

## Review

Rubén Gómez Rioja\*, Débora Martínez Espartosa, Marta Segovia, Mercedes Ibarz, María Antonia Llopis, Josep Miquel Bauça, Itziar Marzana, Nuria Barba, Monserrat Ventura, Isabel García del Pino, Juan José Puente, Andrea Caballero, Carolina Gómez, Ana García Álvarez, María Jesús Alsina and Virtudes Álvarez

# Laboratory sample stability. Is it possible to define a consensus stability function? An example of five blood magnitudes

<https://doi.org/10.1515/cclm-2017-1189>

Received December 26, 2017; accepted March 16, 2018; previously published online May 5, 2018

## Abstract

**Background:** The stability limit of an analyte in a biological sample can be defined as the time required until a measured property acquires a bias higher than a defined specification. Many studies assessing stability and presenting recommendations of stability limits are available, but differences among them are frequent. The aim of this study was to classify and to grade a set of bibliographic

studies on the stability of five common blood measurands and subsequently generate a consensus stability function.

**Methods:** First, a bibliographic search was made for stability studies for five analytes in blood: alanine aminotransferase (ALT), glucose, phosphorus, potassium and prostate specific antigen (PSA). The quality of every study was evaluated using an in-house grading tool. Second, the different conditions of stability were uniformly defined and the percent deviation (PD%) over time for each analyte and condition were scattered while unifying studies with similar conditions.

**\*Corresponding author: Rubén Gómez Rioja**, La Paz – Cantoblanco – Carlos III University Hospital – Laboratory Medicine, Madrid, Spain; and Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain, E-mail: rgrijoa@salud.madrid.org. <http://orcid.org/0000-0002-3429-0427>

**Débora Martínez Espartosa:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Clínica Universidad de Navarra – Laboratory Medicine, Madrid, Spain

**Marta Segovia:** La Paz – Cantoblanco – Carlos III University Hospital – Laboratory Medicine, Madrid, Spain; and Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain

**Mercedes Ibarz:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Universitari Arnau de Vilanova – Laboratori Clínic Territorial ICS, IRB Lleida, Lleida, Catalunya, Spain. <http://orcid.org/0000-0003-0590-946X>

**María Antonia Llopis and Carolina Gómez:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Universitari Germans Trias i Pujol – Laboratory Medicine, Badalona, Spain

**Josep Miquel Bauça:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Department of Laboratory Medicine, Hospital Universitari Son Espases, Palma de Mallorca, Illes Balears, Spain

**Itziar Marzana:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Universitario Cruces – Laboratory, Barakaldo, País Vasco, Spain

**Nuria Barba:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and CatLab – Laboratory, Viladecavalls, Barcelona, Spain

**Monserrat Ventura, María Jesús Alsina and Virtudes Alvarez:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain

**Isabel García del Pino:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Complejo Hospitalario Universitario A Coruña – Area Laboratory, A Coruña, Spain

**Juan José Puente:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Clínico Universitario Lozano Blesa – Servicio de Bioquímica, Zaragoza, Spain

**Andrea Caballero:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Vall d’Hebron – Servei de Bioquímica, Barcelona, Catalunya, Spain

**Ana García Álvarez:** Sociedad Española de Medicina de Laboratorio – Comisión de Calidad Extra analítica, Barcelona, Spain; and Hospital Clínico Universitario San Carlos, Madrid, Madrid, Spain

**Results:** From the 37 articles considered as valid, up to 130 experiments were evaluated and 629 PD% data were included (106 for ALT, 180 for glucose, 113 for phosphorus, 145 for potassium and 85 for PSA). Consensus stability equations were established for glucose, potassium, phosphorus and PSA, but not for ALT.

**Conclusions:** Time is the main variable affecting stability in medical laboratory samples. Bibliographic studies differ in recommendations of stability limits mainly because of different specifications for maximum allowable error. Definition of a consensus stability function in specific conditions can help laboratories define stability limits using their own quality specifications.

**Keywords:** blood specimen collection; pre-analytical phase; specimen handling; stability.

## Introduction

Stability is the capability of a sample material to retain the initial property of a measurand for a period of time within specified limits when the sample is stored under defined conditions [1]. The stability is mainly affected by storage time and also by other factors that increase the metabolism of the analyte or cause the initial property to disappear, such as temperature, light exposure, solvent evaporation or stirring.

The change in the property over time can be expressed as percent deviation (PD%) or bias with respect to the initial value, whereas the deterioration can be defined as a function that represents the variation of PD% over time. It can also be graphically represented in a two-dimensional chart with the PD% in the y-axis and the time in the x-axis, similar to the habitual representation of the dose-response in relation to an interfering substance. In fact, both interferences and instability in a particular sample are the main components of specimen-specific systematic error. The stability limit for an analyte is defined as the point in time when the PD% reaches a given specification of maximum allowable error. Similar to dose-response graphs for interferences, stability limits may also be easily extrapolated from such graphs.

It should be noted that this equation is modified by the physical-chemical factors that facilitate the process of instability, mainly contact with air, temperature, light exposure or agitation. This means that the stability limit needs to be defined for specific conditions (e.g. blood glucose stability in whole blood in closed tube, at room temperature, protected from light and without stirring).

Although a myriad of experimental studies assessing the stability of most laboratory magnitudes in various

conditions may be found in the literature, there are few clinical guidelines with general recommendations for the laboratory. The most widely used is the recommendation of the German Society of Clinical Chemistry (DGKL), published in a WHO report and later updated [1]. These recommendations are based on the collection of published studies, and separated into two basic conditions of stability; the specimen, as obtained from the patient, and the sample, prepared for analysis. The controversy within this recommendation is that stability limits of the reviewed studies are directly assumed or arbitrarily summarized. Noticeable aspects affecting this situation are the variability in the methodology and the quality of the studies included (some of them were performed a long time ago, using materials that are nowadays obsolete) and the different criteria to define the degree of deterioration that hamper the establishment of a rejection recommendation (maximum allowable error), even though the document includes a specific recommendation for the MAE based on biological variation (half of the total allowable error), included studies use different criteria, sometimes purely metrological or statistical. This may be due to the absence of clinical guidelines for the design of stability studies in clinical samples, in contrast with other components of total error, such as imprecision, accuracy or interferences. Therefore, the studies published follow different procedures, have variable number of samples, different statistical approaches and different maximum allowable error criteria.

The aim of this study was to classify and to grade a set of bibliographic studies about the stability of five common blood measurands and subsequently generate a consensus stability function. For this pilot study we selected five magnitudes, four of them with well-known stability problems (glucose, potassium, phosphorus and specific prostate antigen) and one of them usually mentioned as stable (ALT). These models could help medical laboratories to define stability limits based on their own specifications of maximum allowable error for the pre-analytical phase.

## Materials and methods

### Scientific bibliographic search

The analytes selected for the validation of this method were alanine aminotransferase (ALT), glucose, potassium, phosphorus and prostate-specific antigen (PSA).

The search strategy in PUBMED included: "(sample OR samples OR specimen OR specimens) AND (stability OR deterioration OR preanalytical OR pre-analytical)" in the last 15 years. A total of 1084 entries were obtained.

The acceptable articles were selected by two members of the Committee by checking the abstracts. Acceptability included inclusion of ALT, glucose, potassium, phosphorus or PSA in blood in the study. Forty-eight references were finally included [2–50].

## Acceptability of bibliographical references

In order to grade the quality of the papers, a checklist was defined including all variables that could affect the stability of selected magnitudes in blood (Table 1). Every item was rated as 0 if not considered for the study design, 1 if included and 2 if it improved the design. The maximum score was 29 points. References were classified into four grades: excellent if the score was above 21.75 points, acceptable if it was between 14.5 and 21.75 points, doubtful if between 7.25 and 14.5 points and unacceptable if the score was lower than 7.25 points. Scoring and grading of every reference was performed by the members of the Committee, in pairs. In case of discrepancies in grading, a new revision by another member was performed and an average score was obtained.

## Stability conditions

In order to compare the primary data of PD%, it is essential that the different studies have similar experimental conditions. Hence, a second checklist was prepared to compile the performance conditions of each stability experiment included.

The items included in this checklist were:

1. Strict definition of the system: the sample type (whole blood, serum or plasma) and the container/additive used (including manufacturer reference when available).
2. Procedure for blood collection.
3. Temperature of collection and storage.
4. Light exposure.
5. Stirring.
6. Evaporation (i.e. open tube during study).
7. Sample storage until the analysis.
8. Laboratory instrumentation.
9. Analytical methodology.
10. Batch analysis. (i.e. whether analyses were made in one single batch or multiple series).

## Stability models and statistical analysis

PD% for each magnitude and stability condition was modeled by linear or curvilinear regression using the least

squares adjustment over time. The regression equation was forced to include the point (0.0) (the model does not include any constant) as by definition, PD% was zero at the beginning of the experiment. The adjustment equation (linear, polynomial or exponential) was selected by improvement of  $R^2$ . The strength of the fit was assessed by the Pearson's coefficient of correlation, considering  $r > 0.7$  as a good adjustment, and its significance by Snedecor's F-test (ANOVA). Statistical significance was set at 0.01. IBM SPSS v21 software was used for all statistical analyses.

## Results

Forty-eight bibliographical references were selected, 11 of which were discarded as not presenting valid primary data to calculate PD%. In the 37 valid references, up to 130 different experiments were included; 48 in whole blood, 30 in plasma and 52 in serum. Of the five magnitudes assessed, 629 valid stability PD% data were collected: 180 for glucose, 113 for phosphorus, 145 for potassium, 106 for ALT and 85 for PSA (35 for total PSA and 50 for free PSA). Forty data points for glucose stability studies used glycolysis inhibitors, and were excluded due to the heterogeneity of tube types.

In the 48 peer-reviewed references, the average score was 18.3 points (SD 3.83, range 8.0–24.5), and were graded as: three unacceptable, seven doubtful, 28 acceptable and 10 excellent. Only in two cases there was a discrepancy between reviewers, and a third revision was included. The items included in the evaluation checklist are outlined in Table 1, as well as the percentage of studies found for every condition.

Most studies used healthy subjects (60.4%) instead of patient samples. The typical sample size was between 10 and 30 individuals (45.8%), often with a single determination (68%). The procedure of blood collection was frequently defined (73%) and often appropriate in terms of duration and volume (62.9%). The venipuncture devices were usually commercial and vacuum systems were used (90.5%) which could warrant a controlled contact with air during the experiment. However, in 52% of the references this aspect was not clearly reflected.

Temperature was controlled throughout the process in 87.5% of the references, but not agitation (31%), nor light exposure (27%). Centrifugation conditions, including force and time and sample storage until centrifugation were properly defined in 60.5% and 82.1% of the references, respectively.

Once the stability experiment finished, samples were analyzed immediately or deep-frozen to be processed in batch (95.3%). The analytical methods were correctly

**Table 1:** Quality Grading tool.

Item	Options	Score (points)	% Studies
Sample origin	– Healthy subjects	1	60.4
	– Patients	2	39.6
Sample size, n	– <10	0	6.25
	– 10–30	1	45.8
	– >30	2	47.9
Is the blood collection procedure defined?	– No	0	27
	– Yes	1	73
Total drawn volumen (mL) and phlebotomy time	– Defined and appropriate (<60 mL, <1 min)	2	62.9
	– Defined, inappropriate (>60 mL, >1 min)	1	14.5
	– Not defined	0	22.6
Are the pre-analytical conditions controlled?	– Yes (fasting, rest, time of collection)	1	60
	– No	0	40
Is the blood collection tube defined?	– Commercial vacuum tube	2	90.5
	– Tube without vacuum	1	1
	– Not defined	0	8.5
Is the temperature controlled during the experiment?	– Yes	1	87
	– Not defined	0	13
Is stirring controlled?	– Yes	1	31
	– Not defined	0	69
Is light exposure controlled?	– Yes	1	27
	– Not defined	0	73
Is evaporation controlled?	– Yes	1	48
	– Not defined	0	52
Repose before centrifuging (serum), min	– Serum <60, plasma <30	2	82.1
	– Serum >60, plasma >30	1	14
	– Not defined	0	3.9
Type of centrifugation	– Recommended	2	60.5
	– Different of recommended	1	17
	– Not defined	0	22.5
Sample storage between end of experiment and analysis	– Analyzed immediately or deep frozen until analysis	2	87.4
	– Inappropriate storage until analysis	1	7.9
	– Not defined	0	4.7
Analyzer	– Defined	1	88
	– Not defined	0	12
Method	– Commercial	1	85
	– Not defined	0	15
Metrological quality	– Defined, appropriate	2	62
	– Defined	1	6.2
	– Not defined	0	31.8
Repetition of analysis	– >Duplicate	2	5.5
	– Duplicate	1	26.5
	– Simple	0	68
Is confidence interval of the difference calculated?	– Yes, in each case	2	63.2
	– Yes, only in some cases	1	6.3
	– Not defined	0	30.5
Signification of differences	– Defined	1	73
Statistical hypothesis testing	– Not defined	0	27

Description of the variables that affect the stability and scores for each, together with the percentage of studies for every item.

defined in 85% of the references, although the quality assurance of the process was not mentioned in one third of cases (31.8%). The confidence interval for PD% was lacking in 30.5% of the articles. Moreover, 27% of the studies did not perform any analysis of the differences.

Maximum allowable difference specifications were based on biological variation in 17 studies, on metrological criteria in 12 studies and exclusively statistical in 13 studies.

Twelve studies did not include any kind of data regarding PD% for their magnitudes, neither graphs that

allowed an estimate, therefore they were excluded from stability models [7, 11, 14, 22, 27–29, 41, 44, 46–48].

### With regard to stability conditions

1. Strict definition of the system. Thirteen different conditions were defined, including 57 studies in whole blood, 51 in serum and 22 in plasma. In the whole blood studies, seven different conditions were considered (without additives 5, EDTA 4, heparin 12, with glycolysis inhibitors 16, non-thrombin procoagulant and gel 15, non-thrombin procoagulant without gel 4, and with thrombin procoagulant 1). In the serum studies, two conditions were suggested; non-thrombin procoagulant (49 studies) or thrombin procoagulant (two studies). In the plasma studies there were 16 with heparin, two with glycolysis inhibitors and one with EDTA. Due to the great variety of conditions, we moved to a basic classification in two types; whole blood vs serum/plasma. The great difference between these conditions is cell contact, which is complete in whole blood, and reduced to the presence of residual cells in serum or plasma. Studies using a glycolysis inhibitor were separated in the case of glucose.
2. Blood collection procedure. The most frequent design was the collection in primary tubes and used directly for analysis (74 studies, 57%). In 39 studies, collection was performed in the primary tube, but an aliquot was used to perform the analysis. In seven cases, the drawn tubes were mixed to prepare a pool and in seven cases all the necessary volume was collected in a different device and secondary tubes were prepared for the stability study. Blood collection procedure was not defined in three cases.
3. Temperature of collection and storage. Regarding the collecting temperature, the great majority of studies (118) were carried out at room temperature. With respect to storage, temperature was usually controlled in a narrow range, according to its influence on stability. The most studied temperatures were 4 °C (28 experiments), 22 °C (32 experiments) and 27 °C (14 experiments), although many studies used other temperatures. A basic classification was set as refrigerated (range 4–9 °C) or room temperature (RT) (15–27 °C).
4. Light exposure. As already mentioned as a quality parameter, only 22 experiments specified protection from light during the period assessed.
5. Stirring. It was only controlled in 16 studies.
6. Evaporation. In four studies, the effect of evaporation was studied (open tube), in 58 studies the tubes were

kept capped to avoid evaporation and in 68 studies it was not specified.

7. Batch analysis. One hundred and one studies were analyzed in several batches and 29 in a single batch. Studies made in a single batch are less prone to imprecision, but need an extra storage time that could bias the results.
8. Sample storage until analysis. In 29 single batch studies, samples needed to be stored until batch analysis, usually frozen or deepfrozen storage [16], however, 14 were refrigerated.
9. Laboratory instrumentation/Analytical methodology. More than 50 different instruments and a great variety of methods were registered.

### Stability models in basic conditions

With the PD% data collected from the studies, a scattergram of PD% over time expressed in hours was prepared. A regression was made using least squares and the significance and strength of the adjustment were verified. We evaluated the points obtained under similar stability conditions for each magnitude. Initially, only sample type (whole blood versus serum/plasma) and the most common temperatures (room and refrigerated) were considered. Table 2 shows the number of references and points included, the goodness of fit (Pearson's  $r$  for each model), the statistical significance ( $p$ -value) and the selected equation.

Glucose stability depends mainly on cell contact. The uptake and metabolism depend on the activity of the glycolytic enzymes, which explains why refrigeration is commonly used as a preservation method [51]. In whole blood, there is a decrease of approximately 1%/h at room temperature and 0.3%/h in whole refrigerated blood. In both cases, the equation that best adjusts is a quadratic model although the linear model could also be considered as acceptable (Figure 1). There is a clear increase in the rate of degradation with temperature that can be verified including the quantitative variable in a multiple regression model where both factors are significant (coefficient for time:  $B = -0.6\%/h$ , temperature:  $B = -1.6\%/^{\circ}C$ ). Although the average equation obtained shows a slope of 1%/h, studies at temperatures around 20 °C are close to 2%/h. Studies of accelerated degradation at 30–37 °C yield a decrease approaching 4%/h (Figure 2). Degradation is also observed in serum/plasma, with no marked effect on temperature. The decrease is around 0.1%/h.

The case of phosphorus is compelling; only a clear stability loss in whole blood at room temperature is

Table 2: Stability models in basic conditions.

Magnitude	Stability condition	n	PD% points	Adjust model	p-Value	r	Slope (Standard error)
Glucose	Whole blood RT	9	55	Linear	<0.001	0.788	-0.944 (0.094)
				Quadratic	<0.001	0.889	-1743/0.008 (0.135/0.001)
Glucose	Whole blood refrigerated	2	15	Linear	<0.001	0.820	-0.285 (0.053)
				Quadratic	<0.001	0.948	-0.699/0.003 (0.083/0.001)
Glucose	Serum/plasma RT	7	30	Linear	<0.001	0.835	-0.071 (0.010)
				Quadratic	<0.001	0.875	-0.128/0.000 (0.024/0.0001)
Glucose	Serum/plasma refrigerated	6	29	Linear	<0.001	0.685	-0.108 (0.022)
				Quadratic	<0.001	0.773	-0.216/0.001 (0.042/0.0001)
Phosphorus	Whole blood RT	10	59	Linear	<0.001	0.937	3.776 (0.187)
				Quadratic	<0.001	0.959	1.826/0.026 (0.394/0.005)
				Cubic	<0.001	0.975	-1.438/0.135/-0.001 (0.635/0.019/0.0001)
Phosphorus	Whole blood refrigerated	3	12	Linear	<0.001	0.877	0.209 (0.035)
				Quadratic	<0.001	0.992	-0.110/0.002 (0.028/0.0001)
Phosphorus	Serum/plasma RT	4	20	Linear	<0.001	0.888	0.071 (0.008)
				Quadratic	<0.001	0.921	0.126/0.000 (0.022/0.0001)
Phosphorus	Serum/plasma refrigerated	5	18	Linear	<0.001	0.902	0.028 (0.003)
				Quadratic	<0.001	0.904	0.033/-0.00004 (0.10/0.0001)
Potassium	Whole blood RT	12	78	Linear	<0.001	0.824	2.366 (0.185)
				Quadratic	<0.001	0.986	-0.791/0.031 (0.124/0.001)
Potassium	Whole blood refrigerated	3	12	Linear	<0.001	0.978	2.999 (0.191)
				Quadratic	<0.001	0.999	4.758/-0.13 (0.156/0.001)
Potassium	Serum/plasma RT	6	25	Linear	<0.001	0.839	0.811 (0.108)
				Quadratic	<0.001	0.910	-0.295/0.028 (0.282/0.007)
Potassium	Serum/plasma refrigerated	6	25	Linear	0.001	<b>0.614</b>	0.070 (0.018)
				Quadratic	0.002	<b>0.647</b>	0.137/-0.001 (0.055/0.001)
ALT	Whole blood RT	7	47	Linear	0.007	<b>0.385</b>	-0.066 (0.023)
				Quadratic	0.018	<b>0.404</b>	-0.026/0.000 (0.05/0.0001)
ALT	Whole blood refrigerated	3	15	Linear	<b>0.182</b>	<b>0.352</b>	-0.019 (0.013)
				Quadratic	<b>0.272</b>	<b>0.426</b>	0.017/0.000 (0.040/0.0001)
ALT	Serum/plasma RT	4	20	Linear	<b>0.752</b>	<b>0.073</b>	0.009 (0.029)
				Quadratic	<b>0.022</b>	<b>0.588</b>	0.176/-0.001 (0.060/0.000)
ALT	Serum/plasma refrigerated	3	16	Linear	<b>0.238</b>	<b>0.302</b>	0.020 (0.016)
				Quadratic	<b>0.160</b>	<b>0.480</b>	0.087/0.000 (0.045/0.0001)
fPSA	Whole blood RT	3	11	Linear	0.001	0.839	-0.370 (0.076)
				Quadratic	<0.001	0.967	-1.047/0.032 (0.126/0.006)
fPSA	Whole blood refrigerated	1	4	Linear	0.007	0.966	-0.252 (0.039)
				Quadratic	<b>0.066</b>	0.967	-0.245/0.000 (0.201/0.009)
fPSA	Serum/plasma RT	3	11	Linear	0.002	0.806	-0.094 (0.022)
				Quadratic	0.002	0.861	-0.196/0.001 (0.061/0.0001)
fPSA	Serum/plasma refrigerated	3	10	Linear	<0.001	0.977	-0.153 (0.011)
				Quadratic	<0.001	0.987	-0.239/0.001 (0.038/0.0001)
PSA	Whole blood RT	3	9	Linear	<b>1</b>	<b>0</b>	
				Quadratic	<b>0.99</b>	<b>0.025</b>	
PSA	Whole blood refrigerated	0	0		-	-	-
PSA	Serum/plasma RT	2	7	Linear	<0.001	0.948	-0.041 (0.006)
				Quadratic	0.001	0.964	-0.064/0.000 (0.017/0.0001)
PSA	Serum/plasma refrigerated	2	6	Linear	<b>0.061</b>	0.732	-0.010 (0.004)
				Quadratic	<b>0.191</b>	0.750	-0.020/0.000 (0.021/0.0001)

n, number of included studies; RT, room temperature. Slope: regression unstandardized coefficient for time (h) (linear model), time (h)/time (h)<sup>2</sup> (quadratic model), time (h)/time (h)<sup>2</sup>/time (h)<sup>3</sup> (cubic model). Bold: non significant.

observed. In refrigerated whole blood, a very slight elevation was observed (0.2%/h). Serum and plasma produced an even lower elevation of 0.07%/h and 0.03%/h at room

temperature and refrigerated, respectively. In the case of whole blood at room temperature, the adjustment by a cubic model significantly improved the linear adjustment

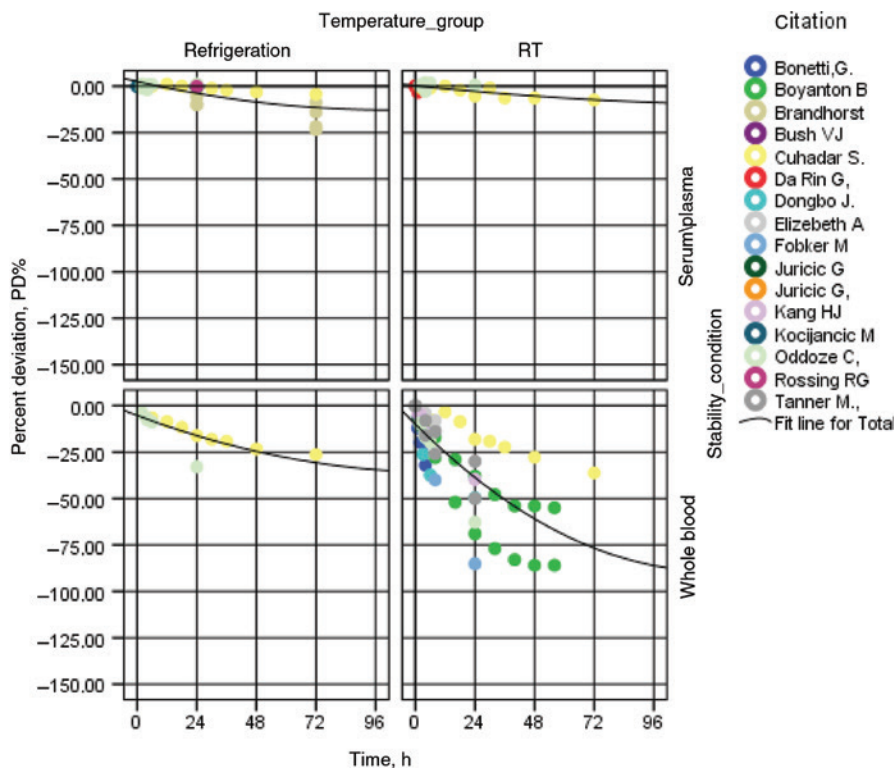


Figure 1: Glucose stability in basic conditions.

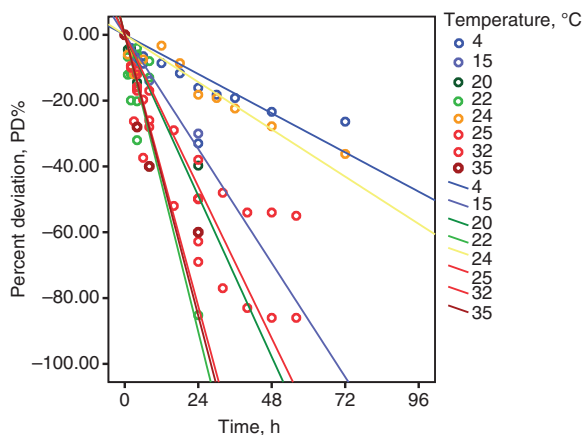


Figure 2: Glucose stability dependency on temperature.

due to a different behavior in the first 24 h and in the later time points. There is an initial period of up to 12–18 h in which the phosphorus decreased (possibly explained by entry into the red blood cells), from 18 to 24 h phosphorus increased exponentially, rising to 400% in 2–3 days (probably by degradation of the phosphate esters) [2, 3] (Figure 3).

For potassium, a temperature-dependent behavior was observed, similar to glucose, but in the opposite direction. In refrigerated whole blood, it undergoes a greater

increase of approximately 3%/h, which is clearly observed in the first hour of storage. Nevertheless, at room temperature, no significant modification was seen in the first 24 h. After 24 h, the function changes to a linear increase similar to that observed in the refrigerated samples. This behavior also seems to occur in serum/plasma at room temperature, however, a clear increase in the refrigerated sample was not observed, with a weaker adjustment of the equation ( $r < 0.7$ ) (Figure 4).

For ALT, there are some studies indicating an increase and others indicating a decrease in both whole blood and serum/plasma. Only in the case of whole blood at room temperature a significant regression line may be estimated, with a weak fit ( $r = 0.3$ ), in the rest of conditions it was not possible to adopt a consensus equation (Figure 5).

For free PSA, a very similar decrease in serum/whole blood was observed in both refrigeration and room temperature. It could even be simplified in a single model with a decrease of 0.25%/h. This was significant in all cases with an acceptable linear model. In the case of PSA, there were few studies, with no apparent loss of stability over time. A significant adjustment was not achieved except in the serum/plasma sample at room temperature, where there was a slight decrease of 0.04%/h (Figure 6).

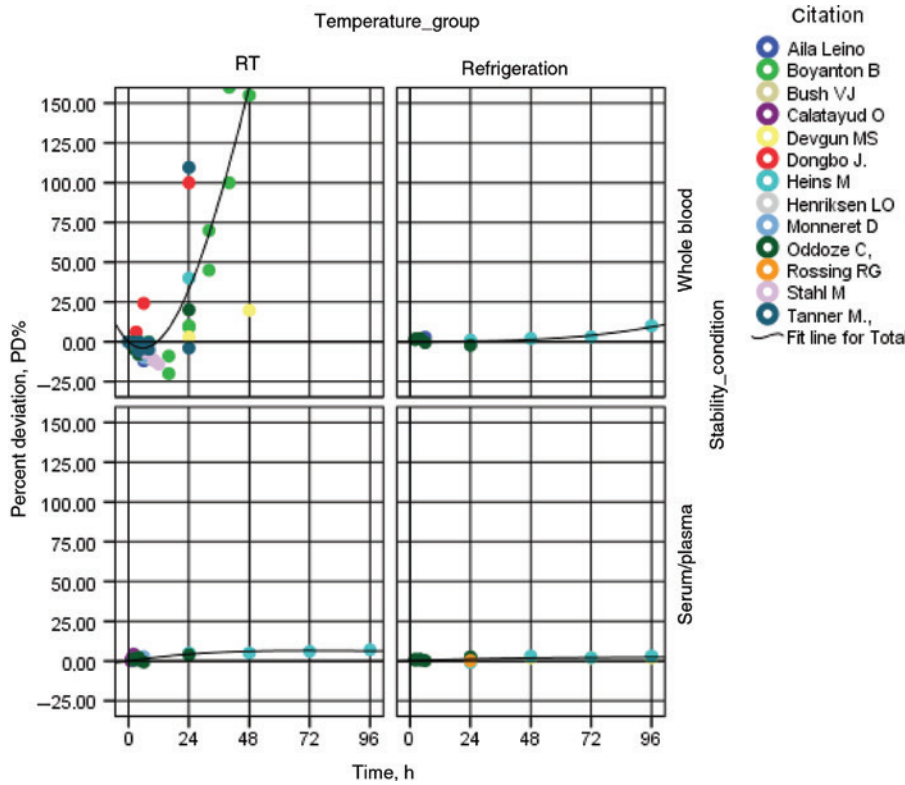


Figure 3: Phosphorus stability in basic conditions.

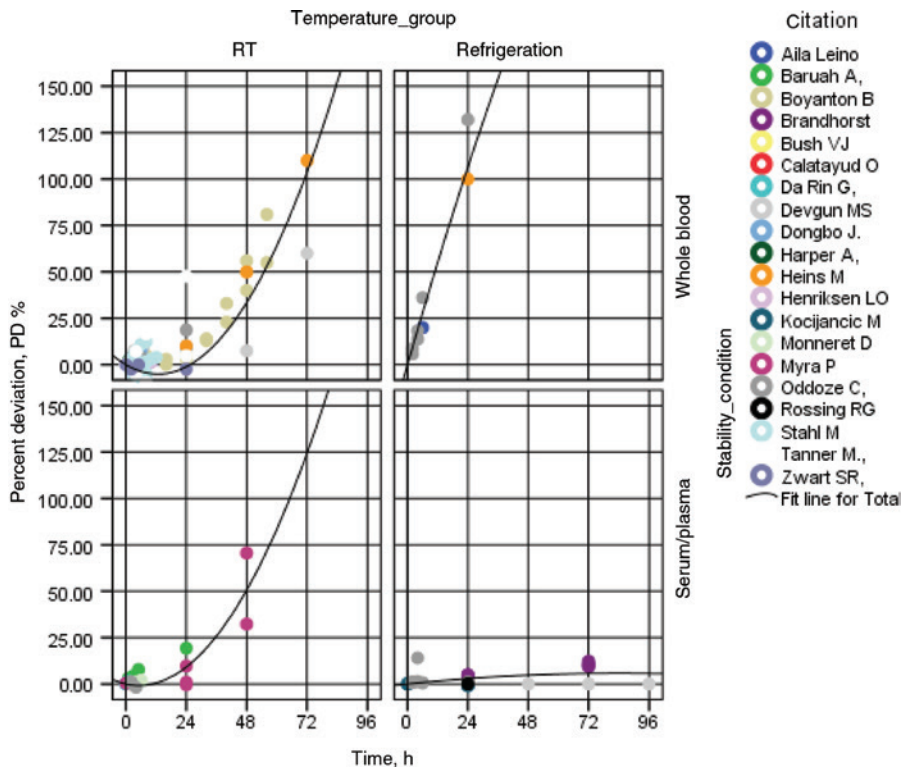


Figure 4: Potassium stability in basic conditions.

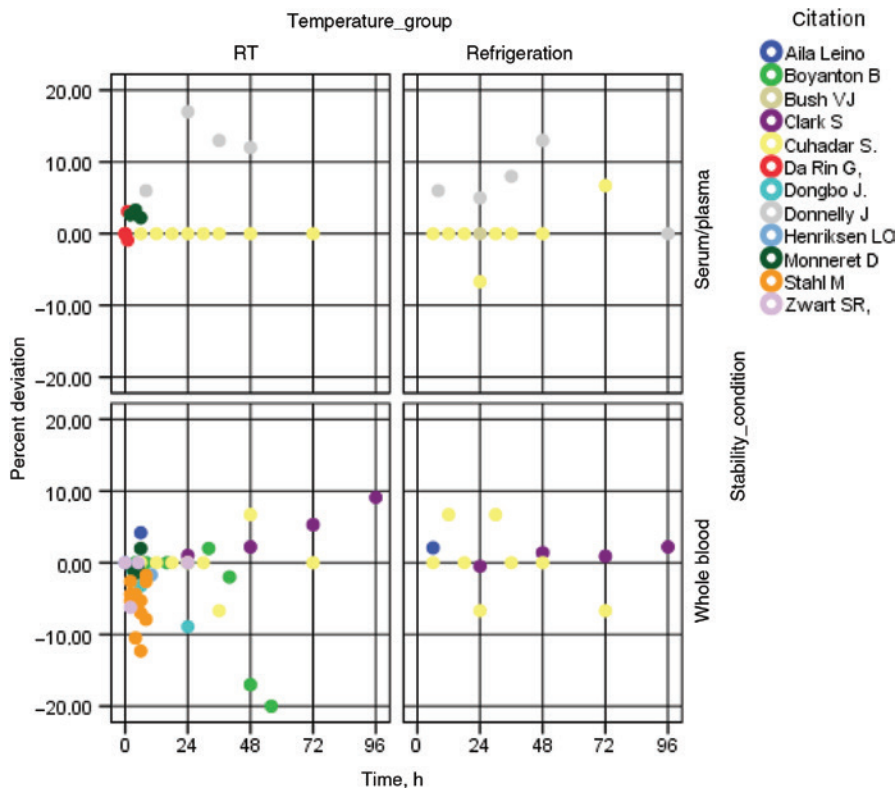


Figure 5: ALT stability in basic conditions.

These functions can be used to define stability limits applying several specifications for maximum allowable error. For glucose, when considering very demanding specifications (desirable systematic error based on biological variation 2.34% [52]) or mild specifications (minimum consensus specifications from Spain 12% [53]) the stability limit would range between 2 and 13 h in whole blood at RT and between 22 and 111 h in refrigerated serum (Table 3).

## Discussion

In laboratory medicine, traceability of the pre-analytical phase is essential. Two critical aspects are the storage temperature and the time from sample collection until its processing. The loss of stability is likely to involve an unacceptable error in the result of a measured component that may lead to wrong clinical decisions. That is why laboratories should define stability limits for each analyte and sample, and use them as a cause of rejection before processing the sample, in spite of no apparent visible deterioration [54].

In this study we considered the possibility of using the cumulative information from multiple studies to

generate a consensus function of stability in specific situations of stability for some magnitudes with known stability problems. Four of the selected analytes to validate this method were well known to be stability-affected; glucose, potassium, phosphorus and PSA and one analyte with doubts about stability (ALT).

The first step was to grade the quality of the stability studies. The quality of the published stability studies with the designed checklist is acceptable/excellent in the majority of cases (80%), although some usual defects were detected. A high percentage of studies did not consider essential variables such as evaporation (52%), light exposure (73%), or agitation (69%). In the design of the checklist, it was originally thought to consider the absence of control of these variables as a reason for direct rejection of the study, as well as the absence of statistical study. We observed later that the main problem was the absence of primary data and this frequently concurred with low quality articles. In 11 of 48 studies (23%) no usable primary data were available, neither in table nor in figure form. This prevents the results from being confirmed by other studies. Rejected studies included the three ones classified as unacceptable.

A possible explanation to this variability is the absence of guidelines or recommendations regarding the

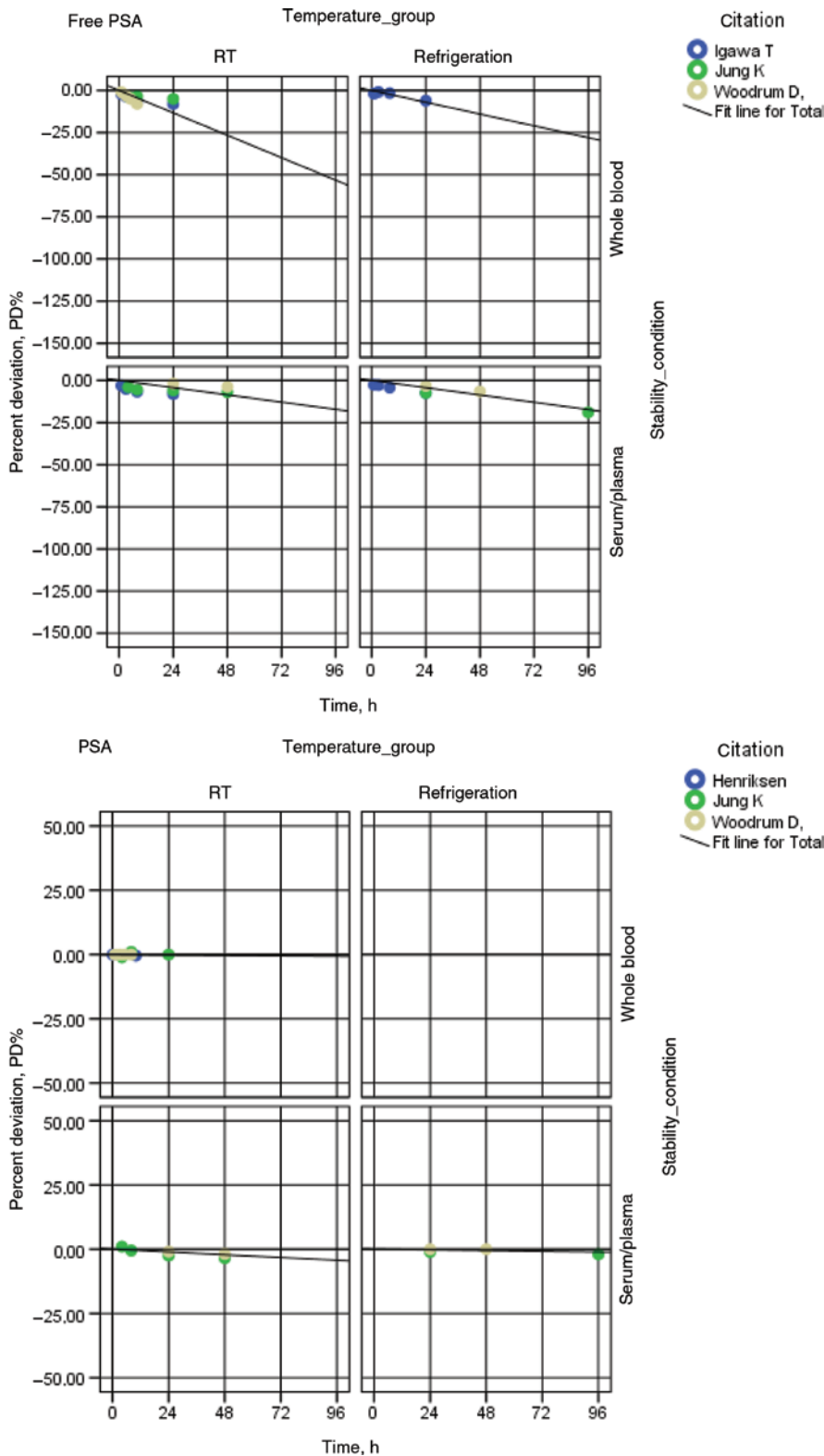


Figure 6: PSA stability in basic conditions.

way a stability study should be carried out. An agreement on the metric to be used is the minimum deemed necessary. When dealing with a systematic error, it would

be reasonable to use the percent deviation in relation to time, as recommended by the German Society of Clinical Chemistry [1] or by CLSI for reagent estability [55].

**Table 3:** Stability limit for glucose applying different specifications for maximum allowable error to the consensus stability model.

Specification	Maximum allowable error, %	Stability limit, h	
		Whole blood RT	Serum/plasma refrigerated
Systematic error (BV)	2.34	2	22
Total error (BV)	6.96	7	64
RCAP	8	8	74
Consensus minimum specifications	12	13	111

In our non-exhaustive literature review, data were collected from about four to eight studies for each minimum study condition (specimen and temperature), with an average of 10–30 points for each model. With this number of studies we achieved reliable consensus functions in most of the magnitudes with a reasonable statistical confidence in the model. In the magnitudes studied, it was possible to define a basic model (whole blood and serum at room temperature or refrigerated) that could cover general working conditions in four of them including PSA, in spite of the low number of studies included. For ALT it was not possible due to the heterogeneity of the results. This magnitude was selected because its instability was only occasionally mentioned. This approach seems ineffective in defining any small loss of stability that should be demonstrated by conclusive experimental studies.

In the case of phosphorus, if strict specifications are applied (desirable systematic error, 3.4%) stability is lost in the first 3 h due to the initial decrease. However, with more extensive specifications, (desirable total error 10%; minimum consensus specification 16%) there would be no loss of stability until 16–18 h, during the second phase of increase. In fact, some authors have reported 3 h stability for phosphate [13, 39], while others have described 24 and even 72 h [2, 3, 24, 26, 49].

In the case of potassium, contradictory results were obtained. In serum, the stability is lower at room temperature than refrigerated, conversely as occurs in whole blood. The theoretical mechanism of onset of potassium is the output of intracellular content that should be dependent on the number of residual cells (greater in plasma than in serum) and dependent on the same manner of inactivation of ATPase by cold [56]. It is very likely that the way of preparation and separation in the experiments is a very important factor to consider in their comparability.

For glucose, contact with cells and temperature were well known factors of instability. In whole blood a difference of 5–10 °C could double the loss of stability. However, the definition of maximum allowable error could be the main factor that justifies the differences in published stability limits.

The stability limit for a magnitude depends on the storage conditions defined by the laboratory, but also largely on the specifications that are selected to define the maximum allowable error.

The usual definition of total analytical error (TAE) does not include components of systematic error due to the sample, which considers the effect of stability loss or deterioration by *in vitro* hemolysis, among others [57]. The design of all laboratory methods includes the selection of the most appropriate specimen/sample and the acceptability conditions of the sample, which usually appear as a rule of thumb; “stable in whole blood with EDTA for 8 h...”, however, it is widely known that deterioration is usually a continuous function over time, depending on the physical-chemical conditions of the sample. For blood samples, time, temperature, contact with atmospheric air, or light exposure and agitation can substantially modify the deterioration function in a predictable amount if complete traceability of preanalytical phase were available. Consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine stated “it is acknowledged that, for patient care, optimizing the quality of the total (pre-analytical/analytical/post-analytical) examination process is the ultimate goal and therefore it would be desirable to go beyond setting analytical performance specifications and to establish examination performance. In principle, the performance specifications for the pre- and post-analytical laboratory processes should follow the models as for analytical performance specifications. When components of these additional phases can be expressed in numerical terms, they should be added in defining examination performance specifications” [58].

We found that in many cases the adjustment of the stability curves is not linear, even in the case of phosphorus a biphasic behavior was detected. It is important to note that in clinical samples there may be multiple simultaneous effects that could provoke unexpected responses in stability of clinical samples.

As limitations of this study it is necessary to emphasize the non-exhaustive nature of the bibliographical

search. The aim was to test a strategy of unification of results that has only been possible to verify in magnitudes with known problems of stability and for basic conditions. There are likely to be many more stability studies that could expand and improve the model. In the case of ALT, which we chose as an example of magnitude considered stable although a decrease in enzymatic activity would be expected over time, the results of the studies are very discordant despite being considered of adequate quality. For these cases, new studies would be necessary, which should be carried out ideally contemplating all the possible variables and a standardized design. The development of an international guide on stability testing would be very useful for the future comparability of these studies.

In conclusion, we propose a quality graduation and unification tool of stability studies on literature that could be useful to define a consensus stability function in specific conditions that could help laboratories to define stability limits using their own quality specifications.

**Author contributions:** All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

**Research funding:** None declared.

**Employment or leadership:** None declared.

**Honorarium:** None declared.

**Competing interests:** The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

## References

- Guder WG, Fiedler M, daFonseca Wollheim F, Schmitt Y, Töpfer G, Wisser H, et al. The quality of diagnostic samples. 4th completely revised ed. Oxford: BD Diagnostics, 2005.
- Boyanton BL Jr, Blick KE. [Stability studies of twenty-four analytes in human plasma and serum.](#) Clin Chem 2002;48:2242–7.
- Oddoze C, Lombard E, Portugal H. Stability study of 81 analytes in human whole blood, in serum and in plasma. Clin Biochem 2012;45:464–9.
- Cuhadar S. [Stability studies of common biochemical analytes in serum separator tubes with or without gel barrier subjected to various storage conditions.](#) Biochimica Medica 2012;22:202–14.
- Woodrum D, French C, Shamel LB. [Stability of free prostate-specific antigen in serum samples under a variety of sample collection and sample storage conditions.](#) Urology 1996;48(6A Suppl):33–9.
- Stahl M, Brandslund I. [Controlled storage conditions prolong stability of biochemical components in whole blood.](#) Clin Chem Lab Med 2005;43:210–5.
- Garcia R, Li G, Wang Z, Cabanero M, Pincus MR. Stability of electrolyte determinations on the Siemens Advia 1800 analyzer. J Clin Lab Anal 2012;26:372–5.
- Fobker M. [Stability of glucose in plasma with different anticoagulants.](#) Clin Chem Lab Med 2014;52:1057–60.
- Da Rin G, Lippi G. The quality of diagnostic testing may be impaired during shipment of lithium-heparine gel tubes. Clin Chem Lab Med 2014;52:1633–7.
- Frank EA, Shubha MC, D'Souza CJ. Blood glucose determination: plasma or serum? J Clin Lab Anal 2012;26:317–20.
- Li G, Cabanero M, Wang Z, Wang H, Huang T, Alexis H, et al. Comparison of glucose determinations on blood samples collected in three types of tubes. Ann Clin Lab Sci 2013;43:278–84.
- Igawa T, Takehara K, Onita T, Ito K, Sakai H. Stability of [-2]Pro-PSA in whole blood and serum: analysis for optimal measurement conditions. J Clin Lab Anal 2014;28:315–9.
- Monneret D, Godmer A, Le Guen R, Bravetti C, Emeraud C, Marteau A, et al. Stability of routine biochemical analytes in whole blood and plasma from lithium heparin gel tubes during 6-hr storage. J Clin Lab Anal 2016;30:602–9.
- Taylor EC, Sethi B. Stability of 27 biochemistry analytes in storage at a range of temperatures after centrifugation. Br J Biomed Sci 2011;68:147–57.
- Zwart SR, Wolf M, Rogers A, Rodgers S, Gillman PL, Hitchcox K, et al. [Stability of analytes related to clinical chemistry and bone metabolism in blood specimens after delayed processing.](#) Clin Biochem 2009;42:907–10.
- Henriksen LO, Faber NR, Moller MF, Nexø E, Hansen AB. Stability of 35 biochemical and immunological routine tests after 10 hours storage and transport of human whole blood at 21 °C. Scand J Clin Lab Invest 2014;74:603–10.
- Kocijancic M, Cargonja J, Delic-Knezevic A. [Evaluation of BD Vacutainer RST blood collection tube for routine chemistry analytes: clinical significance of differences and stability study.](#) Biochimica Medica 2014;24:368–75.
- Cuhadar S, Koseoglu M, Atay A, Dirican A. [The effect of storage time and freeze-thaw cycles on the stability of serum samples.](#) Biochimica Medica 2013;23:70–7.
- Kang HJ, Jeon SY, Park JS, Yun JY, Kil HN, Hong WK et al. Identification of clinical biomarkers for pre-analytical quality control of blood samples. Biopreservation Biobanking 2013;11:94–100.
- Baruah A, Goyal P, Sinha S, Ramesh KL, Datta R. Delay in specimen processing-major source of preanalytical variation in serum electrolytes. J Clin Diagn Res 2014;8:1–3.
- Harper A, Lu Ch, Sun Y, Garcia R, Rets A, Alexis H, et al. [Reproducibility of serum potassium values in serum from blood samples stored for increasing times prior to centrifugation and analysis.](#) J Clin Lab Anal 2016;30:244–7.
- Shi RZ, Seeley ES, Bowen R, Faix JD. [Rapid blood separation is superior to fluoride for preventing in vitro reductions in measured blood glucose concentration.](#) J Clin Pathol 2009;62:752–3.
- Bonetti G, Cancelli V, Coccoli G, Piccinelli G, Brugnoli D, Caimi L, et al. Which sample tube should be used for routine glucose determination? Prim Care Diabetes 2016;10:227–32.
- Tanner M, Kent N, Smith B, Fletcher S, Lewer M. [Stability of common biochemical analytes in serum gel tubes subjected to various storage temperatures and times pre-centrifugation.](#) Ann Clin Biochem 2008;45:375–9.
- Gambino R, Piscitelli J, Ackattupathil TA, Theriault JL, Andrin RD, Sanfilippo ML, et al. Acidification of blood is superior to sodium fluoride alone as an inhibitor of glycolysis. Clin Chem 2009;55:1019–21.

26. Heins M, Heil W, Withold W. Storage of serum or whole blood samples? Effects of time and temperature on 22 serum analytes. *Eur J Clin Chem Clin Biochem* 1995;33:231–8.
27. Foucher B, Pina G, Desjeux G, Prevosto JM, Chaulet JF, Cheminel V. Influence of temperature and delayed centrifugation: stability studies of 28 analytes currently analysed. *Ann Biol Clin (Paris)* 2005;63:93–100.
28. Dimeski G, Masci PP, Trabi M, Lavin MF, de Jersey J. Evaluation of the Becton-Dickinson rapid serum tube: does it provide a suitable alternative to lithium heparin plasma tubes? *Clin Chem Lab Med* 2010;48:651–7.
29. Smets EM, Dijkstra-Lagemaat JE, Blankenstein MA. Influence of blood collection in plastic vs. Glass evacuated serum-separator tubes on hormone and tumour marker levels. *Clin Chem Lab Med* 2004;42:435–9.
30. Clark S, Youngman LD, Palmer A, Parish S, Peto R, Collins R. [Stability of plasma analytes after delayed separation of whole blood: implications for epidemiological studies.](#) *Int J Epidemiol* 2003;32:125–30.
31. Bush VJ, Janu MR, Bathur F, Wells A, Dasgupta A. [Comparison of BD Vacutainer SST plus tubes with BD SST II plus tubes for common analytes.](#) *Clin Chim Acta* 2001;306:139–43.
32. Donnelly JG, Soldin SJ, Nealon DA, Hicks JM. Stability of twenty-five analytes in human serum at 22 degrees C, 4 degrees C, and –20 degrees C. *Pediatr Pathol Lab Med* 1995;15:869–74.
33. Juricic G, Bakliza A, Saracevic A, Kopcinovic LM, Dobrijevic S, Dmic S, et al. Glucose is stable during prolonged storage in un-centrifuged Greiner tubes with liquid citrate buffer, but not in serum and NaF/KOx tubes. *Clin Chem Lab Med* 2016;54:411–8.
34. Juricic G, Kopcinovic LM, Saracevic A, Bakliza A, Simundic AM. [Liquid citrate acidification introduces significant glucose bias and leads to misclassification of patients with diabetes.](#) *Clin Chem Lab Med* 2016;54:363–71.
35. Brandhorst G, Engelmayer J, Götz S, Oellerich M, von Ahsen N. Pre-analytical effects of different lithium heparin plasma separation tubes in the routine clinical chemistry laboratory. *Clin Chem Lab Med* 2011;49:1473–7.
36. Calatayud O, Tenias JM, Castells X. Agreement in the determination of electrolytes in samples of serum and plasma. Stability of these samples kept for 2 hours at room temperature with uncovered tubes. *Ann Biol Clin (Paris)* 2006;64:157–61.
37. Winter T, Greiser A, Nauck M, Petersmann A. Long-term stability of glucose: 96-h study using Terumo Glycaemia tubes. *Clin Chem Lab Med* 2016;54:407–10.
38. Rossing RG, Foster DM. The stability of clinical chemistry specimens during refrigerated storage for 24 hours. *Am J Clin Pathol* 1980;73:91–5.
39. Zhang DJ, Elswick RK, Greg Miller W, Bailey JL. Effect of serum-clot contact time on clinical chemistry laboratory results. *Clin Chem* 1998;44:1325–33.
40. Woodrum D, York L. [Two years stability of free and total PSA in frozen serum samples.](#) *Urology* 1998;52:247–51.
41. David Seamark D, Backhouse S, Barber P, Hichens J, Salzmann M, Powell R. [Transport and temperature effects on measurement of serum and plasma potassium.](#) *J R Soc Med* 1999;92:339–41.
42. Cartledge JJ, Thompson D, Verril H, Clarkson P, Eardley I. The stability of free and bound prostate-specific antigen. *Br J Urol Int* 1999;84:810–4.
43. Leino A, Koivula MK. [Stability of chemical and immunochemical analytes in uncentrifuged plasma samples.](#) *Ann Clin Biochem* 2009;46:159–61.
44. Joshi AR. Variations in serum glucose, urea, creatinine and sodium and potassium as a consequence of delayed transport/processing of samples and delay in assays. *J Nep Med Assoc* 2006;45:186–9.
45. Jung K, Lein M, Brux B, Sinha P, Schnorr D, Loening SA. [Different stability of free and complexed prostate-specific antigen in serum in relation to specimen handling and storage conditions.](#) *Clin Chem Lab Med* 2000;38:1271–5.
46. Mathew G, Zwart SR, Smith SM. Stability of blood analytes after storage in BD SST tubes for 12 mo. *Clin Biochem* 2009;42:1732–4.
47. Lima-Oliveira G, Lippi G, Salvagno GL, Montagnana M, Gelati M, Volanski W, et al. [Effects of vigorous mixing of blood vacuum tubes on laboratory test results.](#) *Clin Biochem* 2013;46:250–4.
48. Fernandez L, Jee P, Klein MJ, Fischer P, Perkins SL, Brooks SP. A comparison of glucose concentration in paired specimens collected in serum separator and fluoride/potassium oxalate blood collection tubes under survey 'field' conditions. *Clin Biochem* 2013;46:285–8.
49. Devgun MS. Delay in centrifugation and measurement of serum constituent in normal subjects. *Clin Physiol Biochem* 1989;7:189–97.
50. O'Keane MP, Cunningham SK. [Evaluation of three different specimen types \(serum, plasma lithium heparin and serum gel separator\) for analysis of certain analytes: clinical significance of differences in results and efficiency in use.](#) *Clin Chem Lab Med* 2006;44:662–8.
51. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Sue Kirkman M. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Diabetes Care* 2011;34:e61–99.
52. Biological Variability database. Analytical Quality Committee. Spanish Society of Laboratory Medicine (SEQ ML). [http://www.seq.es/es/comisiones/comision-de-calidad-analitica/\\_id:4/](http://www.seq.es/es/comisiones/comision-de-calidad-analitica/_id:4/).
53. Blázquez R, Prada E, Ricós C, Gutiérrez-Bassini G, Morancho J, Salas Á, et al. Quality analytical specifications obtained by consensus through intercomparison programs AEFA/AEBM, SEQ y SEHH. *Rev Calid Asist* 2015;30:341–3.
54. UNE-EN ISO 15189:2012. Medical laboratories – Particular requirements for quality and competence. [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=42641](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=42641).
55. CLSI. Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline. CLSI document EP25-A. Wayne, PA: Clinical and Laboratory Standards Institute, 2009.
56. Asirvatham JR, Moses V, Bjornson L. [Errors in potassium measurement: a laboratory perspective for the clinician.](#) *North Am J Med Sci* 2013;5:255–9.
57. Oosterhuis WP, Bayat H, Armbruster D, Coskun A, Freeman KP, Kallner A, et al. [The use of error and methods in the medical laboratory.](#) *Clin Chem Lab Med* 2018;56:209–19.
58. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53:833–5.