Transfusion and Apheresis Science 56 (2017) 914–919

Contents lists available at ScienceDirect

Transfusion and Apheresis Science

journal homepage: www.elsevier.com/locate/transci

Review

Sampling methods to the statistical control of the production of blood components

Paulo Pereira[a,*], Jerard Seghatchian[b,*], Beatriz Caldeira[c], Paula Santos[d], Rosa Castro[d], Teresa Fernandes[d], Sandra Xavier[e], Gracinda de Sousa[f], João Paulo de Almeida e Sousa[g]

[a] Department of Quality Assurance, Portuguese Institute of Blood and Transplantation, Avenida Miguel Bombarda 6, 1000-208 Lisbon, Portugal
[b] International Consultancy in Blood Components Quality/Safety Improvement, Audit/Inspection and DDR Strategy, London, UK
[c] Head for the Laboratory of Quality Control, Portuguese Institute of Blood and Transplantation, Portugal
[d] Medical Technician, Laboratory of Quality Control, Portuguese Institute of Blood and Transplantation, Portugal
[e] Department of Health, Polytechnic Institute of Beja, Beja, Portugal
[f] Directive Board, Portuguese Institute of Blood and Transplantation, Portugal
[g] President of the Directive Board, Portuguese Institute of Blood and Transplantation, Portugal

ARTICLE INFO

Article history:

Keywords:
Lot
Normality
Outliers
Rational subgroup
Sampling
Specifications

ABSTRACT

The control of blood components specifications is a requirement generalized in Europe by the European Commission Directives and in the US by the AABB standards. The use of a statistical process control methodology is recommended in the related literature, including the EDQM guideline. The control reliability is dependent of the sampling. However, a correct sampling methodology seems not to be systematically applied. Commonly, the sampling is intended to comply uniquely with the 1% specification to the produced blood components. Nevertheless, on a purely statistical viewpoint, this model could be argued not to be related to a consistent sampling technique. This could be a severe limitation to detect abnormal patterns and to assure that the production has a non-significant probability of producing nonconforming components. This article discusses what is happening in blood establishments. Three statistical methodologies are proposed: simple random sampling, sampling based on the proportion of a finite population, and sampling based on the inspection level. The empirical results demonstrate that these models are practicable in blood establishments contributing to the robustness of sampling and related statistical process control decisions for the purpose they are suggested for.

Contents

* Corresponding authors.
E-mail addresses: paulo.pereira@ipst.min-saude.pt (P. Pereira), jseghatchian@btopenworld.com (J. Seghatchian).

https://doi.org/10.1016/j.transci.2017.11.022
1473-0502/

1. Introduction

1.1. General principles of statistical process control and sampling in the production of blood components

The European Directive 2002/98/EC [1] requires the control of the production of blood components and in the United States is mandatory by the standards for blood banks and transfusion services [2]. The European Committee (Partial Agreement) on Blood Transfusion (CD-P-TS) is hosted by The European Directorate for the Quality of Medicines & HealthCare (EDQM) of the Council of Europe. The EDQM publishes the "Guide to the preparation, use and quality assurance of blood components" [3]. This guideline includes in Appendix D a policy for statistical sampling on a statistical process control (SPC) methodology perspective. Such as in other blood components specifications [4–6], there is a lack of standardized practices in the European Union.

The follow-up of the production conformity is critical to classify the production at least at an acceptable level of specifications similar to the industry levels. Its goal is to detect issues that could have a major impact on the production, affecting not only the conformity of the blood products but also the overall budget.

Regularly, a manufacturing process is classified as "world class quality" when the average results are "on target" with a "minimum variance." Working with "minimum variance" is feasible uniquely when the process shows a reasonable degree of statistical control. In blood establishments, the validation of the process of blood components' production is well established. The establishment must prove that the output delivers compliant components not only in validation stage but also individually according to specifications and per lot. This is reported to show the requirements are continuously satisfied. The SPC is the standard method to monitor the production processes [7].

SPC is a process built on objective evidence to improve the delivery of conforming lots of products. Predominantly, this methodology is used to improve the processes continuously. Therefore, it is proposed to decrease the variance, to control and report the stability of the operations, to follow-up the methods, to demonstrate the performance of the procedure, to previously detect for non-conformities about lots specifications, to document the process according to the law and other application requests. For further details on SPC on the production of blood components see [8].

In practice, skilled and successfully trained personnel are required. Moreover, it should be acknowledged that operative personnel with demonstrated competence in blood components specifications, statistics basics, and SPC involves a long-term formation by a trainer able of answering to the main requests. Thus, a consultant with experience in SPC in the components production should be hired when it is not possible to have some operative personnel fulfilling these requirements. In practice, it is known that is hard to have this employer. Furthermore, according to what is happening in blood establishments, it is suggested as a first step the support of a consultant and in a second phase to start the education and training of someone that could be the responsible to maintain and develop the SPC methodology.

Models to the determine the sample size, such as to select the samples in a particular production lot seems not to be systematically applied. EDQM requires a minimum of 1% of the produced blood components to be tested. Let's refer this specification as the "1% rule." This simple principle is not related to the number of the components, statistical sampling method, and consistency of the production process. Immediately, the statistician could argue that this model cannot be used to assure a proportional-to-size sampling related to a specific error. We understand that the 1% is associated with stable production processes, i.e., on operations where the variation does not have a significant impact on the specifications a small sampling size is suggested. However, the "1% rule" could be improved merely. Three statistical models are discussed: simple random sampling, sampling based on the proportion of a finite population and sampling based on the inspection level.

1.2. Sampling method applied to the lot of production

The goal of control of the manufacture of blood components could be considered to be the identification and measurement of the error arising from the variables on the collection and production stages. A sampling method is only practical when the statistical control of the production is not per component but lot. A common practice is to separate the produced blood components in lots to allow a more accessible and economical practice. On a statistical view, the lot is referred as the population on a period. A realistic sampling of the production is dependent on the selection of an adequate rational subgrouping method. A rule of thumb is to use only subgroups with independently, identically, and normally distributed (IIND) data, whenever possible. Hypothetically, the estimate of a subgrouping is regularly biased when compared to the production. Nevertheless, the application of an adequate methodology guarantees that the results are not significantly biased (non-significant bias). Some of the sampling models that could be applied are the simple random, systematic, and stratified sampling [9]. A significant risk could/can happen when a particular cause affecting a large number of distributed blood components with a long term lot, such a monthly lot. The number of samples should comply with the total number of blood components of the lot. Systematic sampling or interval sampling are proposed to be applied to the production. A sampling is designated from a lot per blood component at regular intervals (e.g., daily, weekly). This model is used in a continuing series of lots produced on a single blood establishment using one production methodology. The method is applied independently when existing more than one producer or process. Usually, it is employed as a measurable parameter on a continuous scale, e.g., on a daily scale.

This article discusses sampling models applied to the lots SPC of produced blood components on a proper statistical practice perspective. Data from the Quality Control lab archive of the Portuguese Institute of Blood and Transplantation, Portugal, is used where applicable.

2. Material and methods

2.1. Simple random sampling

The Council of Europe guide [3] proposes the number of samples n equal to 1% with a minimum between four and 10 components (according to the measured analyte) of the monthly production per blood component. Mostly, this approach is applied in European blood establishments regularly without a statistical sampling method. From the statistician standpoint, this technique could be criticized since the sample collection is not representative of the full production. This could be viewed as a severe limitation to the sample representativeness. For instance, if the 1% is collected in the first days and the results are conforming to the specifications, these results cannot be statistically inferred to the entire production and the risk associated to sampling cannot be recognized since a hypothesis test is not established. In this case, if the rest of the output is nonconforming, it is merely not identified since it is not sampled. Also, if an SPC method is applied to this sampling's results, it is useless due to the estimates are not based on a representative sampling. A suggested application of this approach could be the use of

the simple random sampling of blood components. This method assures that each sample is selected randomly by probability, signifying that each of the blood components of a particular lot has the same chance to be chosen. Since it is the most uncomplicated probability of sampling methods, it does not require complicated statistical skills to be computed. By means of Excel® (Microsoft®, Redmond, Washington, USA), a simple random sampling is calculated using the pseudo-random generator function RANDBETWEEN. Further information can be found elsewhere [9].

2.2. Sampling based on the proportion of a finite population

However, the concerns about the number of samples consistency remain considering the EDQM requirements [3]. A simple approach to be an alternative the "1% rule" could be used. The laboratorians' goal is to estimate a population proportion p so that the error is no larger than a certain margin of error ε. On an alternative view, it could be considered a sample proportion $\hat{p}$instead p; however, the statistical principles and outcome are similar. Since the population is equal to the lot, p is assumed. For instance, if a weekly lot was part of the monthly lot, it should be considered $\hat{p}$. The goal is to calculate a confidence interval such that $P \pm \varepsilon$. Usually, a confidence interval of 95% is assumed. The sample size necessary for estimating a p of a finite population with $(1-a)100\%$ confidence and error no larger than ε is:

$$n = \frac{Z^2 \cdot p(1-p)}{\varepsilon^2} / 1 + \frac{Z^2 \cdot p(1-p)}{\varepsilon^2 N} \quad (1)$$

where Z is the Z-score, p is the population proportion, ε is the margin of error, and N is the population size.

p is the proportion of the population with a specific characteristic of interest. For instance, a conforming/in-control hemoglobin result on the red blood cells population. It could be measured as follows:

$$p = N_1 / N \quad (2)$$

where N_1 is the number of population elements with a specific characteristic of interest, and N is the population size.

Table 1 displays a n series for ε of 0.01 (1%) to 0.05 (5%) based on a p of 0.5 and confidence intervals of 95% and 99%. As predictable, the sample size is directly related to N for the same confidence interval, p, and ε. For further details see [9].

2.3. Sampling based on the inspection level

ISO 3951-1 includes sampling procedures for inspection by variables. This guideline is intended to determine a particular n according to the size of a lot and the inspection level. It depends on the inspection level of the acceptance quality limit (AQL) defined as the "worst tolerable process fraction nonconforming when a continuing series of lots is submitted for acceptance sampling" (entry 3.6 of [10]). Its methodology is only applicable when is assumed that data is normally distributed. Therefore, the normality is verified before sampling begins (entry 18.1 of [10]). Usually, it is checked in the previous validation of the production method per blood component using tests such as the D'Agostino-Pearson K^2 normality test (see 5.1) [11].

The sampling starts with the s-method and if the production process shows to have a long-term stability moves to the σ-method. The stability is understood not only as the statistical stability of the central tendency measures, but also the systematical fulfillment of the individual requirements of each sample. s-method requires a more substantial number of samples since it is based on short-term data. Otherwise, σ-method is found in the long-term part of a systemically stable process of production, requiring shorter samplings. Consequently, the cost per blood components produced is higher using s-method. Even considering that each sample is validated individually, the use of σ-method since the product beginning due to being less costly represents a substantial risk to accept or reject lots erroneously. It could be the starting method also when the s-method is not used, but there is favorable considerable long-term evidence about the stability of the process. For instance, testing the retrospective standard deviation. The selection between σ-method and s-method must be clearly demonstrated, i.e.; it should be exposed a long-term stable standard deviation if σ-method is selected (entry 11 of [10]). A set of tables are applied to determine the appropriate n (see Tables 2–4).

The sampling includes three models: normal, tightened, and reduced inspection. Pereira (2016) proposed a model complementary to ISO 3951-1 [8]. On this methodology, the inspection level is according to the process capability index $\hat{C}_{pk}$ as follows:

a) For each of the methods the inspection level chosen in the starting of the sampling plan is the normal level (level II);
b) The change from normal to tightened inspection (level III) happens when the inspection results demonstrate that the capability of the process is insufficient ($\hat{C}_{pk} < 1$);
c) The switch from normal to reduced inspection (level I) takes place when the capacity of the process is systematically satisfactory ($\hat{C}_{pk} \geq 1.33$);
d) The change from tightened to normal inspection happens when the inspection results determine that the ability of the process is constantly capable ($1 \leq \hat{C}_{pk} < 1.33$), and;
e) The change from reduced to normal inspection occurs when the capacity of the process is classified as capable ($1 \leq \hat{C}_{pk} < 1.33$).

Sigma-metric is the $\hat{C}_{pk}$ multiplication by three. The capability index mathematical models are out of the scope of this article. These models can be found anywhere [7]. For further details on the application of a sampling based on the inspection level and the computation of $\hat{C}_{pk}$ on the production of blood components sampling see [8].

3. Results

3.1. Case 1: Simple random sampling for a lot of plasma, fresh frozen

Let consider the case where plasma, fresh frozen is tested for Factor VIII after freezing and thawing "as determined by SPC on units in the first month of storage" (entry Table 5D-1 of [3]). In this case, it is understood that the EDQM guideline requires that the sampling is based on the capability of the production to meet the specifications, i.e., plasmas with not less than 70 IU Factor VIII per 100 mL. For instance, if the one-month production is demonstrated to be capable, i.e., $1 \leq \hat{C}_{pk} < 1.33$ or higher, it could be assumed the "1% rule" in the first month of storage. However, also considering the European Pharmacopoeia, a minimum $n = 10$ is required to "carry out the test using a pool of not fewer than 10 units" (entry 07/2017:1646 of [12]). Therefore, if it is previewed to produce in the next month 300 plasmas, a minimum of 10 components are tested. The days of the month are randomly selected. For instance, = RANDBETWEEN(1,31) to choose a random number between the 1st and 31st. Therefore, $n = 10$ is chosen randomly, and the condition is repeated if the sample number is a duplicate by pressing the keyboard F9. In this example, the sampling occurs on the 2nd, 5th, 8th, 10th, 11th, 14th, 19th, 23th, 25th, and 27th day. In each day one plasma sample will be collected. Again, it uses the simple sampling to

Table 1
Required sample size n necessary for estimating a population proportion p with 95%, amd 99% confidence according to the margin of error ε.

Population Size, N	Confidence = 95%, p = 0.5				Confidence = 99%, p = 0.5			
	Margin of Error, ε				Margin of Error, ε			
	0.05	0.035	0.025	0.01	0.05	0.035	0.025	0.01
10	10	10	10	10	10	10	10	10
20	19	20	20	20	19	20	20	20
30	28	29	29	30	29	29	30	30
50	44	47	48	50	47	48	49	50
75	63	69	72	74	67	71	73	75
100	80	89	94	99	87	93	96	99
150	108	126	137	148	122	135	142	149
200	132	160	177	196	154	174	186	198
250	152	190	215	244	182	211	229	246
300	169	217	251	291	207	246	270	295
400	196	265	318	384	250	309	348	391
500	217	306	377	475	285	365	421	485
600	234	340	432	565	315	416	490	579
700	248	370	481	653	341	462	554	672
800	260	396	526	739	363	503	615	763
1000	278	440	606	906	399	575	727	943
1200	291	474	674	1067	427	636	827	1119
1500	306	515	759	1297	460	712	659	1379
2000	322	563	869	1655	498	808	1141	1785

Table 2
Sample size code letters for general inspection levels for inspection s-method and the sampling percentage.

Lot size, n		I			II			III		
		Letter	Higher [%]	Lower [%]	Letter	Higher [%]	Lower [%]	Letter	Higher [%]	Lower [%]
2	8	B	100.0	37.5	B	100.0	37.5	B	100.0	37.5
9	15	B	33.3	20.0	B	33.3	20.0	C	44.4	26.7
16	25	B	18.8	12.0	C	25.0	16.0	D	37.5	24.0
26	50	C	11.5	6.0	D	23.1	12.0	E	34.6	18.0
51	90	C	5.9	3.3	E	17.6	10.0	F	25.5	14.4
91	150	D	3.3	2.0	F	14.3	8.7	G	19.8	12.0
151	280	F	4.0	2.1	G	11.9	6.4	H	16.6	8.9
281	500	F	2.1	1.2	H	8.9	5.0	J	12.5	7.0
501	1 200	G	1.8	0.8	J	7.0	2.9	K	10.0	4.2
1 201	3 200	H	1.1	0.4	K	4.2	1.6	L	5.8	2.2
3 201	10 000	J	0.6	0.2	L	2.2	0.7	M	3.0	1.0
10 001	35 000	K	0.2	0.1	M	0.9	0.3	N	1.2	0.4
35 001	150 000	L	0.1	0.0	N	0.4	0.1	P	0.5	0.1
150 001	500 000	M	0.0	0.0	P	0.1	0.0	Q	0.1	0.0
500 000	+∞	N	0.0	0.0	Q	0.0	0.0	R	0.1	0.1

Table 3
Sample size code letters for general inspection levels for inspection σ-method and the sampling percentage.

Lot size, n		I			II			III		
		Letter	Higher [%]	Lower [%]	Letter	Higher [%]	Lower [%]	Letter	Higher [%]	Lower [%]
2	8	B	100.0	25.0	B	100.0	25.0	B	100.0	25.0
9	15	B	22.2	13.3	B	22.2	13.3	C	33.3	20.0
16	25	B	12.5	8.0	C	18.8	12.0	D	25.0	16.0
26	50	C	7.7	4.0	D	15.4	8.0	E	23.1	12.0
51	90	C	3.9	2.2	E	11.8	6.7	F	15.7	8.9
91	150	D	2.2	1.3	F	8.8	5.3	G	11.0	6.7
151	280	F	2.6	1.4	G	6.6	3.6	H	7.9	4.3
281	500	F	1.4	0.8	H	4.3	2.4	J	5.3	3.0
501	1 200	G	1.2	0.5	J	3.0	1.3	K	3.6	1.5
1 201	3 200	H	0.7	0.3	K	1.5	0.6	L	1.7	0.7
3 201	10 000	J	0.3	0.1	L	0.7	0.2	M	0.8	0.3
10 001	35 000	K	0.1	0.0	M	0.2	0.1	N	0.3	0.1
35 001	150 000	L	0.0	0.0	N	0.1	0.0	P	0.1	0.0
150 001	500 000	M	0.0	0.0	P	0.0	0.0	Q	0.0	0.0
500 000	+∞	N	0.0	0.0	Q	0.0	0.0	R	0.0	0.0

select the plasma from each of the daily series. The plasmas are listed from 1 to +∞ according to the order of production. For instance, = RANDBETWEEN(1,59) to choose a random number between the 1st and 59th plasma of a daily lot. In this case is the 37th plasma.

3.2. Case 2: Sampling for a proportion of a lot of platelets, recovered, pooled, leucocyte-depleted

Let consider a confidence interval of 95% (0.95), p = 0.999, ε = 0.05, and N = 100 for the volume (mL). p is computed for a

Table 4
Sample size code letters per inspection method.

Letter	*s*-method		σ-method	
	Normal and tightened inspection	Reduced inspection	Normal and tightened inspection	Reduced inspection
B	3	3	2	2
C	4	3	3	2
D	6	3	4	2
E	9	4	6	3
F	13	6	8	4
G	18	9	10	6
H	25	13	12	8
J	35	18	15	10
K	50	25	18	12
L	70	35	21	15
M	95	50	25	18
N	125	70	32	21
P	160	95	40	25
Q	200	125	50	32
R	250	160	65	40

total of 1452 *platelets, recovered, pooled, leucocyte-depleted*. The sampling applies to any of the parameters (entry Table 5C-3 of [3]). Let's consider the case that two non-conforming volumes are reported. Two non-conforming components are reported. Thus, $p = 1450/1452 = 0.999$. Typically, the worst margin of error is 0.05. Alpha divided by two is computed as follows: $(1 - 0.95)/2 = 0.025$. *Z*-score for a confidence interval of 95% is 1.96. However, a more accurate estimate could be computed on a spreadsheet using the following function: =NORM.S.INV(1-A1) (note: A1 cell is the Z-$\alpha/2$ result). This function returns the inverse of the standard normal cumulative distribution for a mean equal to zero and a standard deviation of one. The next parcels are calculated on the same worksheet:

$$f(x) := (\mathrm{A2}\hat{2}) * (\mathrm{A3} * (1 - \mathrm{A3}))/(\mathrm{A4}\hat{2})$$

where A2 cell is the Z-score, A3 cell is p, and A4 cell is ε, and;

$$f(y) := 1 + (\mathrm{A2}\hat{2}) * (\mathrm{A3} * (1 - \mathrm{A3}))/(\mathrm{A4}\hat{2} * \mathrm{A5}) \quad \text{where A5 cell is } N.$$

n is computed dividing $f(x)$ by $f(y) = 2$ samples (rounded up). The random sampling is used to identify the two samples of the lot (see 2.1).

The percentage (2% of the lot) is in agreement with the minimum 1% claimed by EDQM [3]. If the rate is below 1%, the sampling is increased accordingly.

3.3. Case 3: Sampling of a lot of red cells, leucocyte depleted in additive solution based on the inspection level

Let consider a daily lot of *red cells, leucocyte depleted in additive solution* components with $N = 84$. The sampling applies to any of the parameters (entry Table 5B-6 of [3]). The number of samples is determined as follows:

a) The lot size is selected choosing the letter of normal inspection level in the *s*-method (Table 2), i.e., letter E in level II (if it is starting used the σ-method Table 3 is applied), and;
b) Using Table 4, is selected the n according to the normal inspection level in the *s*-method, which is equal to nine samples.

The nine *red cells, leucocyte depleted in additive solution* components could be identified using a simple random sampling generator (see 2.1).

4. Discussion

It is confirmed that the simple random sampling application is an essential contribution to the random selection of blood components in blood establishments that follow the "1% rule." A non-statistical approach for this selection is not suggested since represents not only a bad practice but a significant chance to not detect deviations that could be a cause of non-conforming products.

Such as the previous sampling model, sampling for a proportion of a lot could be easily computed using a spreadsheet. If p is unknown, the most conservative value equal to 0.5 (largest sample size in order to p) should be used. A higher margin of error could be interpreted as the quantity of random sampling error in results of a study. On the other hand, larger margins of error represents less confidence (less realistic) that the samples reported results are close to the true numbers of the population. Therefore, ε expresses the chance of error per sampling. Using Table 1 to determine n should be easier than employing a formula; however, these tables are dependent from p for what they are not generally suited to be applied in a systematically stable industrial production, such as the cases of most of the processes to produce blood components. For instance, on the 2nd case $n = 80$ to a $p = 0.5$, $\varepsilon = 0.05$, and $N = 100$. So, use Table 1 in the example causes oversampling representing a serious negative impact on the blood establishment's budget. The population proportion is a simple retrospective calculation, for what its computation is strongly encouraged. Compared to the following of the "1% rule", the sampling for a proportion of a lot is an improvement since it is dependent on the proportion reported to a specific parameter in a particular establishment. Note that it is determined using the specification with the weakest performance (e.g., lowest capacity index or sigma level) to assure that the sampling is statistically not significative to the lot [13,14]. On the real-world case, the percentage is very close to the "1% rule." For what it is favored the interpretation that "1%" could be a percentage near to the sampling required to a stable production process of blood components.

Sampling based on the inspection level approach show to have several advantages over the use of uniquely the simple random sampling along with the "1% Mainly." Principally because it considers the retrospective data before starting a sampling plan and also due to enabling the user to change the sampling plan level according to the capability index or sigma level. Furthermore, the sample size is pre-determined, making it simpler to the user. Compared to the sampling for a proportion of a lot it requires a higher sample size. For instance, using the same data than on the 3rd case, $n = 2$ instead nine. Nevertheless, the use of the sampling based on

the inspection level is preferred since it is based on principles of control of industrial manufacturing and assures that the minimum of 1% blood components per lot is tested. Also due to this model to be depended from the capability index or sigma level, it is the most complex methodology to apply. Therefore, a suitable control of changes technique should be considered [3,15].

Just in any other quality control approach, there are some concerns related to a reliable sampling. The first is the normality importance of data distribution. It should be verified before to the start the sampling. The D'Agostino-Pearson K^2 normality test is a statistical inference technique designed to the examination that the underlying distribution of a random variable is normally distributed [11]. For testing that the blood components production is normally distributed, the skewness and kurtosis statistics are combined in the D'Agostino-Pearson K^2 test outcome. The test first computes the skewness and kurtosis to quantify how far from the normal distribution is regarding the asymmetry and shape. It measures how considerably each of these values differs from the value expected for a normal distribution, computing a single p-value from the sum of these differences. Moreover, the test should be applied to a number of samples higher than 20. Using the previous data on MedCalc® (Medcalc Software bvba, Ostend, Belgium), the p-value is $\geq$0.05 to the D'Agostino-Pearson test ($p = 0.5099$) for what the normality distribution is not rejected.

The data variability is another significant issue. Allegedly, the blood components result of specifications are biased or/and have a statistically significant variance caused by different interferents on the whole blood collection or due to the variability due to the used of apheresis equipment (e.g., volume). The laboratorian should verify statistically if these differences are significant. The significance of bias could be easily confirmed using the Student t-test between two groups. The analysis of variance (ANOVA) is an approach to check if there is a significant difference of the mean within and between groups at least in three categories. The impact of variation could be explored using merely the two-sample hypothesis F-test. If the null hypothesis H_0 is rejected, there is not a statistically significant difference between the mean at the 95% confidence level. For instance, in this scenario, a set of lots could be reunited in one single lot allowing a shorter n. Note that the normality of the data distribution should be verified in order to to apply the statistical tools correctly. Otherwise, the laboratorian should use non-parametric models. For further details on analyzing the variability of normally distributed data see [16].

A third concern is the data robustness. The detection of outliers should be implemented to assure that any observation that seems to diverge noticeably from others in a sampling is corrected or rejected [17]. A single outlier could increase the variation, changing the average and could even alter the acceptance of a lot. Outliers should be removed to avoid the risk to affect the statistical analysis of the sampling. Often the "common-sense" judgment of the analyst (or pair of analysts) is used. However, the decision is affected by human error, principally on the verification of large samplings. It is recommended to use alternatively statistical tests such as the Grubbs test to detect a single outlier [18]. When an outlier is removed it is reported the basis of the decision. If the result is a true result, it should not be treated equally. For instance, a true result cannot be treated as an outlier when is higher than 3s. When it is supposedly more than one outlier, the Generalized Extreme Studentized Deviate (ESD) test [19] or the Tukey test [20] could be used. For instance, considering the hematocrit results (ratio) in a normally distributed sampling with 90 *red cells, buffy coat removed, in additive solution* components. In this case the maximum and minimum values are 0.56 and 0.63. The normality could be tested using three tests in MedCalc® software the p-value is $\geq$0.05 to the Grubbs double-sided and ESD tests for what none outlier is identified which is also verified by the Tukey's test. The p-value value is understood as the hypothesis of a sample in this rational subgroup to be an outlier is not significant at the 0.05 significance level.

Summarizing, the "1% rule" considered outside of a statistical sampling plan is not suggested. Alternatively, it is recommended to compute the sampling using one of the discussed statistical approaches. The sampling based on the proportion of a finite population is favored in relation to the simple random sampling, and the sampling based on the inspection level is preferred over the two other models since it is directly related to the capacity index or sigma level of the production. The distribution and the variability of normally distributed data, such as on the outliers, should be evaluated to assure consistent input data.

References

[1] Commission Directive 2002/98/EC of the European Parliament and of the Council as regards setting standards of quality and safety for the collection, testing, processing, storage and distribution of human blood and blood components and amending Directive 2001/83/EC. Official Journal of the European Union; 2001. L33/30–40.

[2] American Association of Blood Banks. Standards for blood banks and transfusion services. 30th ed. Bethesda (MD): American Association of Blood Banks; 2016.

[3] European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components. 19th ed. Strasbourg: EDQM; 2017.

[4] Pereira P, Westgard J, Encarnação P, Seghatchian J, de Sousa G. Scheme for the selection of measurement uncertainty models in blood establishments' screening immunoassays. Transfus Apher Sci 2014;52(1):42–7, http://dx.doi.org/10.1016/j.transci.2014.12.018.

[5] Pereira P, Westgard J, Encarnação P, Seghatchian J, Sousa G. Quality management in the European screening laboratories in blood establishments: a view on current approaches and trends. Transfus Apher Sci 2015;52(2):245–51, http://dx.doi.org/10.1016/j.transci.2015.02.014.

[6] Pereira P, Westgard J, Encarnação P, Seghatchian J, de Sousa G. The role of uncertainty in results of screening immunoassays in blood establishments. Transfus Apher Sci 2015;52(2):252–5, http://dx.doi.org/10.1016/j.transci.2015.02.015.

[7] Montgomery D. Introduction to statistical quality control. 7th ed. New York (NY): John Wiley and Sons, Inc.; 2012.

[8] Pereira P. Statistical control of the production of blood components using variables and attributes charts, and capability indexes. In: Luna A, editor. Blood donations: attitudes, safety and future directions. Hauppauge (NY): Nova Science Publishers; 2016. p. 27–72.

[9] Thompson SK. Sampling. 3rd ed. New York (NY): John Wiley and Sons, Inc.; 2012.

[10] International Organization for Standardization. ISO 3951 Sampling procedures for inspection by variables – Part 1: specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL. 2nd ed. Geneva: ISO; 2013.

[11] D'Agostino R. Transformation to normality of the null distribution of g1. Biometrika 1970;57(3):679–81, http://dx.doi.org/10.1093/biomet/57.3.679.

[12] European Directorate for the Quality of Medicines and HealthCare. European Pharmacopoeia. 9.4th ed. Strasbourg: EDQM; 2017.

[13] Kotz S, Lovelace C. Process Capability Indices in Theory and Practice. London: 477 Arnold; 1998.

[14] George M. Lean Six Sigma: combining Six Sigma quality with Lean production speed. New York (NY): The McGraw-Hill Companies, Inc.; 2012.

[15] International Organization for Standardization. ISO 9001 Quality management systems – requirements. 5th ed. Geneva: ISO; 2015.

[16] Matthews DE, Farewell VT. Using and understanding medical statistics. 5th ed. Basel: S. Karger AG; 2015.

[17] International Organization for Standardization. ISO 16269 Statistical interpretation of data – part 4: detection and treatment of outliers. Geneva: ISO; 2010.

[18] Grubbs F. Procedures for detecting outlying observations in samples. Technometrics 1969;11(1):1–21, http://dx.doi.org/10.2307/1266761.

[19] Rosner B. Percentage points for a generalized ESD many-outlier procedure. Technometrics 1983;25(2):165–72, http://dx.doi.org/10.2307/1268549.

[20] Tukey J. Exploratory data analysis. Reading (MA): Addison-Wesley Publishing Company; 1977.