REVISTA BRASILEIRA
DE HEMATOLOGIA
E HEMOTERAPIA

Analysis of manual reticulocyte counts in the clinical laboratories of Ponta-Grossa and Campos Gerais, Brazil

Contagem manual de reticulócitos em laboratórios de análises clínicas de Ponta Grossa e Campos Gerais, PR, Brasil

Mackelly Simionatto[1]
Josiane P. Paula[1]
Aguinaldo J. Nascimento[2]
Maria Suely S. Leonart[3]
Domenic Cicchetti[4]

Reticulocyte counts are widely used in laboratories to evaluate bone marrow erythropoietic activity and have great diagnostic and prognostic importance in the treatment of anemias. Reticulocytes are supravitally stained with new methylene blue or brilliant cresyl blue, which highlight the characteristic aspect of the reticulum visible by light microscopy. Known criteria were observed for good manual reticulocyte counting with special attention being paid to the preparation of reticulocyte slides, to counting fields without overlapping cells, and to the number of evaluated cells. The aim of this study was to evaluate the inter-observer variation and also to analyze the statistical error of manual reticulocyte counting. The Intraclass correlation coefficient (ICC) according to Bartko was used to evaluate the level of agreement among 12 laboratory technicians who evaluated the same 25 blood smears. The results of statistical analyses showed that the amount of random error, defined as (1-r2), varied from 4% to 60% among the technicians. Although imprecision occurred, the overall profiles were quite similar, and intraclass correlation coefficients indicated that the results obtained have clinical significance. Rev. Bras. Hematol. Hemoter.

Key words*: Haematology; statistical methods; manual reticulocytes count.*

Introduction

Reticulocytes are precursory to the erythrocyte cells in the blood. They are anucleated cells, more rounded in shape and 20% greater, in volume, than the erythrocytes. However, when stained with panoptic dyes (Romanowski) they produce polychromatic slides, due to the presence of mature red blood cells with hemoglobin, synthesized during maturation, and reticulocytes with ribonucleic acid residues. These residues are stained with new methylene blue or brilliant cresyl blue dyes which confer the characteristic aspect of reticulum, when observed in optic microscopy. Such staining is not permanent.[1] The reticulocytes can be classified as mature or immature depending on the amount of granules or reticula they contain.[2]

Reticulocyte counting is routinely and widely used in the laboratory to evaluate bone marrow erythropoietic activity. It is of great diagnostic and prognostic value in hemolytic anemias, in acute hemorrhage, in response to iron, folic acid and vitamin B12 therapy,[3] as well as, after chemotherapy or bone marrow transplant4. The manual method of reticulocyte counting, described in 1930, is still very frequently used today due to its low cost in comparison to the automated method that has seen widespread

[1]*Farmacêutico-Bioquímico.*
[2]*Professor sênior da Universidade Federal do Paraná – Curitiba-PR.*
[3]*Professora titular da Universidade Federal do Paraná – Curitiba-PR.*
[4]*Senior Professor, Department of Biometry, Yale University School of Medicine, New Haven - USA.*

Programa de Pós-Graduação em Ciências Farmacêuticas, Universidade Federal do Paraná – Curitiba-PR.

Correspondence: *Aguinaldo José do Nascimento*
Universidade Federal do Paraná
Av. Lothário Meissmer, 632 – Jardim Botânico
802210-170 – Curitiba-PR – Brasil
E-mail: ajnasc@ufpr.br
Doi:

application since 1990. However, manual counting presents some factors that cause great variation.

Peebles *et al*.[4] reported interobserver variation coefficients between 25% and 50%, considered higher than would be expected. Therefore, manual counting must follow standardized criteria of execution and identification in order to produce more precise results. As suggested by the H-44 technique norms of the National Committee of Clinical Laboratory Standards (NCCLS) and of the International Committee for Standardization in Hematology (ICSH), some criteria must be observed for a good performance of reticulocyte manual counting. These would include paying special attention to the differentiation among reticulocyte granules, Pappenheimer, Heinz, and Howell-Jolly bodies, or remaining dyes; taking special care in the preparation of the reticulocyte slides; restricting the counting to fields that do not contain overlapping cells; and also taking into account the number of evaluated cells.[5]

Automated reticulocyte counting is based on the rapidly expanding technology of flow cytometry that makes the determination fast, accurate, and efficient by counting a great number of blood cells; and the methodology also provides information regarding the reticulocytes. Parameters such as amount of hemoglobin and degree of reticulocyte maturity allow more trustworthy diagnostic and therapeutic monitoring, mainly for patients with lower reticulocyte counts.[6] However, this methodology is still expensive, and this limits its routine use in small and medium size laboratories.

Considering the relevance of the quality control of reticulocyte countings for the detection of hematologic illnesses, the aim of this study was to evaluate interobserver variations; to determine the amount of statistical error in manual reticulocyte counting, using known criteria for its determination; and also to demonstrate the limitations of this method.

Material and method

Manual reticulocyte counting

Manual reticulocyte counting was carried out in accordance with Lewis *et al.*[7] A new methylene blue or brilliant cresyl blue solution (100µl) was added to 200 - 250µl of blood, and incubated at 37ºC for 20-25 min. The blood volume to be added to the dye depends on the volume of packed red blood cells volume (PCV), 200µl for PCV equal or superior to 30% and 250µl for less than 30% PVC. After the incubation with the dye, blood smears on glass slides were prepared for reticulocyte counting, for up to 72 hours.

Criteria for adequate blood smear counts

Blood smear was observed under optic microscope objective immersion (100x). Both mature erythrocytes and well defined reticulocytes could then be observed. The mature reticulocyte possesses few RNA filaments or granules, while the youngest possess a great amount of granules. Erythrocytes that present at least two small corpuscles or stained filaments are still considered to be reticulocytes.

Adequate supravital staining should not be very weak at the point of not seeing the reticule, nor very strong, at the point of becoming blurred or with thick reticule. The area to be analyzed cannot have cells overlapping, or very fine fields, so that very few cells are observed that show empty spaces between the cells.

According to Lewis *et al.*,[7] when the count is less than 10% a convenient method is to survey successive fields until at least 100 reticulocytes have been counted; and to count total cells in at least ten fields in order to obtain an average number of cells per field. So, when the percentage of reticulocytes decreases, the examiner must count more cells to obtain an acceptable error. When the reticulocyte number exceeds 10%, the examination of a small number of cells will be enough to get a standard error of less than 10%. It is consensual that there must be about 200 erythrocytes in each field so that 20-50 fields are counted with less than 1% reticulocytes.

Blood smear screening

A total of 25 blood smears was screened by professional biochemists in 10 laboratories of Clinical Analyses in Ponta-Grossa and Campos Gerais, PR, Brazil. The technologists were blinded and identified only as A to J. In a first approach, the samples were carefully examined by technologists K and L. Each laboratory received the slides numbered from 1 to 25 and they were asked to screen from 2 to 3 times a week, during a 5 week period, between June 25th and July 30th 2008, and within 48 hours after receiving the slides.

Statistical analyses

The Intraclass correlation coefficient (ICC), due to Bartko[8,9] was used to evaluate the level of agreement or concordance among technologists for parametric measurements, for more than 2 raters.[10] Analyses were performed with the use of the on-line statistical package of Hong Kong University[11] at the following website: http://department.obg.cuhk.edu.hk/researchsupport/IntraClass_correlation.asp

The data were also analyzed by Pearson correlation, one way ANOVA and with the Levey-Jennings quality control chart.[12] The criterion for statistical significance was set at the nominal probability (p) level of $p \leq 0.05$.

Results and discussions

In order to observe if the two technologists, K and L, were calibrated, by the established criteria, their data were analyzed by Pearson correlation, as shown in Figure 1. The systematic error reflected by the slope value close to unity shows that both technologists produced quite concordant

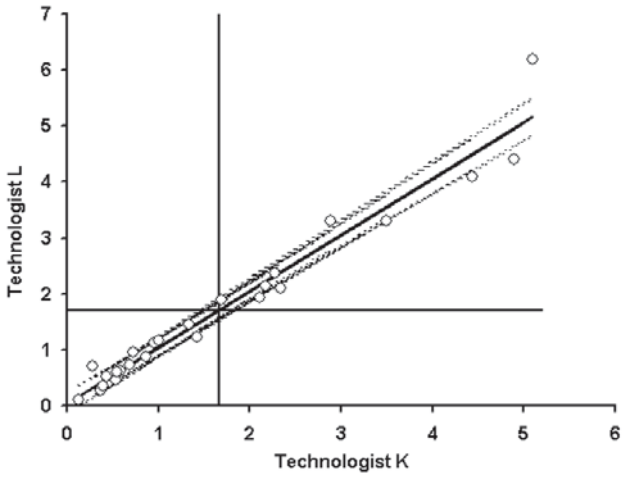

Figure 1. n = 25 blood smears; the Pearson correlation r was 0.96. Solid line - estimated regression line (y = 1.018x); dotted line - 95% limit of confidence for residues. Lines in cross - averages values for x and y

Table 1. Intraclass correlations (ICCs) for technologists "K", "L" and the others for reticulocyte counts (%)

Technologists	r
KL	0.9796
K L G	0.9365
K L B	0.9289
K L C	0.9130
K L F	0.9056
K L D	0.8810
K L J	0.8415
K L H	0.8134
K L A	0.8010
K L E	0.7610
K L I	0.6395

r - Intraclass correlations (ICCs)

values. The random error, reflected by the variability of counts around the regression line shows very little imprecision between them. The "r" value close to unity shows a high degree of correlation and concordance between the two technologists, as evidenced by a random error ($1-r^2$) of only 4.1%. The plot of the 25 blood smears, arranged form lowest to highest values, illustrates the reticulocyte counts of the samples.

The statistical analyses of agreement for intra-class correlation for parametric data, with more than two examiners,[8] using the "on-line" statistical package of the University of Hong Kong,[11] is illustrated in Table I. Highest concordances were obtained between technologists K and L. They were responsible for the blood smears mounting, in accordance with the criteria established by Lewis *et al*.[7]

According to the criteria of Cicchetti and Sparrow,[13] Intraclass Correlation Coefficients (ICCs) can be classified for levels of clinical significance as follows: poor, < 0.40; fair, 0.40 to 0.59; good, 0.60 to 0.74; excellent, 0.75 to 1.00.

The analyses performed on the total data (Table I) show that, except for technologist I, all the technologists can be classified as having excellent ICCs .

With the data fractionated into low, medium and high counts, the sample sizes became small, causing some difficulty in interpreting the full meaning of the correlational analyses. However, the technologists do show certain difficulties in agreement for low counts (not shown).

Considering all the data, the amount of random error ($1-r^2$) was observed as follows: <20% among technologists G, B, C and F; 10-40% among technologists D, J, H and A; >40% for technologists E and I. Technologist I with K and L showed the lowest ICC value (r = 0.64) (Table I).

Using the one-way ANOVA, we compared the mean values of counts for technologists K and L against each

Table 2. One-way ANOVA for data of technologists "K", "L" and the others for reticulocyte counts (%)

Technologists	SS	DF	MS	F	P
KLF	0.0284	2	0.0142	0.0064	0.9936
KLD	0.1655	2	0.0828	0.0320	0.9686
KLB	0.3802	2	0.1901	0.0924	0.9118
KLJ	0.4140	2	0.2070	0.1009	0.9041
KLH	1.7507	2	0.8753	0.3046	0.7384
KLC	2.4932	2	1.2466	0.5794	0.5629
KLE	2.7364	2	1.3682	0.7018	0.4990
KLG	4.1180	2	2.0590	0.8659	0.4251
KLI	3.5230	2	1.7615	1.1371	0.3264
KLA	4.6645	2	2.3323	1.2722	0.2866
Total	29.6283	11	2.6935	1.3156	0.2150

SS - sum of square; DF- degrees of freedom; MS - medium square; F - Test F for variance comparisons; p - probability error (error I in statistical test of null hypothesis)

rater. The result is illustrated in Table II. No significant statistical differences were observed for the mean reticulocyte counts ($p > 0.05$). The mean values obtained for each rater are illustrated in Table III where it can be observed that technologist G showed the highest super-estimation in the reticulocyte countings, while A showed the lowest under-estimation.

The Levey-Jennings quality control plot[12] (Figure 2) for averages shows more clearly that technologists B, D, F, J, K and L had similar results; A, C and I tended to sub-estimate the counts; and E, G and H tended to super-estimate the reticulocyte counts. Figure 2 also shows the quality control chart for amplitude and the high variability of the 25 blood

Table 3. Descriptive statistic for the technologists data of the reticulocyte counts (%)

Technologists	N	Mean	Std.Dev.	Std.Err	-95,00%	+95,00%
G	23	2.200	1.636	0.341	1.493	2.907
E	25	2.094	1.169	0.234	1.611	2.576
H	25	2.012	2.035	0.407	1.172	2.852
J	25	1.843	1.293	0.259	1.309	2.376
B	25	1.836	1.299	0.260	1.300	2.372
L	25	1.714	1.516	0.303	1.088	2.340
F	25	1.681	1.463	0.293	1.077	2.285
K	25	1.667	1.477	0.295	1.057	2.277
D	25	1.599	1.814	0.363	0.851	2.348
C	24	1.300	1.402	0.286	0.709	1.892
I	25	1.232	0.406	0.081	1.065	1.400
A	23	1.148	0.971	0.203	0.728	1.568

Std.Dev.- standard deviation; Std.Err - standard error;
-95.00%, +95.00% - limits of confidence

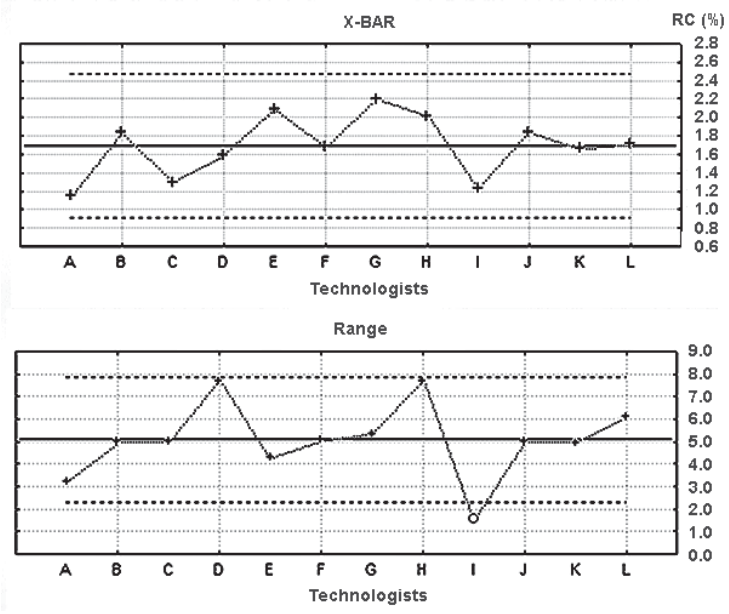

Figure 2. Quality control chart for average and range (Shainin and Shainin, 1993); Solid line - grand mean; dotted line - dispersion limits

smear counts, especially the data of technologist I that was greatly sub estimated ($p < 0.05$). All this said, technologist I did not perform within the 48 hour time span, as requested. These results are quite useful to signal the fact that the blood smears in glass slides are not stable beyond the 48 hour span. The stated period for counts should be observed; otherwise it will likely result in sub estimated counts.

Interesting results were obtained with the scatterplot of the 300 data points distributed across 25 blood smears (Figure 3). Although some numerical discordance under statistical analysis was observed, it can be seen that, as far as the high, low and average counts are concerned, the profile among the technologists is quite similar, and the results obtained are therefore clinically useful. Since great discrepancies are observed for reference values due to variability in sex, age and different pathologies, the interobserver random errors should be around 20%. Applying this criterion, the interobserver random errors of the present work was higher than expected.

Normally erythrocytes do not contain inclusions. However many different inclusions may be seen in various hematological disorders, and in a supravital preparation might be confounded with a reticulocyte:[7,14,15]

1) Pappenheimer bodies are in general a sole granule, staining darker blue then reticulocytes granule. Pappenheimer bodies are secondary lysossomes, variable in their composition of iron and protein, or mitochondria with iron micelles that appears as small irregular basophilic deposits at cell periphery. The protein matrix of the granule is staining with Romanowsky stains and the iron portion of the granule is staining with Prussian blue;

2) Heinz bodies can be visualized with supravital stains as clearly blue. They are discrete inclusion in contrast to the reticular filament material in a reticulocyte. They are composed of aggregated, denatured hemoglobin and appear as masses just under or attached to the cell membrane. When appearing singly, a Heinz body is large, but when several are present in one cell, they are small;

3) Howell-Jolly bodies are dark purple or violet spherical granules in the erythrocyte, usually occurring singly, rarely more than two per cell. The granules are nuclear DNA fragments;

4) Basophilic stipplings are bluish-black granules distributed intracellular, composed of aggregated ribosomes (RNA), and they are sometimes associated with mitochondria and siderosomes;

5) Protozoan inclusions as *Malaria* and *Babesia* present blue ring with red dot in erythrocyte and others evolution shapes (trophozoite stages) outside cells; vi) Erythrocyte artifacts that occurs when cells are mechanically traumatized, or remaining dyes, can also confound the identification of reticulocyte.

The results obtained in this work suggest that the experienced technologists involved in this work have their own criteria, and they are not calibrated as proposed by Lewis *et al*.[7]

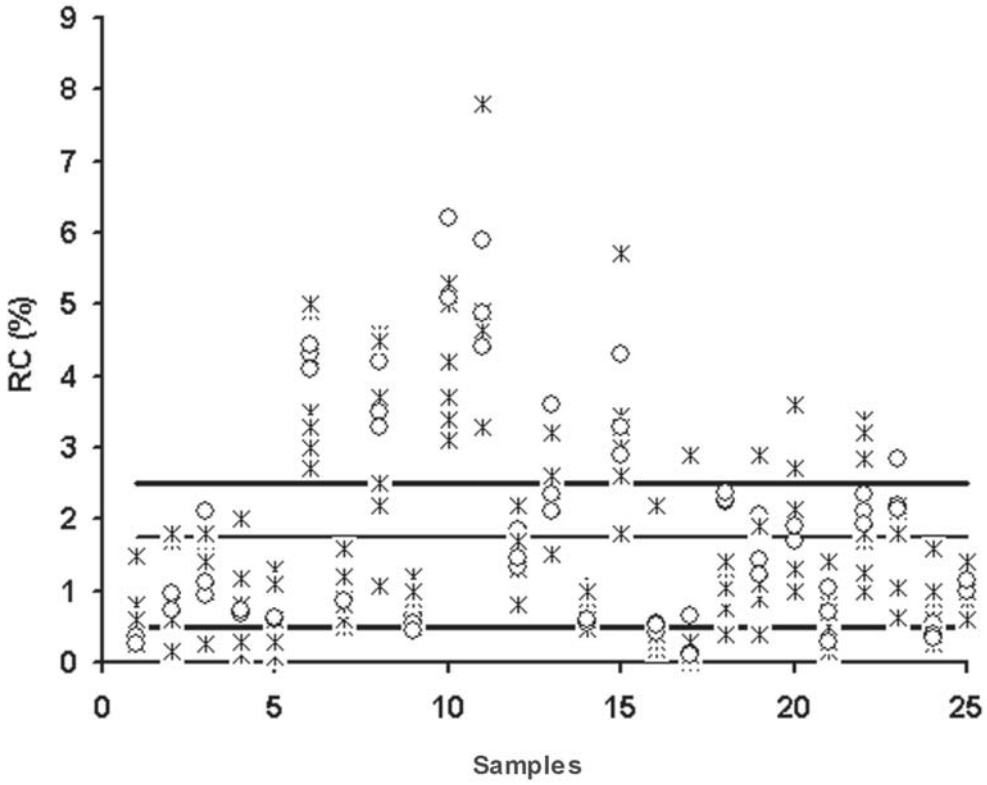

Figure 3. o - Technologists K, L and G; * - Technologists A-J; thick line - limit normal values between 0.5 to 2.5 % RC; thin line - grand average (1.7%). The data of technologist I was not included in this plot

Acknowledgments

The authors wish to thank each of the laboratory participants, for a clinically valuable contribution.

Resumo

A contagem do reticulócitos é usada extensamente na rotina laboratorial para avaliar a atividade eritropoiética da medula óssea, e é de grande importância no diagnóstico e no prognóstico na terapia de anemias hemolíticas. São coradas com o azul de metileno novo e o azul cresil brilhante, o que conferem o aspecto característico de retículo quando observado ao microscópio ótico. Critérios conhecidos foram observados para um bom desempenho da contagem manual de reticulócitos, principalmente, atenção especial nas películas do sangue durante a montagem das lâminas; contagem nos campos que não contém sobreposição celular; e também no número de células avaliadas. O objetivo deste estudo foi avaliar a variação interobservadores, analisar o erro estatístico da contagem manual dos reticulócitos, e demonstrar as limitações deste método. A análise de correlação intraclasses segundo Bartko [Psychol Rep 19:3,1966; 34:418,1974] foi usada para avaliar a concordância entre 12 observadores em um total de 25 lâminas do sangue, com contagens variadas de reticulócitos. Os resultados das análises estatísticas indicam que o erro casual, calculado como (1-r2) variou de 4 a 60% entre os observadores. Embora ocorra imprecisão entre os observadores, o perfil geral entre eles é similar, e o coeficiente de correlação intraclasse indicou que os resultados obtidos são clinicamente úteis. Rev. Bras. Hematol. Hemoter.

***Palavras-chave**: Hematologia; métodos estatísticos; contagem manual de reticulócitos.*

References

1. Wickramasinghe SN, Bain BF. Blood and Bone Marrow. 3. ed., New York: Churchill Livingstone, v. 2, 1986.
2. Lee, GR. et al. Wintrobe's Clinical Hematology. 10. ed. Baltimore: Lippincott William's & Wilkins, 1999.
3. Peebles DA, Hochberg A, Clarke TD. Analysis of manual reticulocyte counting. Am J Clin Pathol. 1981;76(5):713-7.
4. van Hove, L.; Schisano, T.; Brace, L. Anemia Diagnosis, Classification, and Monitoring Using Cell-Dyn Technology Reviewed for the New Millennium. Laboratory Hematology, 2000;6:93-108.
5. Bain, BJ. Células Sanguíneas: um guia prático. 2. ed., Porto Alegre: Artes Médicas, 1997.
6. Davis BH, Bigelow NC. Automated reticulocyte analysis. Clinical practice and associated new parameters. Hematol Oncol Clin North Am. 1994;8(4):617-30.
7. Lewis, SM.; Bain, BJ.; Bates, I. Hematologia Prática. 9 ed., São Paulo: Atmed, 2006

8. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Rep. 1966;19(1):3-11.
9. Bartko, J.J. Corrective note to: "the intraclass correlation as a measure of reliability." Psychological Reports, 1974;34:418.
10. Portney LG, Watkins MP. Foundations of Clinical Research. Applications and Practice Norwalk, Connecticut: Appleton & Lange, ISBN 0-8385-1065-5 p. 509-516, 1993.
11. Hong Kong University. On-line statistical programs. Access in September, 2008.
12. Shining, D; Shining, PD. Statistical control of process. In Jurans Quality Control Handbook. J.M. Juran and Frank M. Gryna Ed. London: McGraw Hill, 1988
13. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic. 1981;86(2):127-37.
14. Mckenzie, SB. Textbook of Hematology. 2. ed. Baltimore: William's & Wilkins, 1996.
15. Stiene-Martin, EA; Lotspeich-Steininger, CA; Koepke, J A. Clinical Hematology: Principles, Procedures, Correlations. 2. ed., Philadelphia: Lippincott, 1998.

Avaliação: Editor e dois revisores externos
Conflito de interesse: sem conflito de interesse

Recebido: 13/12/2008
Aceito após modificações: 16/02/2009