Opinion Paper

Michael Cornes*, Ana-Maria Simundic, Janne Cadamuro, Seán J. Costelloe, Geoffrey Baird, Gunn B. B. Kristensen, Alexander von Meyer, Mads Nybo and Rubén Gómez Rioja, on behalf of the Working Group for the Preanalytical Phase (WG-PRE) of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM)

The CRESS checklist for reporting stability studies: on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group for the Preanalytical Phase (WG-PRE)

https://doi.org/10.1515/cclm-2020-0061
Received January 17, 2020; accepted June 24, 2020; published online July 25, 2020

Abstract: To ensure that clinical laboratories produce results that are both accurate and of clinical utility it is essential that only samples of adequate quality are analysed. Although various studies and databases assessing the stability of analytes in different settings do exist, guidance on how to perform and report stability studies is lacking. This results in studies that often do not report essential information, thus compromising transferability of the data. The aim of this manuscript is to describe the **C**hecklist for **R**eporting **S**tability **S**tudies (CRESS) against which future studies should be reported to ensure standardisation of reporting and easy assessment of transferability of studies to other healthcare settings. The EFLM WG-PRE (European Federation of Clinical Chemistry and Laboratory Medicine Working Group for the Preanalytical Phase) produced the CRESS checklist following a detailed literature review and extensive discussions resulting in consensus agreement. The checklist consists of 20 items covering all the aspects that should be considered when producing a report on a stability study including details of what should be included for each item and a rationale as to why. Adherence to the CRESS checklist will ensure that studies are reported in a transparent and replicable way. This will allow other laboratories to assess whether published data meet the stability criteria required in their own particular healthcare scenario. The EFLM WG-PRE encourage researchers and authors to use the CRESS checklist as a guide to planning stability studies and to produce standardised reporting of future stability studies.

Keywords: checklist; preanalytical; stability.

***Corresponding author: Michael Cornes**, Clinical Chemistry Department, Worcester Acute Hospitals NHS Trust, Worcester, UK, Phone: +44 1905 760843, E-mail: Michael.cornes@nhs.net
Ana-Maria Simundic: Department for Medical Laboratory Diagnostics, Clinical Hospital "Sveti Duh"; Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia
Janne Cadamuro: Department of Laboratory Medicine, Paracelsus Medical University, Salzburg, Austria
Seán J. Costelloe: Department of Clinical Biochemistry, Cork University Hospital, Wilton, Cork, Republic of Ireland. https://orcid.org/0000-0002-2464-3735
Geoffrey Baird: Department of Laboratory Medicine, University of Washington, Seattle, Washington, USA. https://orcid.org/0000-0001-9064-6558
Gunn B. B. Kristensen: Norwegian Quality Improvement of Laboratory Examinations (Noklus), Haraldsplass Deaconess Hospital, Bergen, Norway
Alexander von Meyer: Institute for Laboratory Medicine and Microbiology, Munich Municipal Clinic Group, Munich, Germany
Mads Nybo: Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense, Denmark
Rubén Gómez Rioja Laboratory Medicine Department, La Paz University Hospital, Madrid, Spain. https://orcid.org/0000-0002-3429-0427

Introduction

To ensure that laboratories only analyse samples of sufficient quality it is essential that stability in a variety of different storage and transport conditions and specimen matrices is known for each analyte. Although various studies and databases looking at the stability of analytes in

different settings do exist, guidance on how to perform and report stability studies is lacking [1–3]. This results in a lack of standardisation in the output of stability studies and raises concerns about the quality of data they produce. Often, essential information in studies is missing, thus compromising the applicability, transferability and reproducibility of study results and conclusions.

A good stability study should ideally allow the calculation of the percentage difference (PD%) between the baseline result and a series of results measured at some later time points. This can then be used to display data as an instability equation and obtain the stability limit (SL), defined as the time period within which the loss of stability of a quantity exceeds the specified maximum permissible difference (MPD), i.e. the maximum allowable deviation from the baseline for a given clinical setting [1, 4]. The instability equation is a function of time and will be explained in more details later in the text.

A basic design would involve obtaining multiple samples from a single blood draw (to allow replicates), under optimal preanalytical conditions, in a minimum of 10 different participants [4]. The number of participants must be sufficient to ensure the entire analytical range is covered for the analyte(s) in question. One aliquot per patient is analysed under basal conditions and the remaining tubes/aliquots are analysed after predefined progressive time periods have elapsed under controlled storage conditions. The loss of stability is measured as the difference in measurand concentration between the basal sample and each of the successive test samples. Preferably, the analysis should be run in the same analytical batch to avoid analytical systematic error, and at least in duplicate, preferably in triplicate, to reduce the effect of imprecision [1, 4, 5].

In this type of experiment, the variation of analyte concentration is studied with respect to a single variable, time, to observe the impact on stability. The other variables, known or assumed to affect stability (e.g. temperature, contact with cells, type of container, contact with air, or exposure to light), can be studied by performing separate experiments for each combination of factors. This combination of factors will be the storage condition defined in each case.

There are numerous sources of variability that may affect the study results. Due to the complexity of factors contributing to sample stability it is difficult to generalise a study design that would be applicable across different healthcare settings and populations. However, with sufficient information, appropriately presented in a standardised way, data from different studies could potentially be transferable to a given laboratory scenario or combined to produce a more powerful resource.

The aim of the EFLM WG-PRE (European Federation of Clinical Chemistry and Laboratory Medicine Working Group for the Preanalytical Phase) in this manuscript is to provide a checklist against which future studies should be reported to ensure standardisation of reporting and easy assessment of transferability of studies to other healthcare settings. The CRESS checklist was produced by members of the working group following a detailed literature review by all authors and repeated rounds of extensive discussions resulting in consensus agreement. The checklist follows the recommendations of the European Federation for Laboratory Medicine (EFLM) Biological Variation Working Group (BVWG) [6] and STARD (Standard for Reporting of Diagnostic Accuracy Guidelines), both of which have an aim of improving the quality of reporting the studies in their respective areas [7, 8].

The checklist

The checklist (Table 1) has been produced by EFLM WG-PRE after detailed literature review, extensive discussions and consensus agreement of the working group members. The checklist consists of 20 items covering all aspects of a stability study. The importance and rationale behind each item of the checklist is discussed below.

1. Title/keywords

The title is the first place many readers look and will determine whether or not they will invest further time in reading the main article. It is essential that the title clearly indicates that the article describes a stability study. It should include details of the analytes evaluated, specimen matrices investigated, temperature and the time period over which stability is assessed.

Keywords must be well chosen to ensure that other laboratories searching for stability studies on the same analyte should be able to easily locate the study in the majority of search media. As a minimum the keywords should include:

- Stability
- The name of the analyte(s) if <3 analytes
- Preanalytical
- Sample type

2. Abstract

The abstract needs to expand on the title and must state that the study is a stability study and what the aims of the

Table 1: The CRESS **C**hecklist for **R**eporting **S**tability **S**tudies.

Section	Number	Item
Title/keywords	1.	– Clearly indicate in the title that content is a stability study, – Define in the title what analytes are being evaluated in what matrix, at which temperature, over what period of time, – Use recommended keywords: stability, the name of the analyte(s) if <3 analytes, preanalytical and sample type (we recommend that all of them are used to properly identify a stability study).
Abstract	2.	– State that the study is a stability study, – Clearly define what the aims of the study are, – Include an overview of the study design, – Define which analytes are being examined, in which matrices, in which tube types, at which temperature and over what period of time, as a minimum, – Cover headline results and the key findings of the study, as a minimum, – Draw appropriate conclusions; do not generalize beyond the study results.
Introduction	3.	– Discuss the importance of sample stability, – Provide a thorough overview of the current position and existing studies, – State the knowledge gap being filled.
Aim	4.	– Clearly define the aims of the study – Provide in details how the study will address these aims, – State the analytes, matrix, temperature and period of time involved.
Materials and methods	5.	– State where the study was done, – State the start date and the end date of the study – Thoroughly describe materials and methods to a depth that allows the study to be replicated, – Provide sufficient detail to allow transferability to different healthcare settings.
Measurand/ analyte	6.	– Define the analyte(s) involved in the study.
Samples	7.	– Define the tube type, manufacturer and additive, – Define the sample matrix (e.g. whole blood, serum, plasma, urine, faeces, saliva, cerebrospinal fluid etc), – Define sample volume, – State if samples were specifically collected for the study or leftover samples have been used, – State whether individual participants or pools were used, – If pools have been used, provide the protocol and appropriate reference, according to which pool has been prepared.
Participants	8.	– Define the patient population used including gender, age, geographical location, comorbidities, relevant medication, health status and ethnicity as a minimum. – State how this information was obtained and verified
Preanalytical conditions	9.	– State how samples were collected (i.e. method of phlebotomy), – State how tubes were mixed, – State sample transport conditions (mode of transport, temperature, time), – Specify the exact time and all relevant conditions (temperature, centrifuge conditions, mechanical forces, etc.) the sample was exposed to from the sample collection to the first time point (time zero) of the stability experiment: – Specify time elapsed from the sample collection to the time of separation of the serum or plasma from cells, – If samples were stored prior to time zero, specify sample storage conditions, – If samples were frozen, please state how samples were thawed, – State how temperatures were monitored throughout the process (measurement device, measurement method, measurement time points, device calibration status, etc.), – State any additional preanalytical sources of biases that may have been introduced.
Analytical procedure	10.	– State the method of analysis, analytical platform, relevant details of the reaction including any deviations from manufacturers recommendations, – Define assay traceability to international standards, – State the number of replicate analyses, – State within and between batch variation (CVA%) at the concentration level close to concentrations studied in the stability experiment (state if data are from the manufacturer or from in-house verification study),

Table 1: (continued)

Section	Number	Item
		– State a protocol (provide a reference) according to which CVA (%) was obtained (if in-house data are provided), – Define whether analysis was done as a single batch or multiple batches – If samples were run in multiple batches, state relevant conditions between batches and how bias between batches was controlled.
Spiking studies	11.	– Define how (if any) spiking studies were performed.
Length of study	12.	– State the duration over which stability of the analytes in each sample will be measured in the study and the frequency of time points
Storage conditions	13.	– Clearly define the storage conditions during the experiment and how they were recorded and controlled, – State all relevant variables described in detail under Sections 7 and 9.
Statistical analysis	14.	– Include details of how the sample size and the number of replicates required was justified and how this minimizes the effect of analytical imprecision, – State whether the mean or median of replicate samples was used, – Define the instability equation, – State what statistical tools were and rationale behind their use.
Outliers	15.	– State how many outliers were identified – Define what tools were used to remove outliers and why
Acceptability criteria	16.	– State MPD with evidence for its selection from referenced studies, – Follow the Milan hierarchy to defining analytical performance specifications [22].
Results	17.	– Provide textual, tabulated and graphical presentation of the data, – Plot results on a scatter graph with PD% on the ordinate and the time with appropriate units on the abscissa, – Present data using consistent terminology throughout the manuscript, – Provide raw data in an excel format.
Discussion	18.	– State how the study addresses the original aims, – Demonstrate how it can be applied (or not) to other healthcare systems, – Define any limitations, particularly those that prevent transferability of data.
Funding	19.	– State any funding sources whether financial or in the form of consumables,
Ethics	20.	– State if the study has been granted ethical approval, – State if patient consent (if applicable) was obtained.

study are. It needs to include an overview of how the study is designed, which defines which analytes are being examined, in which matrices, in which tube types, at which temperature and over what period of time. It should summarise the key findings with some headline data and draw appropriate conclusions, without going beyond the results of the study.

3. Introduction

The introduction should define why it is important to understand the stability of analytes under various different conditions and state where in the diagnostic pathology process this information is necessary. It must thoroughly review the current state of knowledge for the analyte(s) being studied in the manuscript. It should provide a summary of what previous studies show and then define the knowledge gap that will be filled by the proposed study.

4. Aims

The introduction should end by clearly defining what the study aims to show, along with brief details about how this will be achieved. The aim should include the analytes, specimen matrices, temperature and the time period over which the stability study is done.

5. Materials and methods

The methodology used in the study should be described in sufficient detail to allow other laboratories to clearly understand what has been done, enable them to reproduce the study if required and allow them to understand how it fits with their own healthcare setting. Authors should state when and where the study has been done (start date and the end date of the study) to allow other laboratories to understand the time of year and allow for any seasonal influences.

6. Measurand/analyte

The analyte(s) must be clearly defined and must remove any possible ambiguity. For example NT-proBNP must not be confused with BNP. Where possible these should follow standard international terminology. Where analytes are non-standard or not commonly used sufficient details should be provided to allow other laboratories to easily replicate the work.

7. Samples

Stability of an analyte largely depends on the sample matrix. For most analytes, plasma or serum is the matrix with the best stability, whilst most analytes are usually much less stable in other matrices like saliva, urine and cerebrospinal fluid [9]. That is why it is important to define the sample matrix, e.g. whether the stability of the analyte was studied in whole blood, serum, plasma, or other matrices (e.g. urine, faeces, saliva, cerebrospinal fluid, etc.).

The sample container type, manufacturer and additive for primary samples and all aliquots used in the study should be described.

It is most preferable to use real patient specimens with analyte levels covering the analytical range. If real patient samples are used, then it is important to state if the samples used were from individual participants or patient/subject pools. If stability was studied on patient pools, details of how the patient pools were prepared and how many samples were included should be stated, including details of any guidelines or protocols followed with references. Deviations from guidelines and/or protocols should be stated. Authors should also state whether the samples were specifically collected for the study in question or if leftover samples following routine analysis were used.

8. Participants

The population in which the analyte is being studied must be clearly defined as this may have an impact on transferability, e.g. if inpatient or outpatient, renal patients or intensive care patients etc. were included. Demographic information should also be stated, to include as a minimum gender, age, geographical location, comorbidities, relevant medications, health status and ethnicity.

Where applicable, details of how the health status, comorbidities and medications information was obtained should be included, i.e. were they self-declared by participants or independently verified. If verified, details of how this was done should be included.

9. Preanalytical conditions (conditions prior to the basal time point)

For each analyte, there should be a definition of the conditions the baseline sample was subjected to, from the time the sample was obtained until the baseline measurement is performed. This should follow, but not be limited to, the SPREC (Standard Preanalytical Code) criteria for defining the main pre-analytical factors that could impact on a sample's integrity [10].

The method of phlebotomy is an important preanalytical step which needs to be described and properly referenced in the manuscript. Authors should state which guidelines were followed and if there was any deviation from this or manufacturer guidelines. EFLM WG-PRE recommends that venous blood sampling is done in full compliance with the EFLM-COLABIOCLI Recommendations for venous blood sampling [11]. Whether the blood sample was venous, collected via a vascular catheter or another form of sampling should also be included, as well as how and for how long the tubes were mixed. Authors should also provide information on the length of sample storage before the baseline sample was analysed, how long the contact time with cellular components was, and when the sample was separated prior to storage as this can significantly affect sample stability for some analytes [10, 12–14]. Additional information including whether the tubes were kept open or closed during the experiment and whether the sample was exposed to light should be recorded and reported in the manuscript.

Sample transport conditions should be described and include as a minimum, details on the transit time, temperature, exposure to mechanical forces e.g. if samples have been transported via pneumatic tube systems [15].

Authors should provide details about centrifugation (e.g. time, g force, temperature) and storage conditions (temperature, air, light) throughout the process, as sample exposure to air is detrimental to the stability of some analytes (e.g. bicarbonate) [16]. If samples were stored frozen for a certain period of time, the thawing procedure may affect sample integrity, at least for some analytes (e.g. coagulation parameters) [17]. For this reason, if samples were stored frozen, authors should provide the detailed description about how samples were thawed and the number of freeze thaw cycles.

Additionally, temperature for a particular storage or handling period or any other part of the process throughout the study should be reported as mean, minimum, and maximum temperature. How temperatures were monitored

throughout the process should be described in detail and include, but not be limited to, the following: measurement device, measurement method, measurement time points, device calibration status.

Any exclusion criteria and potential sources of bias introduced should be stated. This should include whether haemolysis, icterus, and lipaemia were excluded and if so, how the measurement of these indices was performed and what cut-offs were used, together with evidence for these thresholds [18, 19].

10. Analytical procedure

The methods must clearly define the platform on which each analysis was performed and which exact methodology (assay) was used. Any deviations from manufacturer recommendations should be explicitly stated and the rationale behind this provided. The traceability of the method to international standards should be defined, and within and between batch performance (i.e. analytical coefficient of variation, CV_A%) of the assay for the duration of the study should be stated at relevant levels including how this information was derived, i.e. was it from the manufacturer kit inserts or in house experiments. If it was derived from in-house experiments, details of the protocol followed should be stated and the protocol referenced. Details of the reagent, calibrators, and IQC (internal quality control) used including lot numbers for all and a statement on recent external quality performance should be stated.

Methods should also state whether analysis was completed in a single analytical batch, or in multiple batches. Single batch studies reduce the analytical bias due to method drift or recalibration but add additional time to the storage of samples, which could impact on the results [4]. It also requires the introduction of a preservation process without deterioration (freezing at –70 °C for example), which does not always exist or is possible for a specific analyte/sample matrix and may be itself a stability variable. If samples were run in multiple batches, the authors should state all relevant conditions between batches that may have influenced the results of the experiment and how bias between these batches was controlled.

Finally, in this section, the authors should state the number of replicate analyses for each individual sample and the number of samples at each time point.

11. Spiking studies

Where spiking studies are used, there should be a clear description of how these were performed, which – if any – guidelines were followed and how results were calculated [20]. Spiking studies add additional potential sources of error and bias due to the need to open and manipulate the sample. Moreover, the spiking material may behave differently to the naturally occurring analyte in a given matrix. To avoid any potential source of bias due to sample spiking, it is preferable to use patient specimens with high analyte levels, rather than spiked samples wherever possible.

12. Length of study

The actual duration over which stability of the analytes in each sample will be measured should be documented along with the time points at which analysis will occur. Justification for time points at which analytes will be measured must be stated.

13. Storage conditions

As for the storage of samples prior to measurement of analyte concentration at the basal time point, sample storage conditions for the duration of the stability experiment should be documented in line with all requirements listed under points 7 and 9.

14. Statistical analysis

The methodology behind data analysis must be stated and justified. Provide a justification for the sample size and numbers of replicates required, and confirm that this minimizes the effect of analytical imprecision [21].

The instability equation should be expressed as follows:

$$\mathrm{PD\%} = a \times \mathrm{time} + b \times \mathrm{time}^2 + c \times \mathrm{time}^3 + \cdots + n \times \mathrm{time}^n$$

This complex equation will fit to non-linear behaviour observed for some analytes [4, 22], but is difficult to implement. A first-order equation could fit properly for many equations. Expression of simplified equation should be:

$$\mathrm{PD\%} = a \times \mathrm{time}$$

where PD% is percentage difference between the baseline result and a result measured at a particular time point and given by '*a*', the slope of the equation, multiplied by the time. The instability equation does not have an intercept, as there is no difference at the starting point (time zero) i.e. the intercept would be zero. Reporting an instability equation in addition to a graph facilitates the

Table 2: Example of stability data presentation: stability of glucose in whole blood at room temperature (22 °C) in closed serum tube (manufacturer XX), protected from light.

Definitions			Results (PD% [95% CI])						
Test	Number of participants	Baseline mean value (range)	1 h	2 h	3 h	4 h	5 h	6 h	Instability equation
Glucose, mmol/L	10	5.98 (3.84–9.48)	−2.32% (−1.61%, −3.02%)	−3.71% (−2.8%, −4.62%)	−5.64% (−4.46%, −6.81%)	−7.23% (−6.04%, −8.42%)	−9.07% (−7.46%, −10.67%)	−10.99% (−9.29%, −12.67%)	PD%=−1,834 × time (h) r=0.961, p<0.001

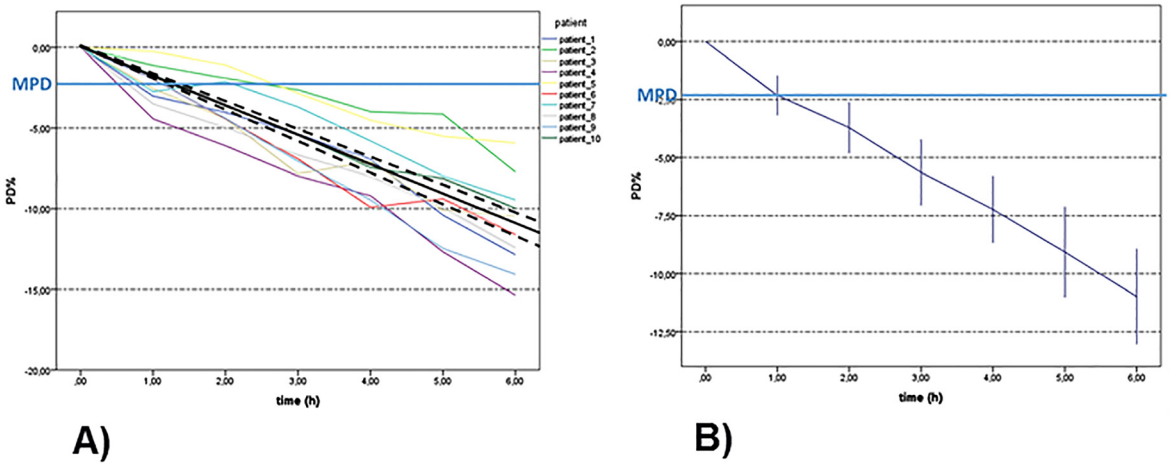

Figure 1: Example of stability data presentation: stability of glucose in whole blood at room temperature (22 °C) in closed serum tube (manufacturer XX), protected from light.
(A) Instability equation calculation (black line) using the least squares adjustment with confidence intervals (dotted lines) for the slope. All patient data is also shown. (B) Stability equation calculation using the point to point estimation with confidence intervals for the mean of patients at every study time. Blue line presents the MPD (maximum permissible difference) for serum glucose based on biological variation (which was 2.34% at the time when the study was performed).

interchangeability of information between laboratories to calculate the stability limit in relation to their MPD specifications based on their performance characteristics to meet their clinical requirements.

There should be sufficient data to estimate confidence intervals for the instability equation or the instability estimation in a given time for each analyte. Any statistical tools used to calculate statistical significance should be justified with details as to why that tool was selected.

15. Outliers

If outliers were observed, the statistical tools used to identify these should be described including justification for their use. Outliers should be looked for in both the samples and in the replicates of the samples. They should not be discarded but should be recorded in the final document. If outliers were removed or not used in any data produced from the study this should also be stated.

16. Performance acceptability criteria

The establishment of the MPD must be very clearly defined. It is critical that future readers are able to clearly understand how the MPD was determined and what the justification was for its value. The MPD should ideally be traceable to published data and should follow the Milan hierarchy for analytical performance. Therefore, ideally the MPD should be derived, in order of preference, from outcome data, biological variation, or state-of-the-art in line with the Milan hierarchy [23]. Whatever the source, justification of why the selected

Table 3: Example of stability data presentation: Raw data for individual participants at each time point demonstrating the stability of glucose in whole blood at room temperature (22 °C) in closed serum tube (manufacturer X), protected from light.

Glucose, mmol/L		**Basal**	**1 h**	**2 h**	**3 h**	**4 h**	**5 h**	**6 h**
Patient 1	Replicate 1	4.41	4.35	4.3	4.25	4.15	4.03	3.92
	Replicate 2	4.53	4.32	4.28	4.21	4.17	3.98	3.87
	Average	4.47	4.34	4.29	4.23	4.16	4.01	3.90
	CV%	1.90%	0.49%	0.33%	0.67%	0.34%	0.88%	0.91%
	PD% from basal		−3.02%	−4.03%	−5.37%	−6.94%	−10.40%	−12.86%
Patient 2	Replicate 1	6.32	6.21	6.18	6.14	6.05	5.98	5.71
	Replicate 2	6.25	6.22	6.15	6.1	6.02	6.07	5.89
	Average	6.29	6.22	6.17	6.12	6.04	6.03	5.80
	CV%	0.79%	0.11%	0.34%	0.46%	0.35%	1.06%	2.19%
	PD% from basal		−1.11%	−1.91%	−2.63%	−3.98%	−4.14%	−7.72%
Patient 3	Replicate 1	5.32	5.15	5.07	4.97	4.98	4.74	4.81
	Replicate 2	5.28	5.18	5.06	4.8	4.87	4.79	4.68
	Average	5.30	5.17	5.07	4.89	4.93	4.77	4.75
	CV%	0.53%	0.41%	0.14%	2.46%	1.58%	0.74%	1.94%
	PD% from basal		−2.55%	−4.43%	−7.83%	−7.08%	−10.09%	−10.47%
Patient 4	Replicate 1	7.5	7.1	7.02	6.87	6.7	6.54	6.37
	Replicate 2	7.41	7.15	6.98	6.85	6.84	6.48	6.25
	Average	7.46	7.13	7.00	6.86	6.77	6.51	6.31
	CV%	0.85%	0.50%	0.40%	0.21%	1.46%	0.65%	1.34%
	PD% fron basal		−4.43%	−6.10%	−7.98%	−9.19%	−12.68%	−15.36%
Patient 5	Replicate 1	6	5.9	5.87	5.84	5.74	5.64	5.66
	Replicate 2	5.98	6.05	5.98	5.8	5.7	5.68	5.61
	Average	5.99	5.98	5.93	5.82	5.72	5.66	5.64
	CV%	0.24%	1.78%	1.31%	0.49%	0.49%	0.50%	0.63%
	PD% from basal		−0.25%	−1.09%	−2.84%	−4.51%	−5.51%	−5.93%
Patient 6	Replicate 1	3.87	3.74	3.65	3.54	3.47	3.45	3.4
	Replicate 2	3.8	3.78	3.68	3.6	3.44	3.5	3.38
	Average	3.84	3.76	3.67	3.57	3.46	3.48	3.39
	CV%	1.29%	0.75%	0.58%	1.19%	0.61%	1.02%	0.42%
	PD% from basal		−1.96%	−4.43%	−6.91%	−9.91%	−9.39%	−11.60%
Patient 7	Replicate 1	9.5	9.25	9.27	9.15	8.87	8.74	8.62
	Replicate 2	9.45	9.18	9.28	9.1	8.98	8.7	8.54
	Average	9.48	9.22	9.28	9.13	8.93	8.72	8.58
	CV%	0.37%	0.54%	0.08%	0.39%	0.87%	0.32%	0.66%
	PD% from basal		−2.74%	−2.11%	−3.69%	−5.80%	−7.97%	−9.45%
Patient 8	Replicate 1	4.39	4.3	4.26	4.12	4.05	4	3.87
	Replicate 2	4.47	4.25	4.16	4.15	4.1	3.98	3.89
	Average	4.43	4.28	4.21	4.14	4.08	3.99	3.88
	CV%	1.28%	0.83%	1.68%	0.51%	0.87%	0.35%	0.36%
	PD% from basal		−3.50%	−4.97%	−6.66%	−8.01%	−9.93%	−12.42%
Patient 9	Replicate 1	7.1	7.02	6.87	6.57	6.42	6.23	6.17
	Replicate 2	7.18	6.98	6.78	6.7	6.51	6.27	6.1
	Average	7.14	7.00	6.83	6.64	6.47	6.25	6.14
	CV%	0.79%	0.40%	0.93%	1.39%	0.98%	0.45%	0.81%
	PD% from basal		−1.96%	−4.41%	−7.07%	−9.45%	−12.46%	−14.08%
Patient 10	Replicate 1	5.37	5.3	5.15	5.04	4.97	4.98	4.78
	Replicate 2	5.35	5.24	5.18	5.1	4.95	4.87	4.87
	Average	5.36	5.27	5.17	5.07	4.96	4.93	4.83
	CV%	0.26%	0.81%	0.41%	0.84%	0.29%	1.58%	1.32%
	PD% from basal		−1.68%	−3.64%	−5.41%	−7.46%	−8.12%	−9.98%
Global	**Average PD%**		−2.32%	−3.71%	−5.64%	−7.23%	−9.07%	−10.99%
	CI 95%		(−1.61%, −3.02%)	(−2.8%, −4.62%)	(−4.46%, −6.81%)	(−6.04%, −8.42%)	(−7.46%, −10.67%)	(−9.29%, −12.67%)

PD%, percentage deviations.

MPD is fit for purpose for the analyte and population in question should be given.

17. Results

Results should be presented textually, in tabulated forms and graphically. The instability equation should always be shown in conjunction with the results from which it was derived. Data should be consistently presented throughout the manuscript using consistent terminology and units (e.g. S.I.). Data interpretation should be easy, allowing key results to be quickly identified and understood. Where replicates have been performed, averages should be calculated using means for duplicates or medians in the case of triplicates, which should then be used for calculation of PD%.

Results should be plotted on a scatter graph with PD% on the ordinate and the time with appropriate units on the abscissa. Examples of data presentation are shown in Table 2 and Figure 1. As an indication of inter-individual variability, Figure 1A represents each patient studied with a different marker or line whereas Figure 1B gives confidence intervals of the mean PD%. Preferably, both types of graph should be provided, given the slightly different information they display and therefore the easy visual information they provide discussed below. Representation of individual plots are useful to identify possible cases of inter-individual variability, since an individual equation of loss of stability could be very different from the overall trend. The relationship of analytical variation with analyte concentration can be assessed by looking at individual patient data. Other inter-individual variation factors, such as cell number or genetic differences in metabolic pathways must be suspected on participants with different relationship patterns. These may require manual adjustment of the instability equation on the stability graph or, preferably, use of a least squares adjustment methodology. The regression equation must pass through the origin of coordinates since at time 0 the specimen is by definition unchanged. Note that the best fit to the measured concentration change over time may not be linear. Confidence intervals for defined time points should be plotted. From the MPD it is possible to extrapolate the SL from the instability equation and graphs. Raw data should also be available, ideally as an Excel file presented as supplemental data. An example format for the raw data is shown in Table 3, which provides percentage deviations (PD%) for each patient and average PD% and its confidence interval (1.96 × standard error of the mean [SEM]).

The time point defined as the SL must be the last time point at which the analyte has been proven to be stable. As its behaviour is not known beyond measured time points, the instability equation should not be extrapolated outside the range of data included in the stability experiment.

18. Discussion

The discussion must discuss not just what is shown by the data but how it addresses the original aims of the study. It should also discuss how the data can be applied to other healthcare settings. Perhaps most importantly it should state what the limitations are and how these will affect transferability to other healthcare systems.

19. Funding

Details of how the study was financially supported should be included. This should incorporate details of any support in terms of provided consumables such as blood collection tubes, instruments, reagents, controls and calibrators.

20. Ethics

Details of any ethical approval obtained should be included. This should state who this was obtained from and whether patient consent was required/obtained.

Discussion

The CRESS checklist provides a list of essential items that should be reported in future stability studies. Adherence will ensure that a study is reported in a transparent way and that it can be easily replicated.

Understanding of the stability of biomarkers is important to ensure that laboratories produce results that have the highest possible accuracy. There is currently a lack of a definition of what should be included in a study of sample stability to ensure that data produced is of a high quality, is transparent, and is reported in such a way as to make the results reproducible and applicable to other laboratories. Of the stability studies found in the literature, many are contradictory or do not have sufficient data to allow transferability to different healthcare systems [2, 3]. In order to address this, the EFLM WG-PRE set out to establish a checklist against which future studies should be reported. It is the ambition of the EFLM WG-PRE that the CRESS checklist is used to guide

transparent and standardised reporting of all future stability studies. This would allow other laboratories to assess whether the CRESS-compliant published data meets the stability criteria required in their own particular healthcare scenario. Laboratories would then be able to input certain criteria to find an instability equation that is immediately applicable to their own requirements.

The ultimate aim is that stability studies will improve as a result of this checklist in much the same way as diagnostic accuracy studies have improved since the introduction of STARD. This should then lead to a future where all stability studies are based on the same high quality standardised data.

Conclusions

The EFLM WG-PRE encourages researchers and authors to use the CRESS checklist as a guide to produce standardised reporting of future stability studies. We recommend consulting this checklist as part of the planning phase of a stability study to stimulate thought about what is required to produce the publication at the end of the study. We also encourage reviewers and journal Editors to use the CRESS checklist as a guide to assess and evaluate the way stability studies are reported and to ensure standardised reporting of future stability studies. Finally, we encourage all readers to look for important details and pieces of information listed in the CRESS checklist when reading stability studies in order to evaluate its transferability to their particular healthcare setting and local conditions.

Research funding: None declared.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: Authors state no conflict of interest.

References

1. Guder W, Fiedler GM, daFonseca-Wollheim F, Schmitt Y, Töper G, Wisser H, et al., editors. The quality of diagnostic samples. 4th completely revised ed. Oxford: BD Diagnostics; 2015.
2. 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.
3. Cuhadar S, Atay A, Koseoglu M, Dirican A, Hur A. Stability studies of common biochemical analytes in serum separator tubes with or without gel barrier subjected to various storage conditions. Biochem Med 2012;22:202–1.
4. Gómez Rioja R, Martínez Espartosa D, Segovia M, Ibarz M, Llopis MA, Bauça JM, et al. Laboratory sample stability. Is it possible a consensus stability function? An example of 5 blood magnitudes. Clin Chem Lab Med 2018;56:1806–18.
5. Lumsden JH. Laboratory test method validation. Revue Méd Vét 2000;151:623–30.
6. Bartlett WA, Braga F, Carobene A, Coşkun A, Prusa R, Fernandez-Calle P, et al. A checklist for critical appraisal of studies of biological variation; biological variation working group, European federation of clinical Chemistry and laboratory medicine (EFLM). Clin Chem Lab Med 2015;53:879–85.
7. STARD guidelines. Available from http://www.equator-network.org/reporting-guidelines/stard/ [Accessed October 2018].
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. Stard 2015: an updated list of essential items for reporting diagnostic accuracy studies. Clin Chem 2015;61:1446–52.
9. Barranco T, Rubio CP, Tvarijonaviciute A, Rubio M, Damia E, Lamy E, et al. Changes of salivary biomarkers under different storage conditions: effects of temperature and length of storage. Biochem Med (Zagreb) 2019;29:010706.
10. https://www.isber.org/page/SPREC [Accessed March 2019].
11. Simundic AM, Bölenius K, Cadamuro J, Church S, Cornes MP, van Dongen-Lases EC, et al. Working group for preanalytical phase (WG-PRE), of the European federation of clinical Chemistry and laboratory medicine (EFLM) and Latin American working group for preanalytical phase (WG-PRE-LATAM) of the Latin America confederation of clinical biochemistry (COLABIOCLI). Joint EFLM-COLABIOCLI recommendation for venous blood sampling. Clin Chem Lab Med 2018;56:2015–38.
12. Abraham RA, Agrawal PK, Acharya R, Sarna A, Ramesh S, Johnston R, et al. Effect of temperature and time delay in centrifugation on stability of select biomarkers of nutrition and non-communicable diseases in blood samples. Biochem Med (Zagreb) 2019;29:020708.
13. Lippi G, Salvagno GL, Lampus S, Danese E, Gelati M, Bovo C, et al. Impact of blood cell counts and volumes on glucose concentration in uncentrifuged serum and lithium-heparin blood tubes. Clin Chem Lab Med 2018;56:2125–31.
14. Juricic G, Bakliza A, Saracevic A, Kopcinovic LM, Dobrijevic S, Drmic S, Simundic AM. 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.
15. Mullins GR, Harrison JH, Bruns DE. Smartphones can monitor medical center pneumatic tube systems. Clin Chem 2016;62:891–3.
16. Nonkes LJP, van Gelder MK, Kemperman H, Abrahams AC, Boereboom FTJ, Ten Berg MJ, Gerritsen KGF. Improving home haemodialysis: stability evaluation of routine clinical chemistry analytes in blood samples of haemodialysis patients. Biochem Med (Zagreb) 2019;29:010709.
17. Bronić A, Coen Herak D, Margetić S, Milić M. Croatian Society of Medical Biochemistry and Laboratory Medicine: National recommendations for blood collection, processing, performance and reporting of results for coagulation screening assays prothrombin time, activated partial thromboplastin time, thrombin time, fibrinogen and D-dimer. Biochem Med (Zagreb) 2019;29:020503.
18. Farrell CJ, Carter AC. Serum indices: managing assay interference. Ann Clin Biochem 2016;53:527–38.
19. von Meyer A, Cadamuro J, Lippi G, Simundic AM. Call for more transparency in manufacturers declarations on serum indices: on

behalf of the working group for preanalytical phase (WG-PRE), European federation of clinical Chemistry and laboratory medicine (EFLM). Clin Chim Acta 2018;484:328–32.
20. Clinical Laboratory Standards Institute. Establishing and verifying and extended measuring interval through specimen dilution and spiking, 1st ed.; 2018. CLSI Document EP34 Approved August.
21. Jones SR1, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J 2003;20:453–8.
22. Pum JKW. Evaluating sample stability in the clinical laboratory with the help of linear and non-linear regression analysis. Clin Chem Lab Med 2020;58:188–96.
23. 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.