, 0.125)	(-0.585, 0.465)	(-0.427, 0.623)	(-0.680, 0.724)	(-0.441, 0.459)	(-0.521, 0.379)	(-0.552, 0.072)	(-0.818, 0.064)	(-0.809, 0.381)	(-0.588, 0.732)

Table E5. 2: Example calculations for RBC Day1-Day4, n=20 with outlier (Y) in row 14 in bold type.

N	Day 1	Day 4	Average Day1-Day 4	Difference D1-D4	Difference squared
1	7.120	7.190	7.155	-0.070	0.005
2	7.460	7.570	7.515	-0.110	0.012
3	5.210	5.210	5.210	0.000	0.000
4	6.450	6.450	6.450	0.000	0.000
5	6.980	7.050	7.015	-0.070	0.005
6	5.750	5.920	5.835	-0.170	0.029
7	7.400	7.450	7.425	-0.050	0.003
8	5.760	5.780	5.770	-0.020	0.000
9	6.830	6.790	6.810	0.040	0.002
10	7.770	7.700	7.735	0.070	0.005
11	7.840	7.880	7.860	-0.040	0.002
12	5.100	5.210	5.155	-0.110	0.012
13	8.060	8.130	8.095	-0.070	0.005
14	**8.320**	**8.530**	**8.425**	**-0.210**	**0.044**
15	6.420	6.530	6.475	-0.110	0.012
16	6.760	6.680	6.720	0.080	0.006
17	7.230	7.380	7.305	-0.150	0.023
18	7.550	7.710	7.630	-0.160	0.026
19	8.230	8.210	8.220	0.020	0.000
20	5.650	5.560	5.605	0.090	0.008
Averages	6.895	6.947	6.921	-0.052	0.010
without Y	6.819	6.863	6.841	-0.044	0.008
SD_{dup}	0.070	**without Y**	0.052		
CV %	1.017	**without Y**	0.753		
Sum Diff sq	0.198	**without Y**	0.154		

Note: Sum Diff sq = Sum of the difference squared; SD_{dup} = Standard deviation of the duplicate measurements.

Table E5. 3: RBC mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

RBC

Time interval	**Day1-Day2**			**Day1-Day3**		**Day1-Day4**		
Analyser number	**1**	**3**	**4a**	**1**	**3**	**1**	**3**	**4a**
Mean difference	0.021	-0.005	0.008	-0.033	0.007	-0.051	0.003	-0.025
95% CI (LL,UL)	(0.009, 0.033)	(-0.024, 0.014)	(-0.007, 0.023)	(-0.046, -0.02)	(-0.026, 0.040)	(-0.071, -0.031)	**(-0.023, 0.030)**	(-0.044, -0.006)
Control limit, upper	0.138	0.178	0.107	0.072	0.267	0.135	0.255	0.092
95% CI (LL, UL)	(0.126, 0.150)	**(0.159, 0.197)**	**(0.092, 0.122)**	(0.059, 0.085)	**(0.234, 0.300)**	(0.115, 0.155)	**(0.229, 0.282)**	(0.073, 0.111)
Control limit, lower	-0.096	-0.188	-0.091	-0.138	-0.253	-0.237	-0.249	-0.142
95% CI (LL, UL)	(-0.108, -0.084)	**(-0.207, -0.169)**	(-0.106, -0.076)	(-0.151, -0.125)	**(-0.286, -0.220)**	(-0.257, -0.217)	(-0.275, -0.222)	**(-0.161, -0.123)**

Bold type does not overlap with reference analyzer confidence intervals;

RBC, red blood cell; RPT-QC, repeat patient testing quality control; LL, lower limit; UL, upper limit.

Table E5. 4: HGB mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

HGB					
Time interval	**Day1-Day2**			**Day1-Day3**	
Analyser number	**1**	**4a**	**4b**	**1**	**3**
Mean difference	-0.039	-0.113	-0.089	-0.013	-0.029
95% CI (LL,UL)	(-0.065, -0.013)	**(-0.155, -0.071)**	(-0.127, -0.051)	(-0.057, 0.031)	(-0.048, -0.010)
Control limit, upper	0.216	0.289	0.262	0.398	0.154
95% CI (LL, UL)	(0.190, 0.242)	**(0.247, 0.331)**	(0.224, 0.300)	(0.354, 0.442)	**(0.135, 0.173)**
Control limit, lower	-0.294	-0.515	-0.440	-0.424	-0.212
95% CI (LL, UL)	(-0.320, -0.268)	**(-0.557, -0.473)**	**(-0.478, -0.402)**	(-0.468, -0.380)	**(-0.231, -0.193)**

Bold type does not overlap with reference analyzer confidence intervals

HGB, haemoglobin; RPT-QC, repeat patient testing quality control; LL, lower limit; UL, upper limit.

Table E5. 5: HCT mean difference, upper and lower RPT-QC limits, and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

	HCT											
Time interval	**Day1-Day2**				**Day1-Day3**				**Day1-Day4**			
Analyser number	**1**	**3**	**4a**	**4b**	**1**	**3**	**4a**	**4b**	**1**	**3**	**4a**	**4b**
Mean difference	-2.250	-3.150	-2.660	-2.890	-3.980	-4.760	-4.420	-4.120	-5.710	-6.000	-5.810	-6.220
95% CI (LL,UL)	(-2.982, -1.518)	(-4.189, -2.111)	(-3.283, -2.037)	(-3.593, -2.187)	(-4.840, -3.120)	(-6.252, -3.268)	(-5.461, -3.379)	(-5.087, -3.153)	(-7.026, -4.394)	(-7.566, -4.434)	(-7.138, -4.482)	(-7.664, -4.776)
Control limit, upper	1.925	2.475	2.240	2.635	2.695	3.315	3.655	3.380	4.640	6.150	4.765	5.280
95% CI (LL, UL)	(1.193, 2.657)	(1.436, 3.514)	(1.617, 2.863)	(1.932, 3.338)	(1.835, 3.555)	(1.823, 4.807)	(2.614, 4.696)	(2.413, 4.347)	(3.324, 5.956)	(4.584, 7.716)	(3.437, 6.093)	(3.836, 6.724)
Control limit, lower	-6.425	-8.775	-7.560	-8.415	-10.655	-12.835	-12.495	-11.620	-16.060	-18.150	-16.385	-17.720
95% CI (LL, UL)	(-7.157, -5.693)	**(-9.814, -7.736)**	(-8.183, -6.937)	**(-9.118, -7.712)**	(-11.515, -9.795)	(-14.327, -11.343)	(-13.536, -11.454)	(-12.587, -10.653)	(-17.376, -14.744)	(-19.716, -16.585)	(-17.713, -15.057)	(-19.164, -16.276)

Bold type does not overlap with reference analyser confidence intervals

HCT, haematocrit; RPT-QC, repeat patient testing quality control; LL, lower limit; UL, upper limit

Table E5. 6: WBC mean difference, upper and lower RPT-QC limits and confidence intervals (CI) for the network analysers (3, 4a, 4b) compared to the reference analyser (1) for time intervals using the same QC rule.

	WBC							
Time interval	**Day1-Day2**		**Day1-Day3**		**Day1-Day4**			
Analyser number	**1**	**3**	**1**	**4b**	**1**	**3**	**4a**	**4b**
Mean difference	-0.111	-0.060	-0.040	-0.214	-0.556	0.022	0.018	0.072
95% CI (LL,UL)	(-0.162, -0.060)	(-0.152, 0.032)	(-0.093, 0.014)	**(-0.321, -0.107)**	(-0.627, -0.484)	**(-0.063, 0.107)**	**(-0.042, 0.078)**	**(0.001, 0.142)**
Control limit, upper	0.287	0.465	0.383	0.381	0.125	0.724	0.072	0.732
95% CI (LL, UL)	(0.235, 0.338)	**(0.373, 0.557)**	(0.329, 0.437)	(0.274, 0.488)	(0.054, 0.196)	**(0.639, 0.809)**	(0.012, 0.132)	**(0.661, 0.802)**
Control limit, lower	-0.509	-0.585	-0.462	-0.809	-1.237	-0.680	-0.552	-0.588
95% CI (LL, UL)	(-0.560, -0.457)	(-0.677, -0.493)	(-0.516, -0.408)	**(-0.916, -0.702)**	(-1.308, -1.166)	**(-0.765, -0.595)**	**(-0.612, -0.492)**	**(-0.659, -0.518)**

Bold type does not overlap with reference analyser confidence intervals

WBC, white blood cell; RPT-QC, repeat patient testing quality control; LL, lower limit; UL, upper limit

Table E5. 7: Comparison of RPT-QC and commercial QCM over a 4-month period for the reference analyzer (analyzer 1).

	RPT-QC (1-2.5s or 1-3s rules based on QC Validation)	**Commercial QCM (1-2.5s rule based on QC Validation)**
OOC	4	0
False rejections	4	0
Instability/deterioration	Not studied after Day1-Day 4	By day 4 and/or 5
Ease of use	Moderate (with training)	Moderate (with training)
Shifts/trends	No trends or shifts in data noted; results fluctuate on each side of the mean difference, as expected	Trends or shifts noted by 4-5th day after opening new vial
Supply chain	Readily available but planning required for extended closure periods	Twice delayed delivery in 12 month period
Cost	Technician time and reagents	€1000 annually per analyzer plus technician time and reagents

RPT-QC, repeat patient testing quality control; QCM, quality control material; OCC, out of control

CHAPTER 6 TABLES

Table E6. 1: Mean and CV(%) range for the erythroid measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

Erythroid Measurands	Very low bin (target HCT 10-15%)		Low bin (target HCT 20-25%)		Normal bin (target HCT 35-60%)		High bin (target HCT 55-65%)		Very high (target HCT 70-75%)	
	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)
RBC (10^{12}/l)	1.26-1.78	0.46-0.93	2.09-2.93	0.54-0.81	5.45-7.33	0.407-0.77	7.71-9.0	0.35-0.47	9.75-14.16	0.25-0.47
HGB (g/dl)	3.11-4.09	0.77-1.70	5.10-6.93	0.98-1.14	12.56-17.47	0.27-0.85	17.99-21.74	0.31-0.38	22.9-29.6	0.27-0.44
HCT (%)	11.09-14.94	0.345-0.806	17.6-20.5	0.56-0.75	38.56-51.76	0.357-0.63	58.9-65.4	0.22-0.36	71.3-80.6	0.16-0.48
MCV (fl)	66.41-89.5	0.228-0.483	69.8-84.2	0.22-0.33	67.68-81.19	0.365-0.84	68.2-78.8	0.17-0.60	54.26-79.3	0.083-0.40
MCH (pg)	21.83-25.9	0.563-1.58	18.7-25.0	0.64-1.17	23.0-24.81	0.40-0.87	22.3-24.3	0.39-0.59	20.92-24.34	0.31-0.68
MCHC (g/dl)	27.31-34.77	1.09-1.74	26.7-33.7	0.53-1.29	29.42-35.32	0.43-1.18	29.5-35.7	0.32-0.63	29.9-38.55	0.33-0.75
RDW-CV	11.44-14.15	0.359-0.600	12.33-16.35	0.39-0.65	13.83-16.42	0.488-1.83	14.8-21.0	0.33-0.87	19.65-25.8	0.29-0.72
RETIC#	5.45-19.94	2.53-22.41	1.32-44.1	7.16-13.67	5.11-87.45	5.48-12.65	8.34-82.6	5.43-8.50	12.08-272.01	4.79-8.94

CV = Coefficient of variation; RBC = Red blood cell count; HGB = haemoglobin; HCT = Haematocrit; MCV = Mean corpuscular volume, MCH = Mean corpuscular haemoglobin; MCHC = Mean corpuscular haemoglobin concentration; RDW-CV = Red cell distribution width-coefficient of variation; RETIC = Reticulocyte number.

Table E6. 2: Mean and CV(%) range for the leukocyte measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

Leucocyte Measurands	Very low bin (target HCT 10-15%)		Low bin (target HCT 20-25%)		Normal bin (target HCT 35-60%)		High bin (target HCT 55-65%)		Very high (target HCT 70-75%)	
	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)	Mean	CV(%)
WBC (10^3/ul)	1.88-3.13	2.67-3.590	2.35-9.30	1.17-3.25	7.32-12.74	1.37-2.23	8.24-23.74	0.99-1.65	7.72-22.0	1.14-4.21
NEUT# (10^9/l)	1.24-2.41	2.628-4.327	1.35-7.75	1.30-4.56	4.61-10.03	1.70-2.59	5.40-19.0	0.82-1.99	5.57-16.6	1.25-3.63
LYMPH# (10^9/l)	0.37-0.673	4.432-9.67	0.44-1.41	3.15-11.03	1.22-2.90	1.23-7.39	1.92-3.10	2.30-3.76	1.04-20.64	2.92-7.08
MONO# (10^9/l)	0.12-0.26	11.75-18.86	0.09-0.62	5.60-18.0	0.227-1.08	5.13-23.29	0.48-1.44	3.17-12.3	0.40-1.19	3.85-20.45
EO# (10^9/l)	0.05-0.35	7.35-107.9	0.06-0.25	7.85-45.5	0.11-0.552	7.59-30.0	0.19-0.86	3.13-23.6	0.31-0.47	3.53-11.70
BASO# (10^9/l)	0.0-0.01	0.0-0.0	0-0.015	n/a	0.0-0.012	0.0-35.13	0.005-0.03	n/a	0.006-0.4	19.36-80.06

CV = Coefficient of variation; HCT = Haematocrit; WBC = White blood cell count; NEUT = Neutrophil count; LYMPH = Lymphocyte count; MONO = Monocyte count; EO = Eosinophil count; BASO = basophil count.

Table E6. 3: Mean and CV(%) range for the platelet measurands generated from replicate analyses of 5 different 'bins' of results (target HCT (%) from manipulated canine samples) across 6 Sysmex XT-2000iV analysers.

	Very low bin (target HCT 10-15%)		Low bin (target HCT 20-25%)		Normal bin (target HCT 35-60%)		High bin (target HCT 55-65%)		Very high (target HCT 70-75%)	
	Mean	**CV(%)**	**Mean**	**CV(%)**	**Mean**	**CV(%)**	**Mean**	**CV(%)**	**Mean**	**CV(%)**
PLT (10^9/l)	38.1-112.5	1.88-6.71	42.2-156.2	1.87-9.57	125-357	1.45-7.39	120.9-378.6	1.01-4.63	328.7-408.4	1.80-5.09
PCT (%)	0.05-0.139	0.0-10.34	0.049-0.153	3.0-11.8	0.147-0.431	1.74-7.20	0.319-0.428	1.47-4.41	0.39-0.458	1.90-7.77

Note PCT high bin calculated from 4/6 analysers

CV = Coefficient of variation; HCT = Haematocrit; PLT = Platelet count; PCT = Plateletcrit (%).

CHAPTER 7 TABLES

Table E7. 1: The number (n) of each measurand evaluated on 7 point of care haematology analysers using different canine EDTA samples that could achieve > 4.0 sigma metric and using the QGI whether those that didn't achieve > 4 sigma was due to bias, imprecision or both bias and imprecision.

Measurand (Units)	*n*	Sigma metric > 4.0	Bias	Imprecision	Both
RBC (x 10^{12}/L)	7	2	5	0	0
HGB (g/L)	7	6	1	0	0
HCT (%)	7	4	2	1	0
MCV (fL)	7	5	2	0	0
MCH (pg)	7	4	3	0	0
MCHC (g/dL)	7	4	1	2	0
WBC (x 10^9/L)	7	3	4	0	0
Platelets (x 10^9/L)	7	3	4	0	0
RDW-CV (%)	5	3	2	0	0
MPV (fL)	7	3	4	0	0

Note: these are based on initial results prior to harmonisation

APPENDIX F – SUPPLEMENTARY FIGURES

Supplementary Figure list:

Chapter 4

Figure F4.1: An example of data calculations using 20 data points for HGB Day1-Day2 showing D1 (day1) measurements, D2 (day2) measurements, Diff (difference from Day1 minus Day2), Aver D1,D2 (average taken of Day1 and Day2 measurement) and Diff sq (Difference squared).

Chapter 6

Figure F6.1 (A-E): Bar charts of single ILC results for the erythroid measurands for the analyser network for each 'bin': **A**; very low 'bin', **B**; low 'bin', **C**; normal 'bin', **D**; high 'bin, **E**; very high 'bin'.

Figure F6.2 (A-E): Bar charts of single ILC results for the leukocyte measurands for the analyser network. **A**; very low 'bin', **B**; low 'bin', **C**; normal 'bin', **D**; high 'bin, **E**; very high 'bin'.

Figure F6.3 (A-E): Bar charts of single ILC results for the platelet measurands for the analyser network. **A**; very low 'bin', **B**; low 'bin', **C**; normal 'bin', **D**; high 'bin, **E**; very high 'bin'.

Figure F6.4: An image of the reference analyser EQA spreadsheet results prior to routine service.

Figure F6.5: An image of the reference analyser EQA (RIQAS) spreadsheet results following the service.

Figure F6. 5 An image of the reference analyser EQA (RIQAS) spreadsheet results following the service.

Chapter 4 Figures

Figure F4. 1: An example of data calculations using 20 data points for HGB Day1-Day2 showing D1 (day1) measurements, D2 (day2) measurements, Diff (difference from Day1 minus Day2), Aver D1,D2 (average taken of Day1 and Day2 measurement) and Diff sq (Difference squared).

	D1	D2	Aver D1,D2	Diff	Diff sq
1	15.8	15.8	15.8	0	0
2	15.3	15.3	15.3	0	0
3	11.7	11.7	11.7	0	0
4	15.4	15.6	15.5	-0.2	0.04
5	16.3	16.4	16.35	-0.1	0.01
6	13.8	13.7	13.75	0.1	0.01
7	14	14	14	0	0
8	14.9	15	14.95	-0.1	0.01
9	11.6	11.7	11.65	-0.1	0.01
10	15.2	15.2	15.2	0	0
11	18.4	18.4	18.4	0	0
12	17.6	17.7	17.65	-0.1	0.01
13	15.3	15.4	15.35	-0.1	0.01
14	14.3	14	14.15	0.3	0.09
15	15.3	15.5	15.4	-0.2	0.04
16	11.7	11.8	11.75	-0.1	0.01
17	16.3	16.6	16.45	-0.3	0.09
18	11.4	11.4	11.4	0	0
19	15.4	15.5	15.45	-0.1	0.01
20	14.9	14.9	14.9	0	0
averages	14.73	14.78	14.755	-0.05	0.017

Chapter 6 Figures

Figure F6. 1 (A- E): Bar charts of single ILC results for the erythroid measurands for the analyser network for each ‘bin’.

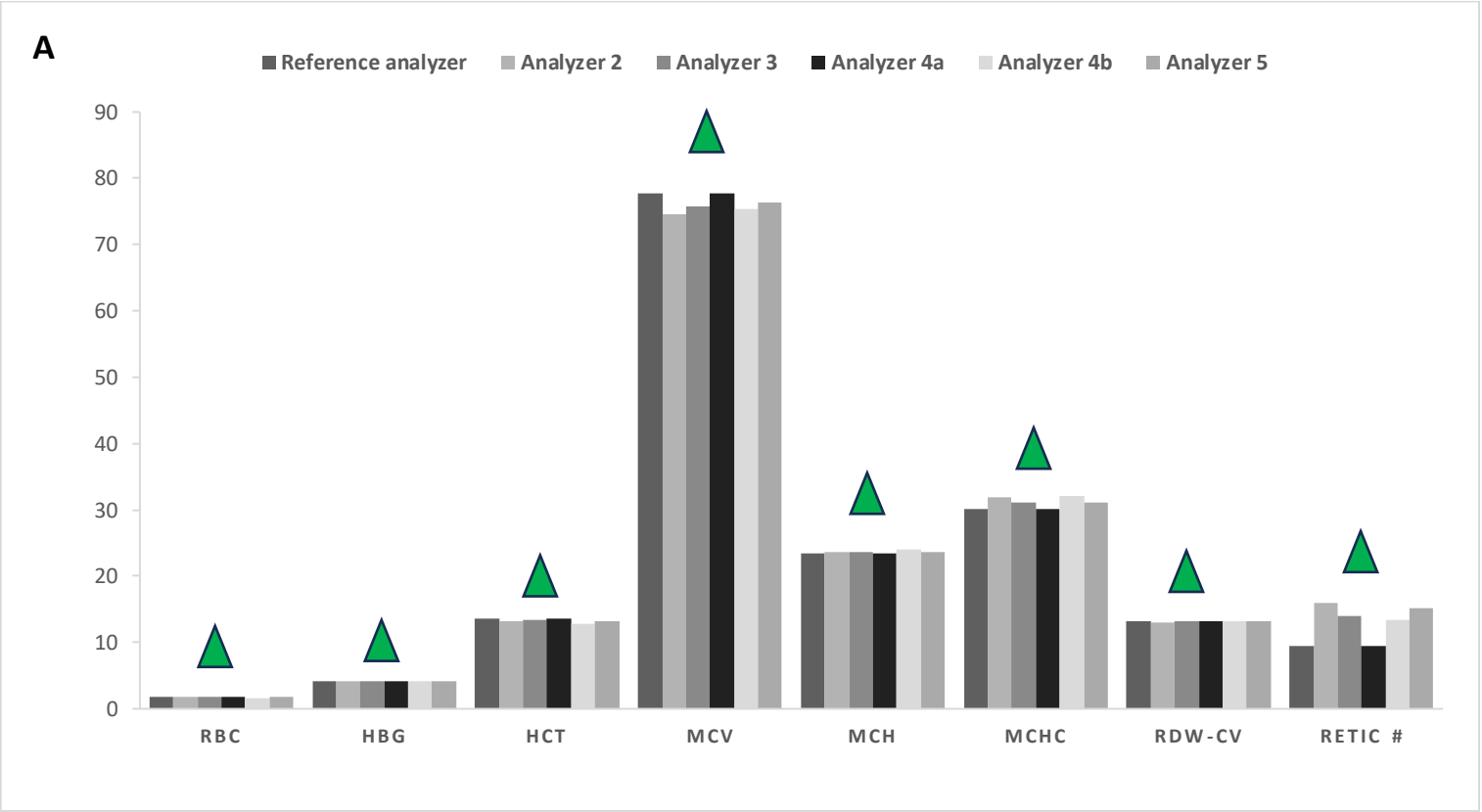


Figure F6.1 A; very low 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

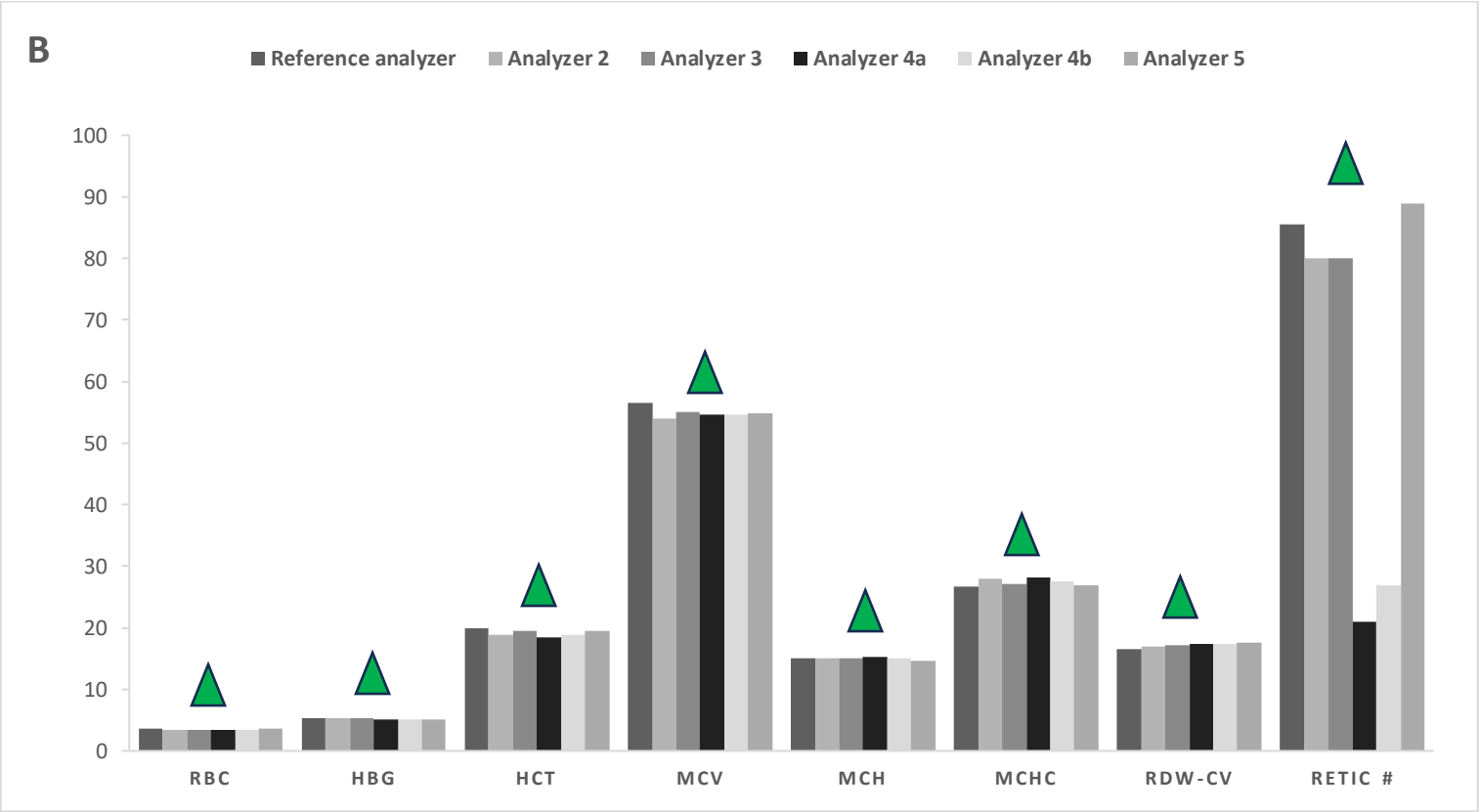


Figure F6.1 B; low 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

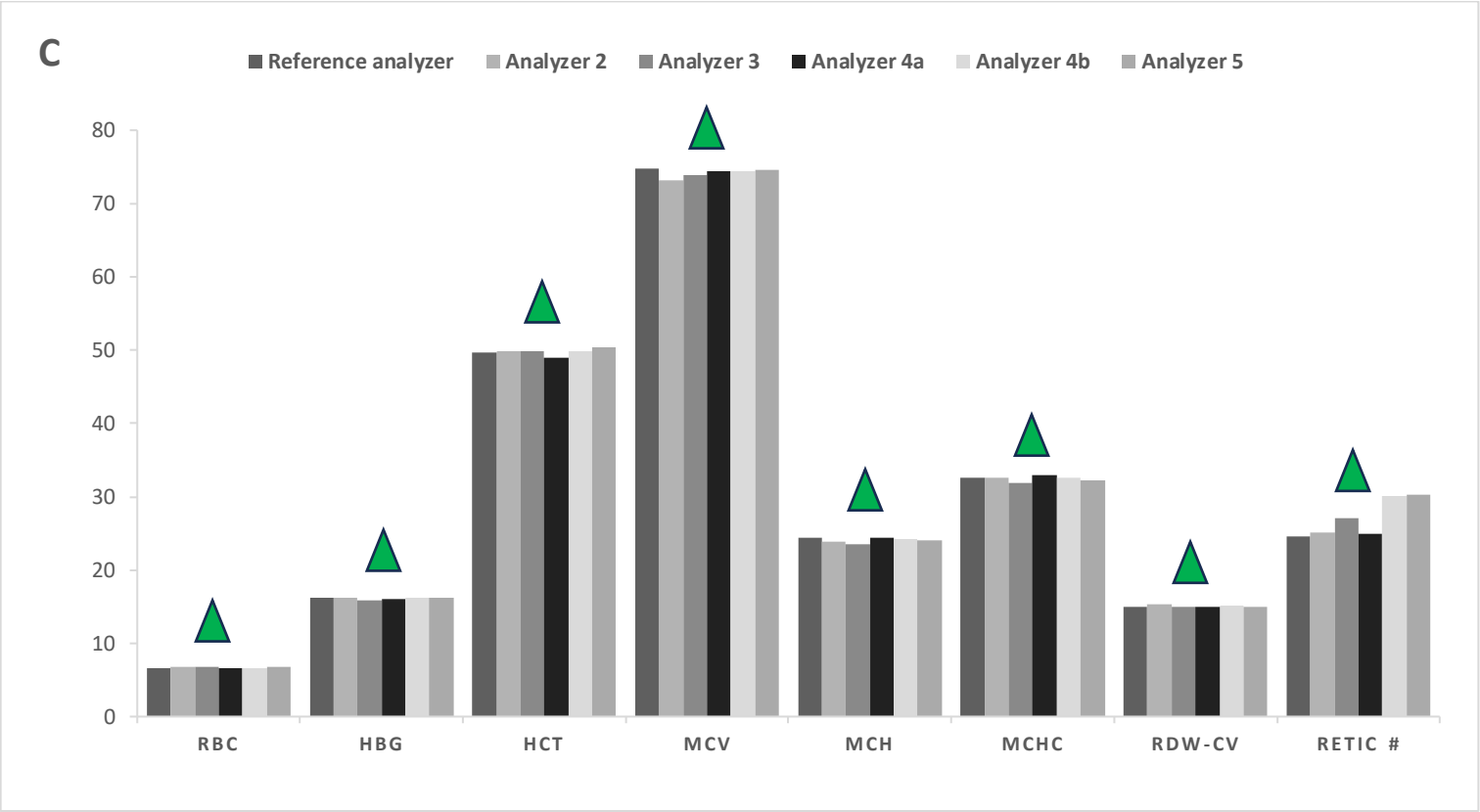


Figure F6.1 C; normal 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

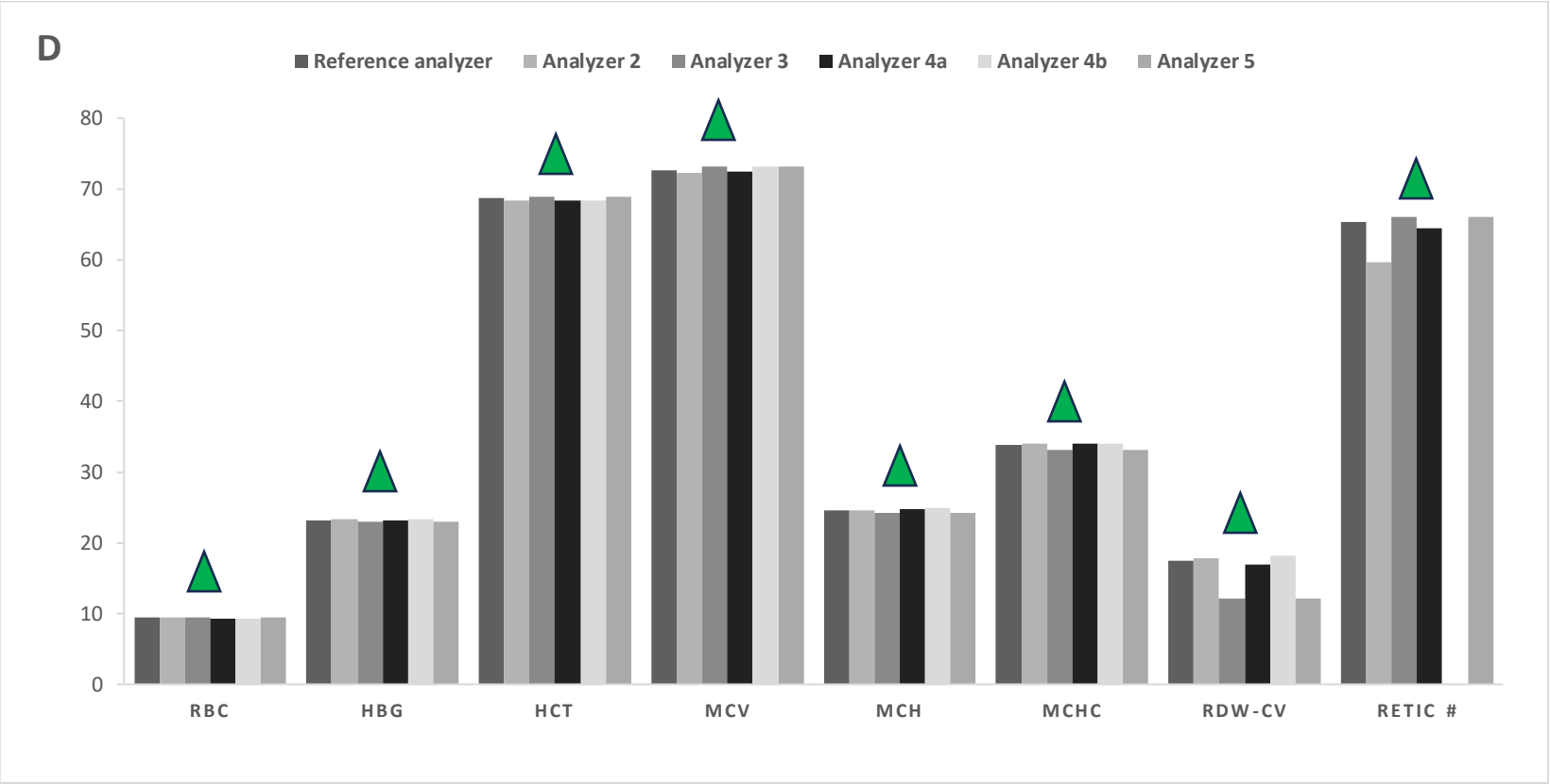


Figure F 6.1 D; high 'bin. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event. **Note:** Analyser 4b did not submit a result for RETIC#.

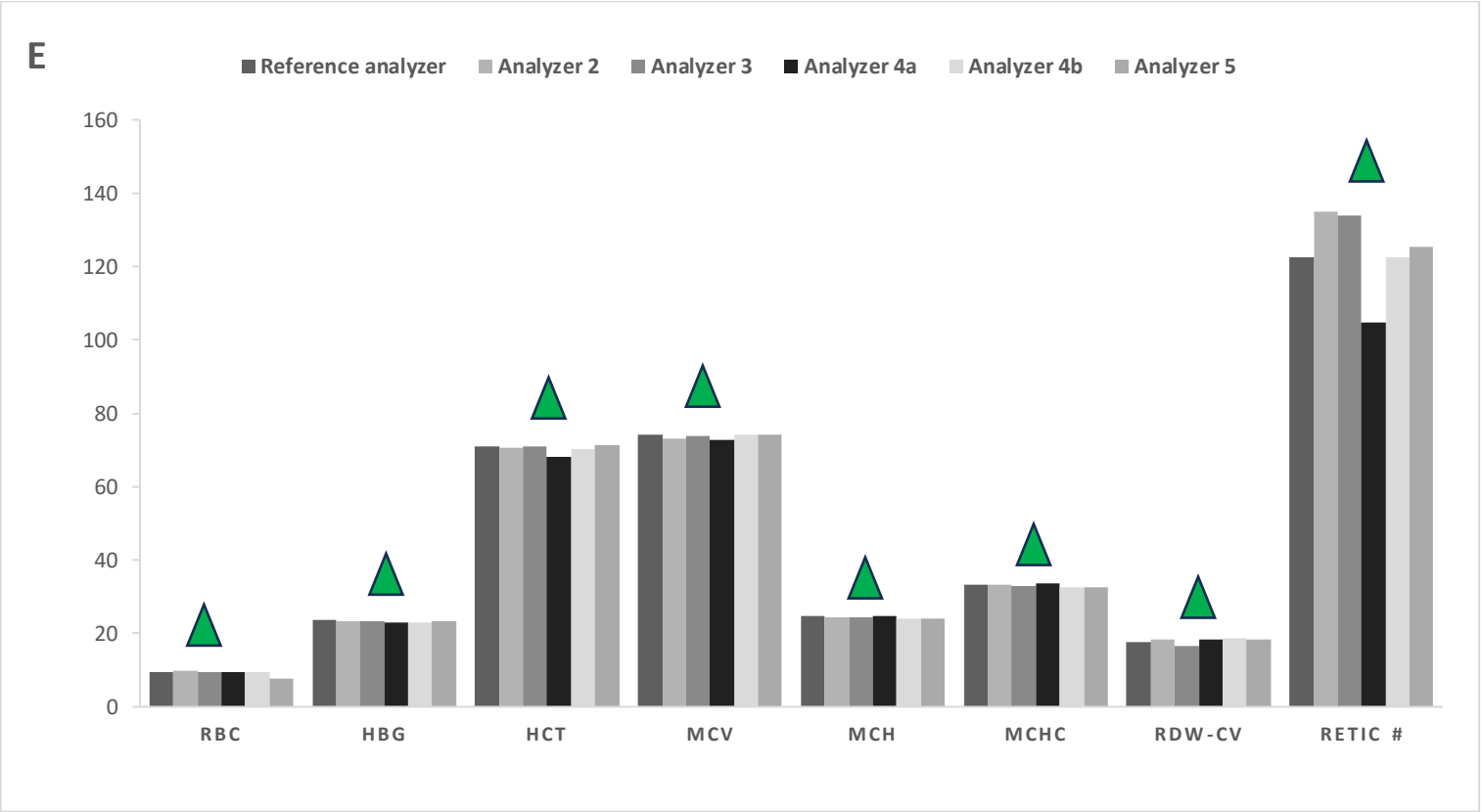


Figure F 6.1 **E**; very high 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

Figure F6. 2 (A-E): Bar charts of single ILC results for the leukocyte measurands for the analyser network.

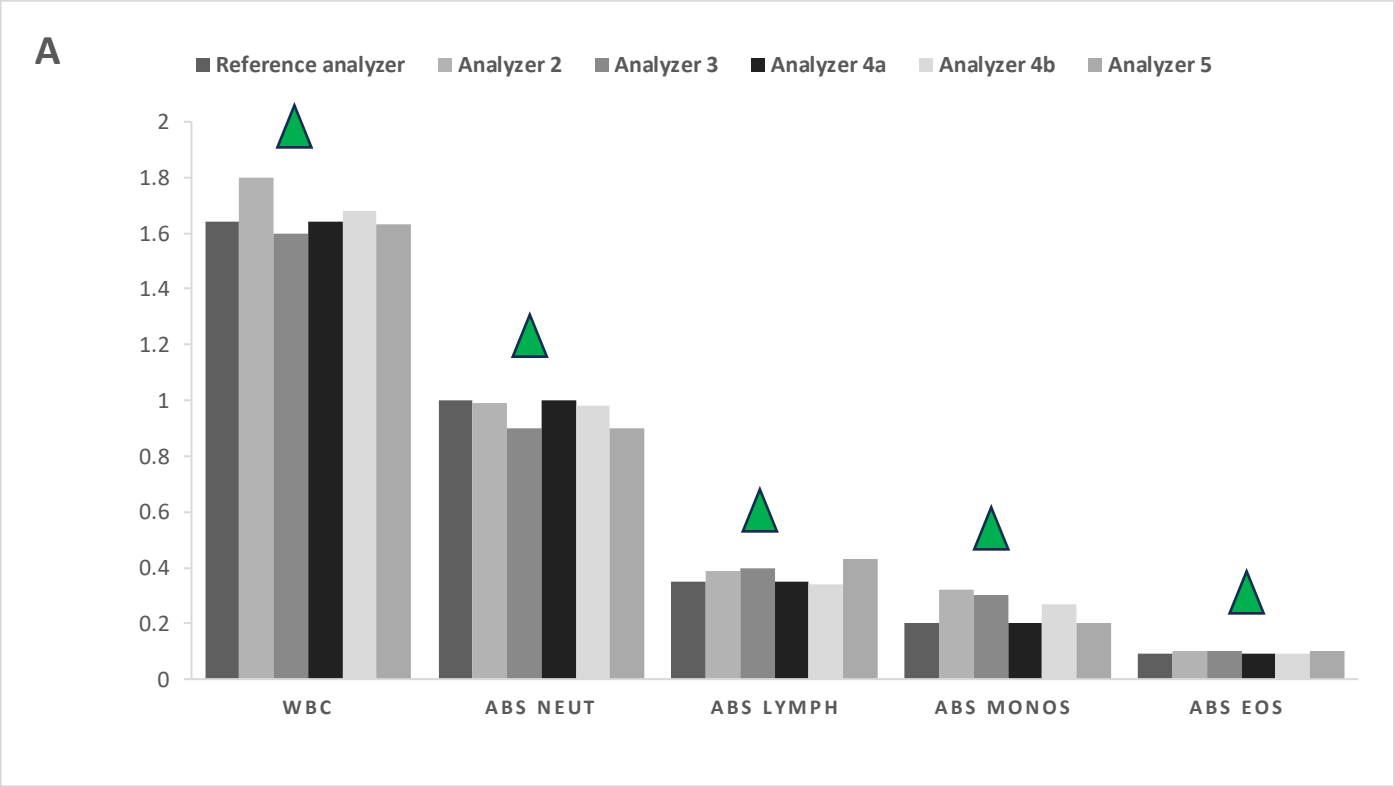


Figure F6.2 A; very low 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

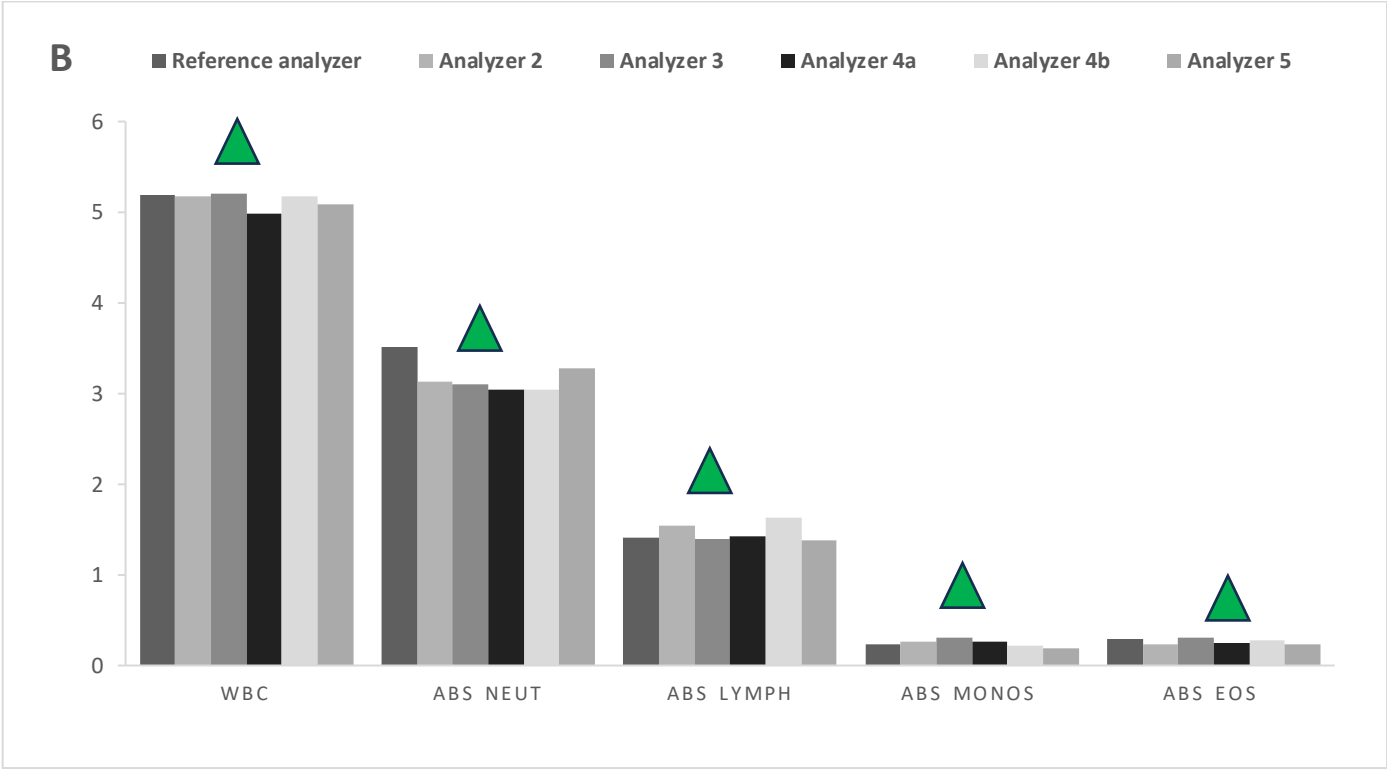


Figure F6.2 B; low 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

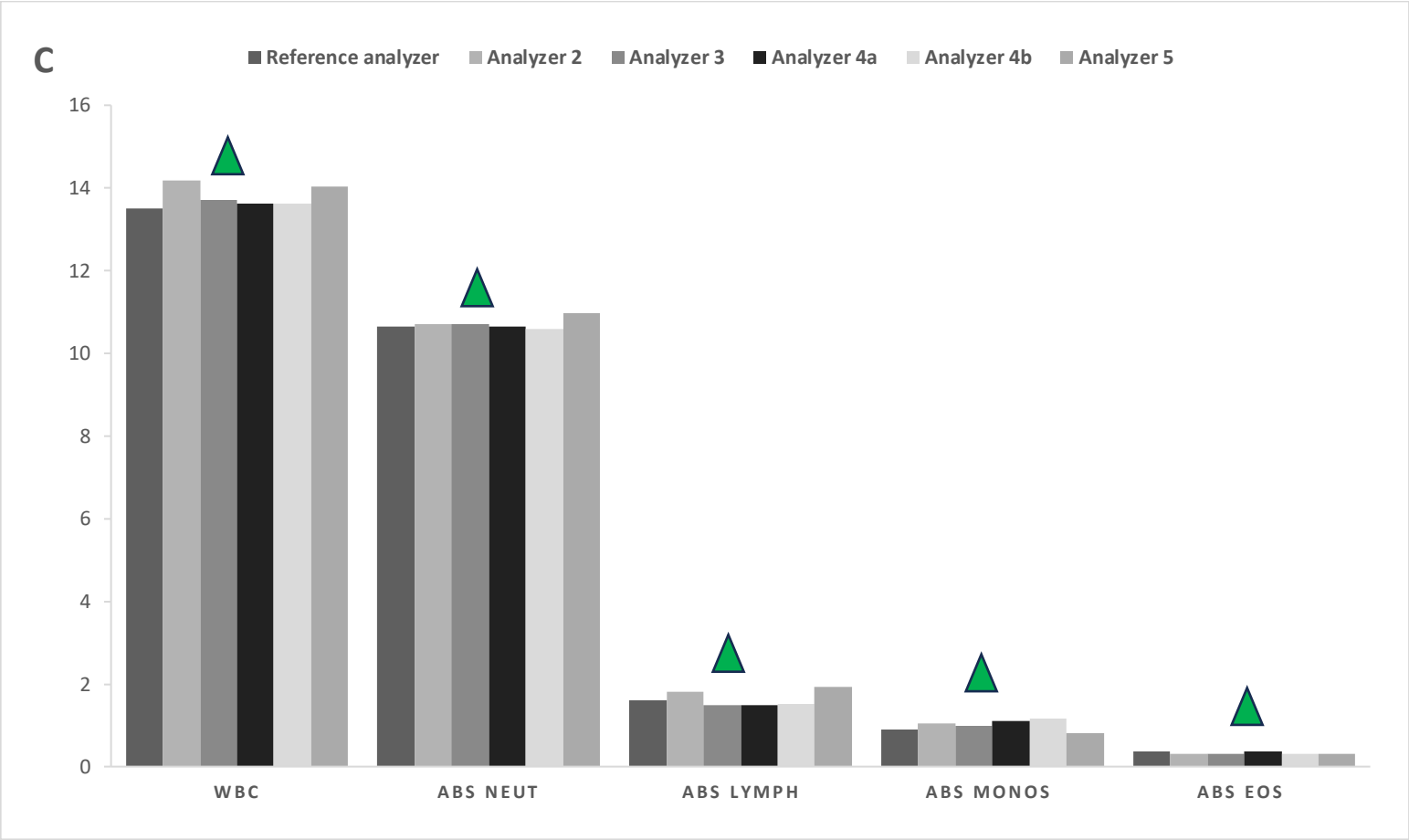


Figure F6.2 C; normal 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

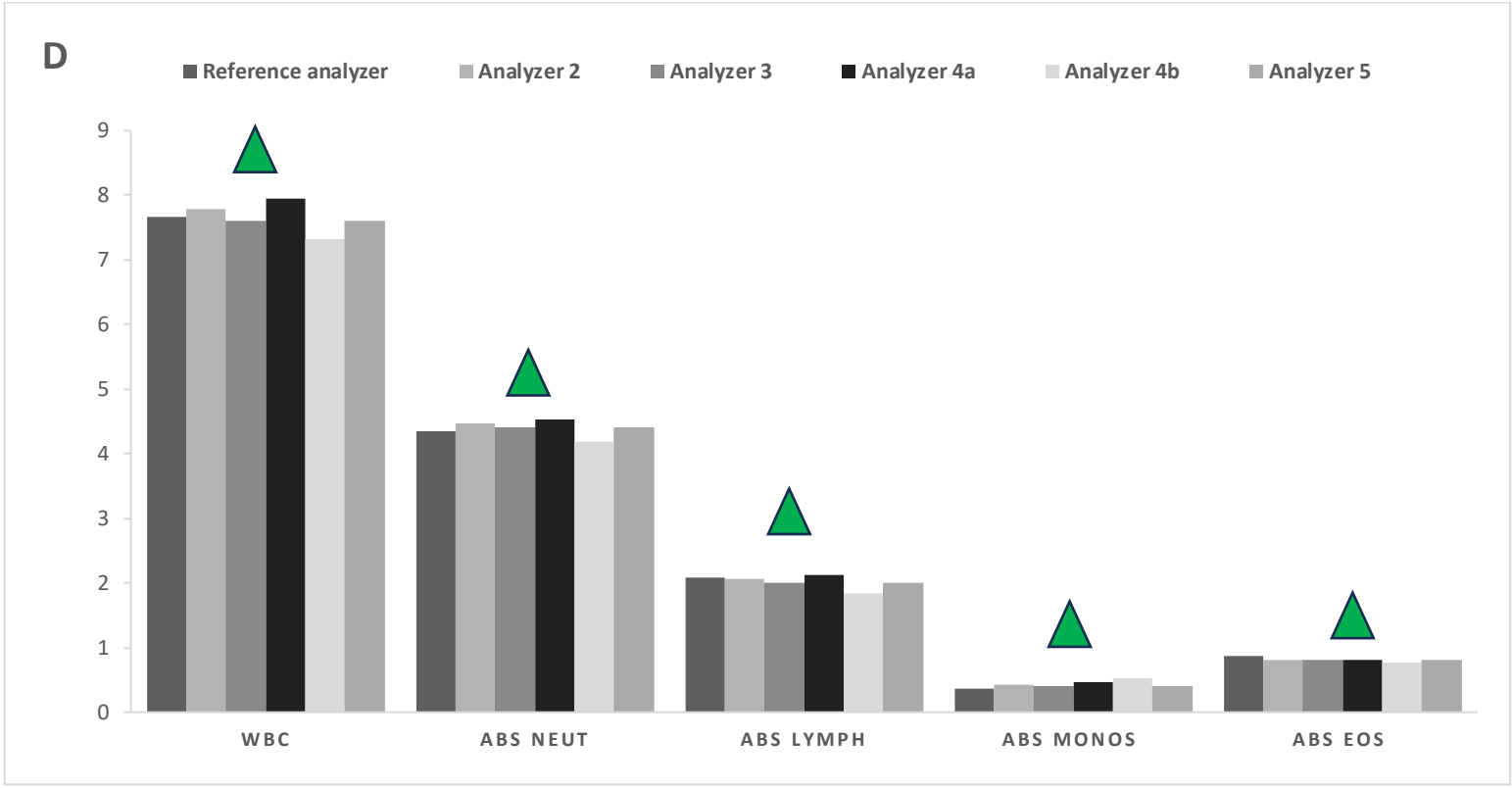


Figure F6.2 D; high 'bin'. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

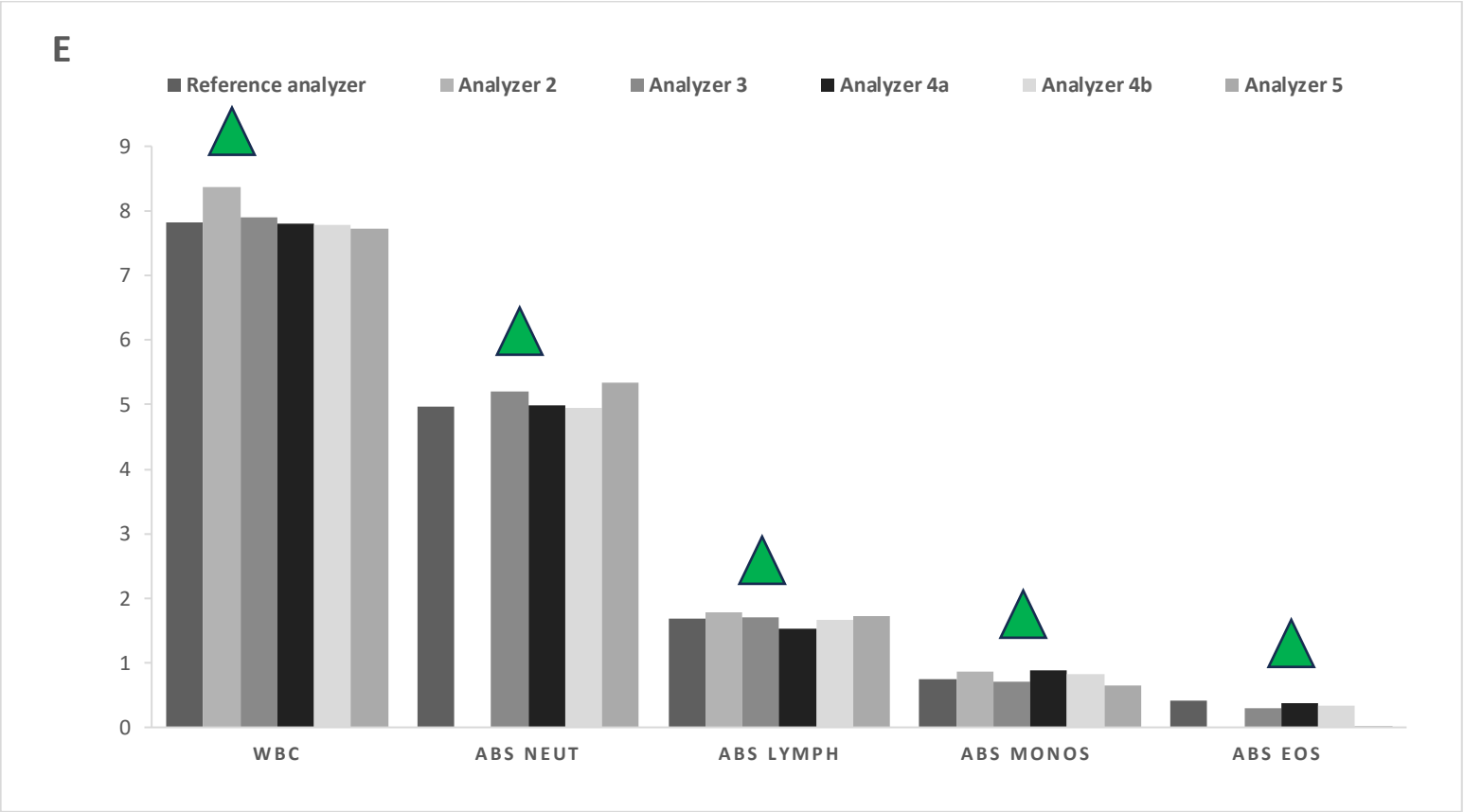


Figure F6.2 E; very high 'bin' analysis events. The X-axis shows the measurands and Y-axis represents the concentration of the measurands. Each bar is colour coded to represent each analyser result (x6). The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event. **Note:** Analyzer 2 did not submit absolute neutrophils and absolute eosinophils.

Figure F6. 3 (A-E): Bar charts of single ILC results for the platelet measurands for the analyser network.

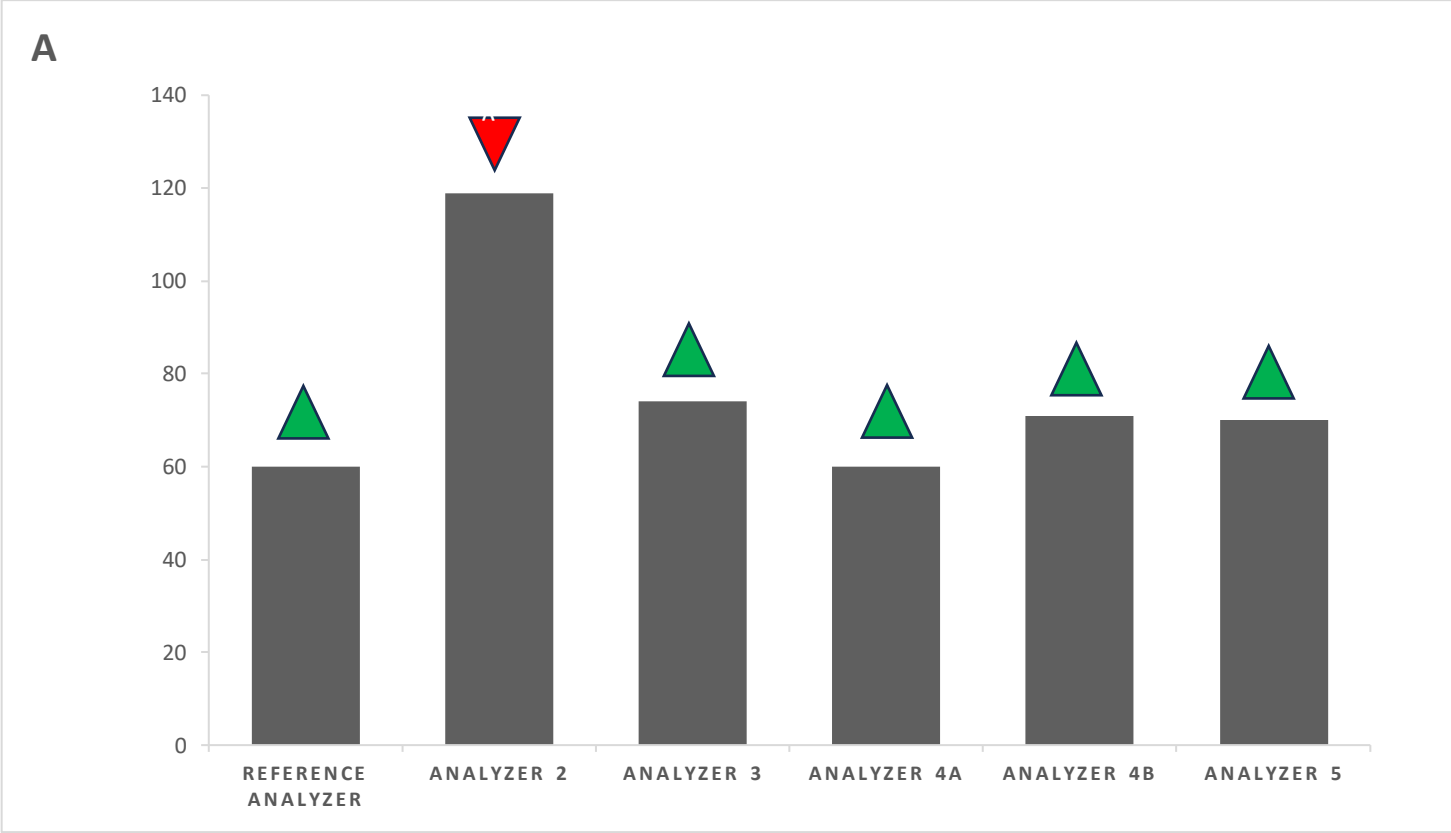


Figure F6.3 A; very low 'bin'. The X-axis represents the individual analyser and Y-axis represents the concentration of the measurands. The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

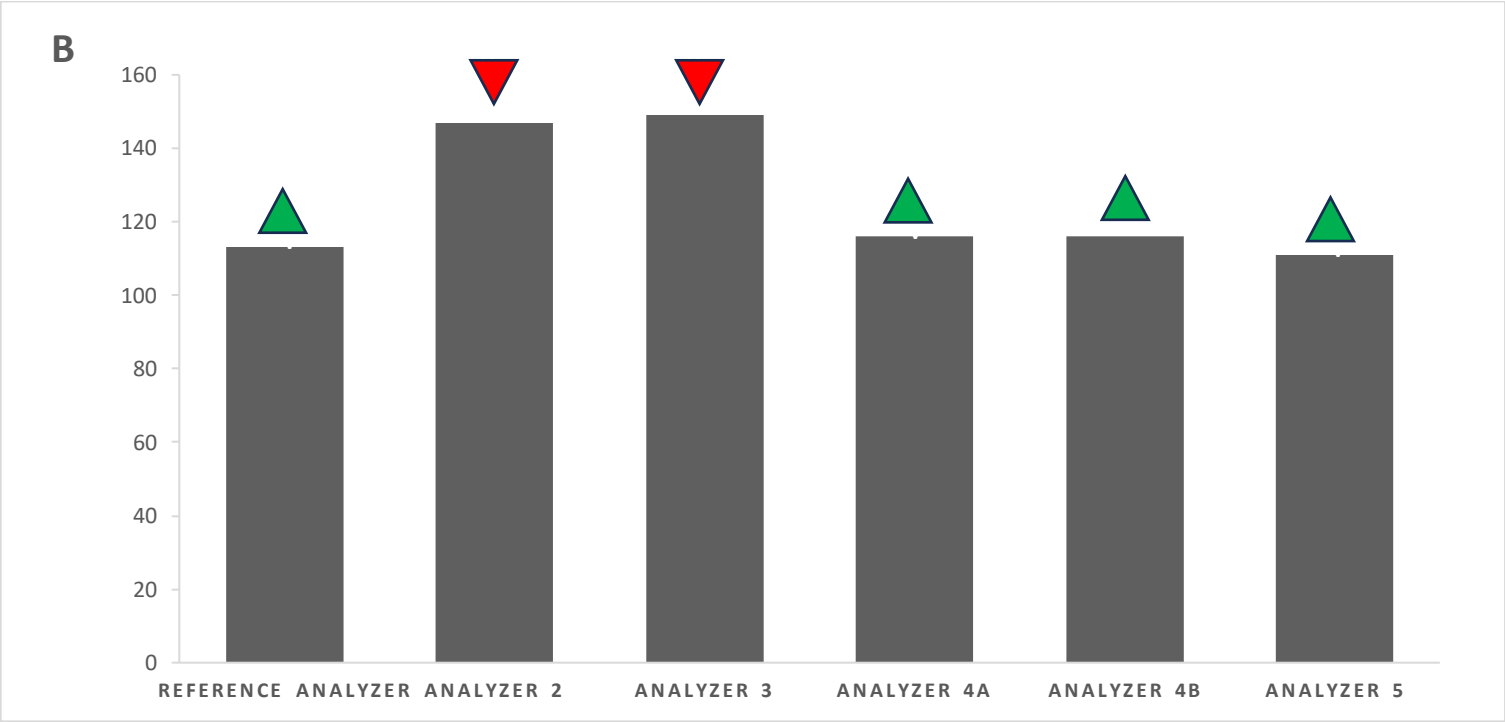


Figure F6.3 B; low 'bin'. The X-axis represents the individual analyser and Y-axis represents the concentration of the measurands. The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

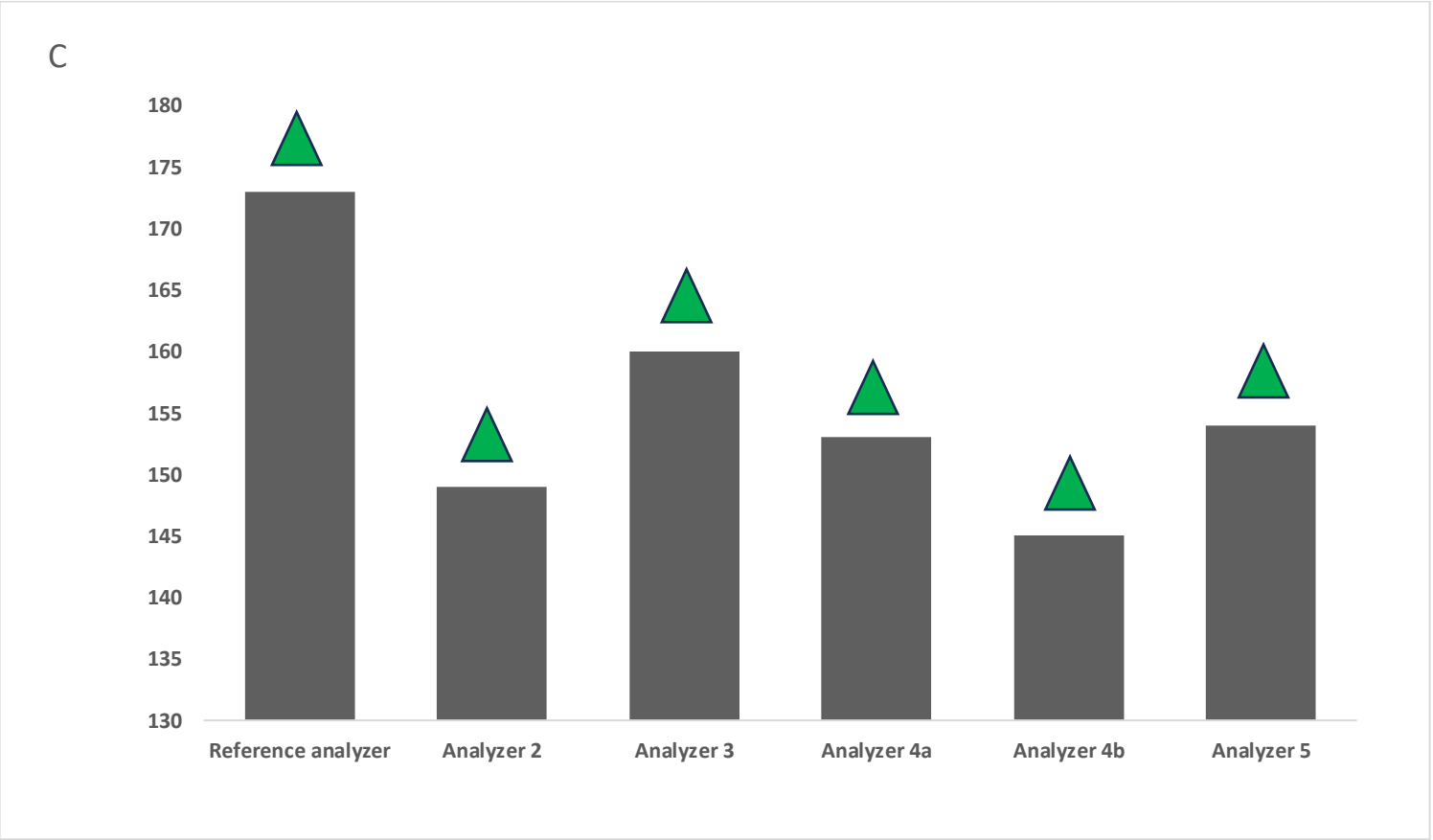


Figure F6.3 C; normal ‘bin’. The X-axis represents the individual analyser and Y-axis represents the concentration of the measurands. The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

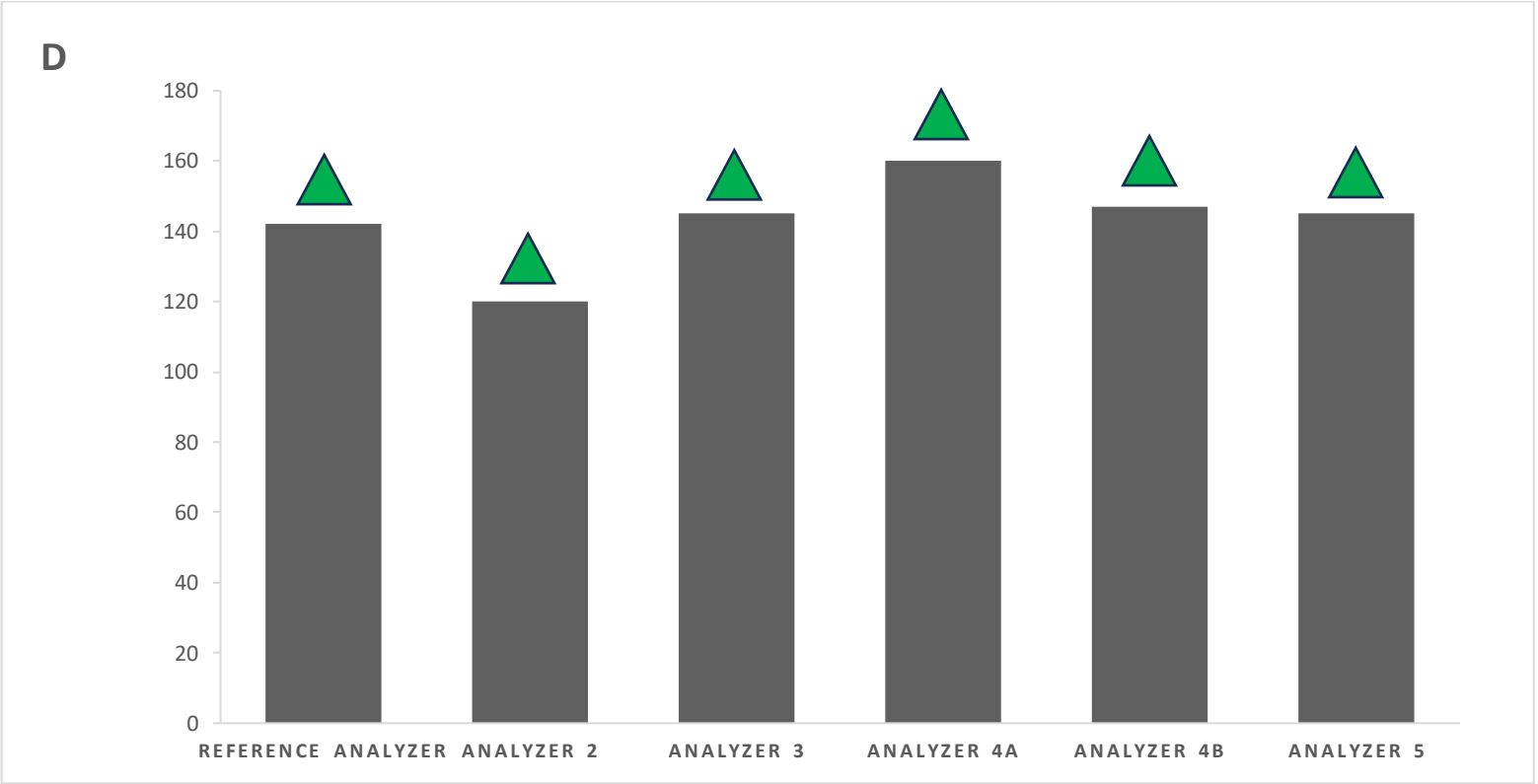


Figure F6.3 D; high ‘bin’. The X-axis represents the individual analyser and Y-axis represents the concentration of the measurands. The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

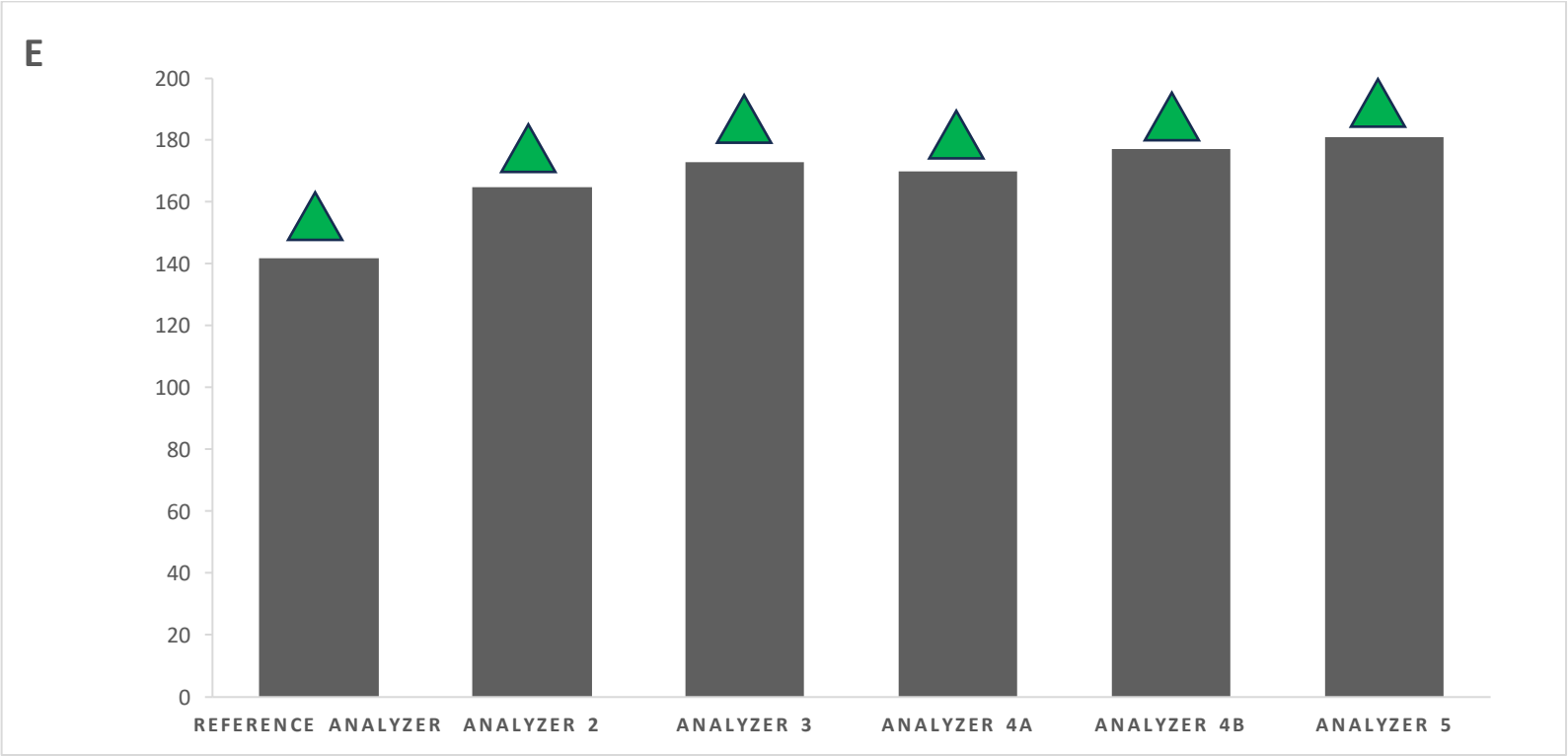
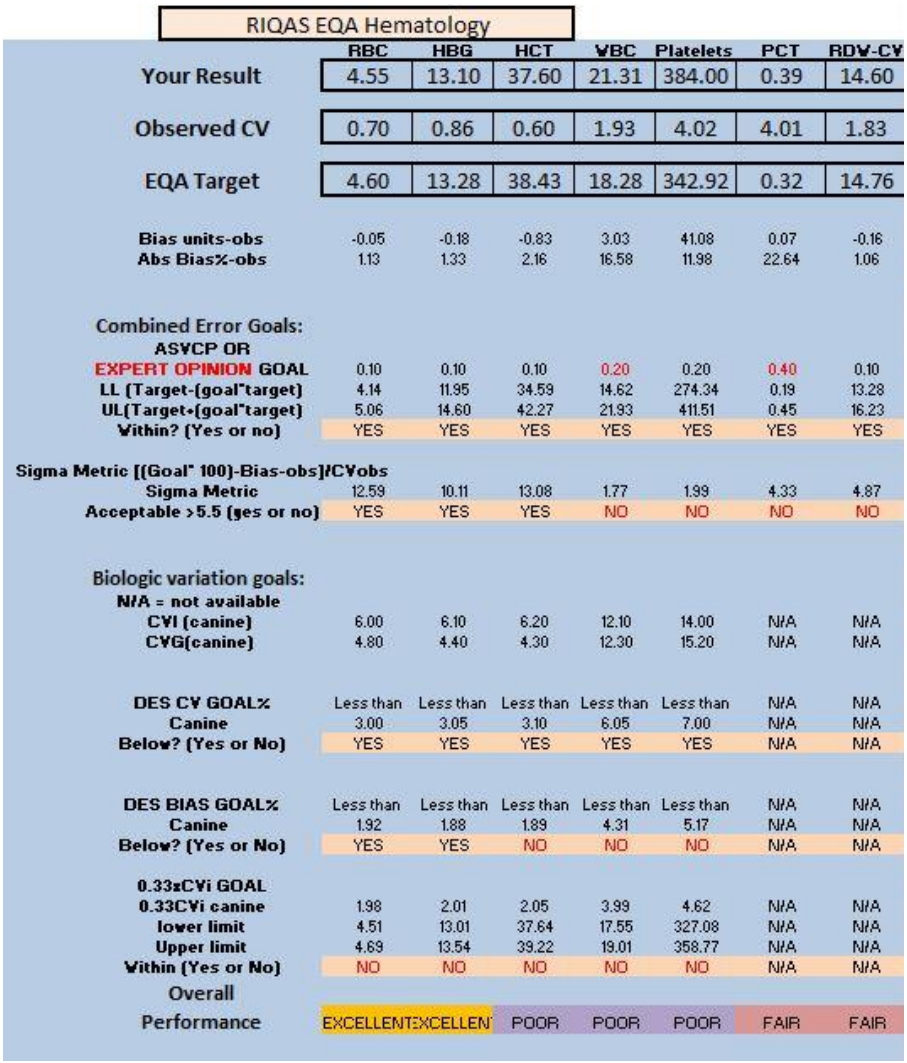


Figure F6.3 E; very high 'bin'. The X-axis represents the individual analyser and Y-axis represents the concentration of the measurands. The green triangle indicates harmonisation was maintained, whereas a red triangle indicates harmonisation was not maintained for that event.

Figure F6. 4: An image of the reference analyser EQA spreadsheet results prior to routine service.

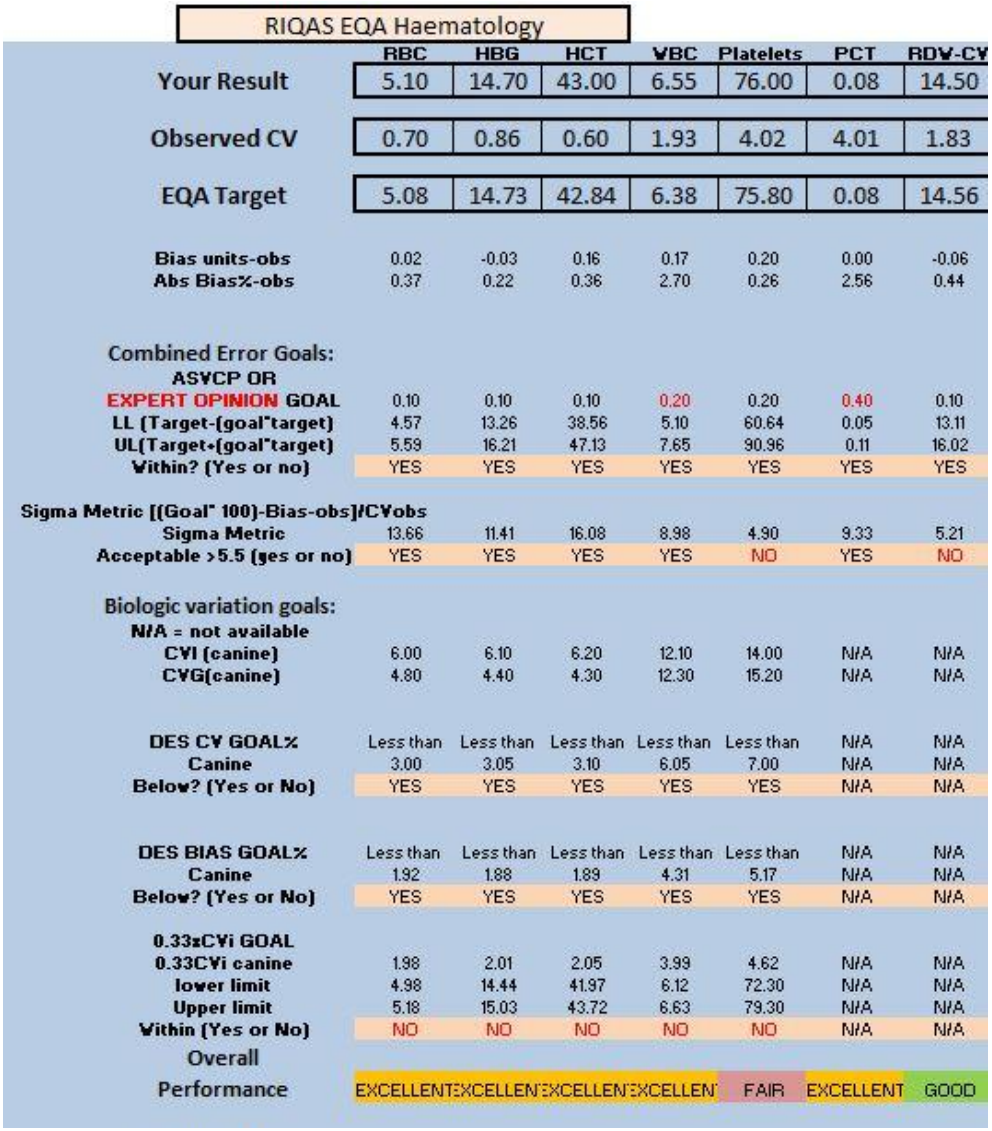


RIQAS EQA Hematology

	RBC	HBG	HCT	WBC	Platelets	PCT	RDW-CV
Your Result	4.55	13.10	37.60	21.31	384.00	0.39	14.60
Observed CV	0.70	0.86	0.60	1.93	4.02	4.01	1.83
EQA Target	4.60	13.28	38.43	18.28	342.92	0.32	14.76
Bias units-obs	-0.05	-0.18	-0.83	3.03	41.08	0.07	-0.16
Abs Bias%-obs	1.13	1.33	2.16	16.58	11.98	22.64	1.06
Combined Error Goals: ASVCP OR							
EXPERT OPINION GOAL	0.10	0.10	0.10	0.20	0.20	0.40	0.10
LL (Target-(goal*target)	4.14	11.95	34.59	14.62	274.34	0.19	13.28
UL(Target+(goal*target)	5.06	14.60	42.27	21.93	411.51	0.45	16.23
Within? (Yes or no)	YES	YES	YES	YES	YES	YES	YES
Sigma Metric [(Goal* 100)-Bias-obs]/CVobs							
Sigma Metric	12.59	10.11	13.08	1.77	1.99	4.33	4.87
Acceptable >5.5 (yes or no)	YES	YES	YES	NO	NO	NO	NO
Biologic variation goals: N/A = not available							
CVI (canine)	6.00	6.10	6.20	12.10	14.00	N/A	N/A
CVG(canine)	4.80	4.40	4.30	12.30	15.20	N/A	N/A
DES CV GOAL%	Less than	Less than	Less than	Less than	Less than	N/A	N/A
Canine	3.00	3.05	3.10	6.05	7.00	N/A	N/A
Below? (Yes or No)	YES	YES	YES	YES	YES	N/A	N/A
DES BIAS GOAL%	Less than	Less than	Less than	Less than	Less than	N/A	N/A
Canine	1.92	1.88	1.89	4.31	5.17	N/A	N/A
Below? (Yes or No)	YES	YES	NO	NO	NO	N/A	N/A
0.33xCVi GOAL							
0.33CVi canine	1.98	2.01	2.05	3.99	4.62	N/A	N/A
lower limit	4.51	13.01	37.64	17.55	327.08	N/A	N/A
Upper limit	4.69	13.54	39.22	19.01	358.77	N/A	N/A
Within (Yes or No)	NO	NO	NO	NO	NO	N/A	N/A
Overall Performance	EXCELLENT	EXCELLEN	POOR	POOR	POOR	FAIR	FAIR

Performance is graded “Excellent” when sigma metric > 5.5 and 2/3 of the biologic variation based quality goals are achieved , “Good” performance when sigma metric > 5.5-5.0 and 2/3 of the biologic variation based quality goals are achieved, “Fair” performance when sigma metric > 4-5 and 1/3 of the biologic variation based quality goals are achieved, “Poor” performance when < 4 sigma, with or without achievement of any of the 3 biologic variation-based goals.

Figure F6. 5: An image of the reference analyser EQA (RIQAS) spreadsheet results following the service.



RIQAS EQA Haematology

	RBC	HBG	HCT	WBC	Platelets	PCT	RDW-CV
Your Result	5.10	14.70	43.00	6.55	76.00	0.08	14.50
Observed CV	0.70	0.86	0.60	1.93	4.02	4.01	1.83
EQA Target	5.08	14.73	42.84	6.38	75.80	0.08	14.56
Bias units-obs	0.02	-0.03	0.16	0.17	0.20	0.00	-0.06
Abs Bias%-obs	0.37	0.22	0.36	2.70	0.26	2.56	0.44
Combined Error Goals: ASVCP OR							
EXPERT OPINION GOAL	0.10	0.10	0.10	0.20	0.20	0.40	0.10
LL (Target-(goal*target)	4.57	13.26	38.56	5.10	60.64	0.05	13.11
UL(Target+(goal*target)	5.59	16.21	47.13	7.65	90.96	0.11	16.02
Within? (Yes or no)	YES	YES	YES	YES	YES	YES	YES
Sigma Metric [(Goal* 100)-Bias-obs]/CVobs							
Sigma Metric	13.66	11.41	16.08	8.98	4.90	9.33	5.21
Acceptable >5.5 (yes or no)	YES	YES	YES	YES	NO	YES	NO
Biologic variation goals: N/A = not available							
CVI (canine)	6.00	6.10	6.20	12.10	14.00	N/A	N/A
CVG(canine)	4.80	4.40	4.30	12.30	15.20	N/A	N/A
DES CV GOAL%	Less than	Less than	Less than	Less than	Less than	N/A	N/A
Canine	3.00	3.05	3.10	6.05	7.00	N/A	N/A
Below? (Yes or No)	YES	YES	YES	YES	YES	N/A	N/A
DES BIAS GOAL%	Less than	Less than	Less than	Less than	Less than	N/A	N/A
Canine	1.92	1.88	1.89	4.31	5.17	N/A	N/A
Below? (Yes or No)	YES	YES	YES	YES	YES	N/A	N/A
0.33xCVi GOAL							
0.33CVi canine	1.98	2.01	2.05	3.99	4.62	N/A	N/A
lower limit	4.98	14.44	41.97	6.12	72.30	N/A	N/A
Upper limit	5.18	15.03	43.72	6.63	79.30	N/A	N/A
Within (Yes or No)	NO	NO	NO	NO	NO	N/A	N/A
Overall Performance	EXCELLENT	EXCELLEN	EXCELLEN	EXCELLEN	FAIR	EXCELLENT	GOOD

Note the improvement in bias (units-obs) and abs bias (%) following the service compared to the event in Figure F.8 and corresponding improvement in overall performance. Performance is graded “Excellent” when sigma metric > 5.5 and 2/3 of the biologic variation based quality goals are achieved , “Good” performance when sigma metric > 5.5-5.0 and 2/3 of the biologic variation based quality goals are achieved, “Fair” performance when sigma metric > 4-5 and 1/3 of the biologic variation based quality goals are achieved, “Poor” performance when < 4 sigma, with or without achievement of any of the 3 biologic variation-based goals.